Augmentative Communication Device
Hardware Development

BME 300
Department of Biomedical Engineering
University of Wisconsin-Madison
December 5, 2009

Prachi Agarwal, BWIG
Erin Devine, Communicator
Brian Mogen, Team Leader
Steve Wyche, BSAC

Client
Dr. Lawrence Kaplan, M.D
Department of Pediatrics, UW School of Medicine and Public Health

Advisor
Professor Thomas Yen, Professor
Department of Biomedical Engineering

Table of Contents

Abstract 3

Background Information 3

Existing Devices 5

Problem Statement 6

Design Requirements 7

Previous Work 8

Design Alternatives 10

Design Matrices 11

Final Design 14

Testing 15

Future Work 18

Limitations 20

Works Cited 21

Abstract

The goal of this project is to develop a novel method of communication for individuals suffering from a wide range of communicative disorders. Currently, there are no devices on the market that are fast, inexpensive, aesthetic, and allow the user to add emotion to their communication. We present work already accomplished, several design alternatives for hardware development, and a comparative weighted matrix to match our product design specifications. We chose to pursue a rechargeable compression driver-based platform to be carried on the hip, or in a pack.

Background Information

Neuromotor dysfunction presents itself in a number of forms, one of the most common being cerebral palsy. This occurs in approximately 1 out of every 500 people and is a result of abnormalities in the growth and functioning of the brain[1]. This leads to uncontrollable reflex movements and moderate to severe muscle tightness. Cerebral palsy can be caused by head trauma after birth, but this is relatively rare. It is more common for the brain to be affected before or during birth.

Four main types of brain damage contribute to the majority of cerebral palsy cases[2]. The first is periventricular leukomalacia, which is damage to the white matter of the brain. This is usually responsible for transmitting signals throughout the brain and body, but small holes in this white matter that form before birth do not allow this to develop properly. Another cause of cerebral palsy is cerebral dysgenesis, or abnormal development of the brain. During the first 20 weeks of development, the fetal brain is very vulnerable. Any interruption in the growth of the brain causes abnormalities that interfere with the transmission of signals. Mutations in genes, infections, fevers, or trauma could contribute to this interruption. Intracranial hemorrhage, or bleeding in the brain, is also a possibility. If blood flow is blocked by blood clots in the placenta, the baby may suffer a stroke, leading to blocked or broken vessels in the brain. The final key development malfunction is hypoxic-ischemic encephalopathy, or intrapartum asphyxia. More commonly referred to simply as asphyxia, this is a lack of oxygen in the brain. Tissue in the brain, most notably in the cerebral motor cortex, can be destroyed, and this causes cerebral palsy.

Motor functions are affected differently in everyone; some have a slight limp, while others are completely wheelchair-bound. Those with spastic hemiplegia are mostly affected in the arms and hands; those with spastic diplegia are more affected in the legs and feet[3]. The most severe form is spastic quadriplegia, where one has severe stiffness in the limbs, is usually completely wheelchair-bound, and has extreme difficulties speaking. Cerebral palsy is a non-progressive disorder, meaning the disease will not worsen, but later physiological disabilities are very common. Many of these neurological conditions cause disarthia which impairs a person's ability to speak[6].

Laryngeal cancer is not as common as cerebral palsy, only affecting about 1 in 22,666 people. This is a disease in which malignant cancer cells develop in the tissues of the larynx[4]. Sometimes called the "voice box", the larynx houses the vocal cords and is found just below the pharynx in the neck. When a person attempts to speak, the vocal cords vibrate as air moves against them, producing sound. This sound echoes through the person's mouth and nose to create a voice. The cancer typically develops in the squamous cells, which line the inside of the larynx. To treat the laryngeal cancer, many people choose to have a laryngectomy. In this procedure, the larynx is surgically removed. Since the vocal cords are located within the larynx, these are also removed. This causes the affected individual to lose all speech capabilities.

Finally, speech can be inhibited with a paralyzed diaphragm. The diaphragm is vital to normal respiration. In normal speech, air is required to be pushed up and out of the body. The diaphragm serves the purpose of pushing this air. When the diaphragm is paralyzed, the individual is not able to produce sufficient air flow to generate normal speech[7].

Existing Devices

The devices currently available on the market today consist mostly of touch screen tablet PCs or handheld devices. These devices have pre-programmed common phrases and keyboards to enter in custom sentences. In order to give the user the ability to speak more quickly, the devices usually prompt possibilities for the next letter or word, but typing what they want to say is still a laborious task and is frustrating because it is time consuming.

One such device is the Tango. It is geared towards children and uses images and icons to direct the child to what they want to say. This device, however, is limited only to pre-programmed phrases. While useful for kids who can't speak, this device stifles a child's creativity in that they cannot produce their own sentences. It does not allow children to expand their vocabulary and limits them to what is programmed into the device.

Another device on the market is the Dasher. This device uses some sort of pointer, whether it is a joystick, a mouse type apparatus or a slider, to point out letters to form words. The program prompts the user with possible and common letters to follow the first in order to speed up the input process. While this is a great interface for someone that cannot use a standard keyboard to type, this is still very slow and makes the user less likely to say something that would be time consuming. In our client's experience, people then tend to limit their speech and vocabulary to the minimum that is required for what they want to communicate. In this way, devices like this limit self-expression in addition to being unable to add inflection or emphasis to the words.

Another manufacturer, DynaVox, produces touch screen devices that have some pre-programmed common phrases and a keypad input system. This interface is also slow, cumbersome, and lacks the ability to add emotion. While these devices come in small handheld versions or full size tablet PC based on the preference of the user, they are still limiting because they are slow and users encounter the same frustration as they do with other devices. Users pay a lot of money for these expensive devices, and then tend to not use them because they are not an efficient means of communication.

Problem Statement

As a continuing project this semester, our goal is to work from our "proof of premise" prototype to develop a smaller, more compact and convenient device that people with speech difficulties will find easy and intuitive to use. We hope to integrate the different hardware components into our own circuit and package it in an aesthetically pleasing way.

Design Requirements

One of the biggest complaints from patients with communicative disorders is that the devices out there to help them speak are slow and lack the ability to add emotion to what the user wants to express. This delay, between when the user thinks a phrase and when they are able to actually communicate, can make the user feel unintelligent or that they are being perceived as unintelligent. It can also leave them out of a conversation since they cannot produce language within the normal pause of a conversation. In addition, they are unable to add emphasis or inflection to what they want to say. These are the issues that our client, Dr. Lawrence Kaplan, has asked us to address. Every day he encounters patients that are frustrated with the means of communication to which they are limited. Many of his patients give up trying to speak and let others do it for them. Our client is looking to break away from the conventional communicative devices that are on the market today. He would like us to come up with something new that allows the user to have more spontaneity when they speak, as well as the ability to demonstrate emotions such as irritation or excitement when they communicate. By doing this, we can hopefully “bridge the gap” for people with communicative disorders and help them to communicate in a way that feels more natural and comfortable.

The main feature that our client is looking for in the device is that it be phonetics-based instead of text-based. With sounds at their disposal instead of words, people have a wider variety of things they can say. They can put more personalization into their speech and incorporate slang. Working from this idea, we hope to make the device fast and intuitive so that the user feels that they are better able to express themselves, as well as feeling more a part of conversation.

Previous Work

This semester's project is a continuation of work completed last semester. Last semester, we established the principal of the idea we are working from now. We used musical hardware in order to generate a "proof or premise" prototype in order to test the principal behind the idea. The hardware used included a Kaossilator as an input source and a talkbox as a mechanical sound source.

The premise of our idea is that sound can be funneled into the mouth and shaped by the mouth much like naturally produced sound is. By using the generated sound the user does not need to produce sound., which can be a problem for individuals with certain speaking difficulties The device replicates the function of the diaphragm and vocal chords by funneling sound into the mouth. This way, the user does not have to use their vocal chords or produce the air force necessary to work the vocal chords. One stipulation we knew would be necessary is that the user has control over their facial muscles so that they have the ability to shape the sound into understandable language.

The device works by taking an electronic input signal and converting it to mechanical sound. For our prototype the input came from the Kaossilator, using a setting that had a setting with vowel sounds. This input was converted by a compression driver in the talkbox to mechanical sound and funneled into the mouth by a vinyl tube. This sound is then shaped into language as the facial muscles modulate the frequencies generated by the talkbox.

Testing of the device showed that the understandability of the speech was best when the tube was placed about a centimeter in the mouth. This way the teeth could be used to occlude the sounds and create consonant sounds. It was found that the tube diameter did not affect the sound quality for the testing that we performed. We determined that a small tube would be best for the future design.

Inflection and emotion were also important factors in proving the principal of our idea. We tested the ability of others to understand whether the expression was a statement, exclamation or question, It was determined that the average rate of understandability was 91.11%, on the second iteration of the expression. This result can be seen in Figure 1.

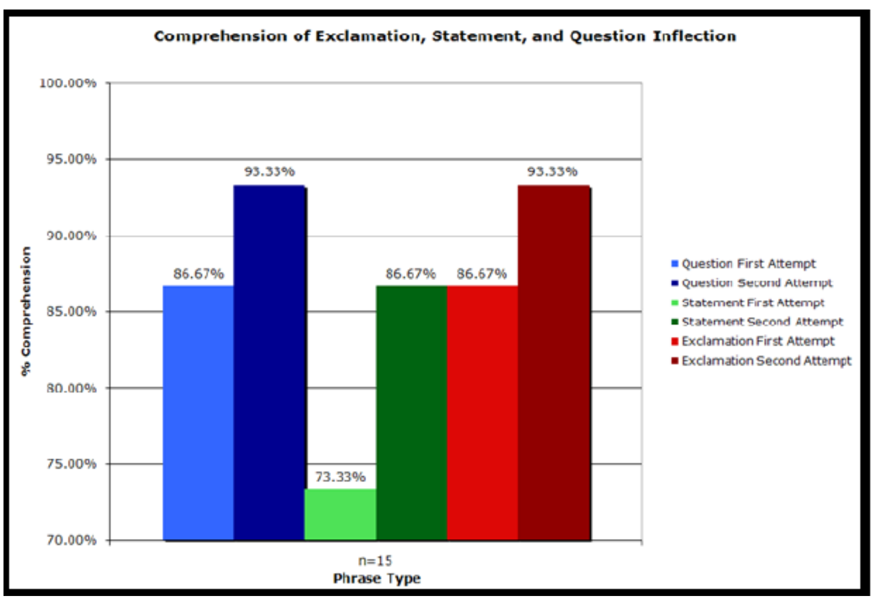

Figure 1. The inflection was understood 91.11% of the time on the second iteration of the expression.

The previous work on this project from last semester has laid the groundwork for the direction we will take this semester. We hope to integrate the hardware into a small, more compact device that is easy and comfortable for the user. Our previous work and testing has demonstrated that the principal behind the idea of this device is valid.

Design Alternatives

There were three aspects of the design we felt were important and required evaluation: the method of attachment, how to power the device, and the type of sound source to use.

Method of Attachment

There were several locations in which it would be feasible to carry around a device. There were a number of positions that needed to be evaluated including on the head as a cap, over the ear, around the neck, on the hip, integrated into a pair of glasses, and on the lapel. These locations were evaluated in a design matrix.

Power Supply

There are three options for powering the device. First, a rechargeable battery could be employed. Second, a set of standard disposable batteries could provide the required voltage. Finally, the device could be powered by whatever input platform is being used.

Sound Source

Finally, there is the source of sound to be used. Most professional sound is created by electromagnetic compression drivers, which provide the best sound quality. Other potential sound sources include miniature piezo-electric speakers, or small speakers found in standard computer speakers or boom boxes.

Design Matrices

Attachment

The design matrix was split into three different aspects, as they could all be independently evaluated. The attachment matrix is evaluated on four categories (Table 1). The first and most important criterion is the aesthetic nature of the device. In designing this device the needs of the end user must be kept foremost. The current market discriminates against users both visually and aurally. Devices both sound unnatural and look out of place. Allowing users to fit into society is a chief concern. The ear, cap and hip designs scored the best in this category, as they are the most easily integrated, or socially acceptable. The cap design can be covered by any standard hat or scarf while the hip design visually mimics a blood glucose monitor and the ear design mirrors a Bluetooth headset.

The second important criterion is the user convenience. For preliminary design there will be wiring connecting the hand control to the device, and tubing carrying sound to the user's mouth. Allowing the user to move freely is critical for user's who may already have motor limitations. The cap design only has one connection down the neck leaving it as the freest device. All other designs have more two connections coming from the user's head thus limiting head and arm movement. Ease of manufacture was also considered, as time is limited

Factor	Weight	Neck	Ear	Cap	Hip	Glasses	Lapel
User Convenience	30%	4	2	7	5	2	6
Aesthetics	40%	3	8	9	8.5	1	4
Ease of Manufacture	20%	7	2	6	8	3	8
Cost	10%	8	4	6	8	3	8
TOTAL	10	4.6	4.6	**8.7**	7.3	1.9	5.8

Table 1. Attachment design matrix shows how designs were evaluated on how socially acceptable the device is (aesthetics), how feasible the device is to manufacture (ease of manufacture), and cost.

to one semester for prototyping the device. Finally cost was given as a final category of evaluation. It was weighted lightly as all of these design alternatives will fall an order of magnitude below any device currently on the market.

Power Supply

The power supply is evaluated based on five criteria (Table 2). The most important factors were convenience and life cycle. Lithium Ion rechargeable batteries scored the highest in lifecycle, as they do not need to be replaced. Disposable batteries can provide the same amount of power as rechargeable Li/Ion batteries, however when they die they must be removed and replaced daily. Convenience was the second important criterion. It would be easy for the user if the device were powered by the input hardware however this would decrease the lifecycle of the device. Because it is assumed that the user will be recharging the input hardware daily, it would be easy for them to recharge the sound hardware at the same time. Size and safety were scored less critically as all possible designs were relatively equal. All

Factor	Weight	Li/Ion	Disposable	Input Device
Size	20%	8	5	10
Safety	10%	7	7	9
Life Cycle	30%	8	2	4
Convenience	30%	8	4	8
Cost	10%	8	3	8
TOTAL	10	**7.9**	3.8	7.3

Table 2. Power supply design matrix shows how designs were evaluated on functional size, user safety, length of battery life (life cycle), ease of use (convenience), and cost.

batteries have the potential to leak; however, if powered by the input device, there would only be one site of leakage thus explaining the slightly higher score in the Input Device category. For the life cycle, user convenience, and size, Li/Ion batteries were selected as the design to pursue.

Sound Source

The last aspect of the design to be evaluated was the sound source. Important factors addressed were size, frequency response, volume and cost (Table 3). Size was critical in creating a portable device that would be minimally visible and minimally invasive. The compression driver being evaluated had a largest dimension of 1”, so all designs scored well. Frequency response was the also critical. The compression driver provided the largest frequency band in the human range, while the speaker's is slightly smaller and the piezo-electric speaker only has a single pitch. As all of these designs were inexpensive compared to existing devices cost was not an important deciding factor and could be used to justify testing multiple sound sources. Finally the output volume of the sound source is critical, as it would be beneficial for the user to be heard in social situations without requiring and external amplifier and microphone system.

Factor	Weight	Compression Driver	Piezo –Electric	Miniature Speaker
Size	30%	7	9	9
Freq. Response	30%	9	1	8
Cost	10%	9	9	9
Volume	30%	9	2	8
TOTAL	10	**8.4**	4.5	**8.4**

Table 3. Sound source design matrix shows how designs were evaluated on the size of sound source, whether the source can provide the human sound frequencies needed, cost, and output volume of the sound source.

Final Design

Our final hardware design consists of two parts: an amplification system and a compression driver (Figure 2). First, we input the human sound into the filter. This human sound is a pre recorded sound with a continuous 'aa' syllable. The sound source produces low amplitude signals which get attenuated when passed through the tube of the compression driver. Also, with a low voltage it is not possible to achieve resonant frequencies to occur at medium range frequencies. Hence we first tested the compression driver for different voltages and frequencies in order to produce a reasonable output. While testing it was found that there is need of at least a gain of 15 to produce an audible sound.

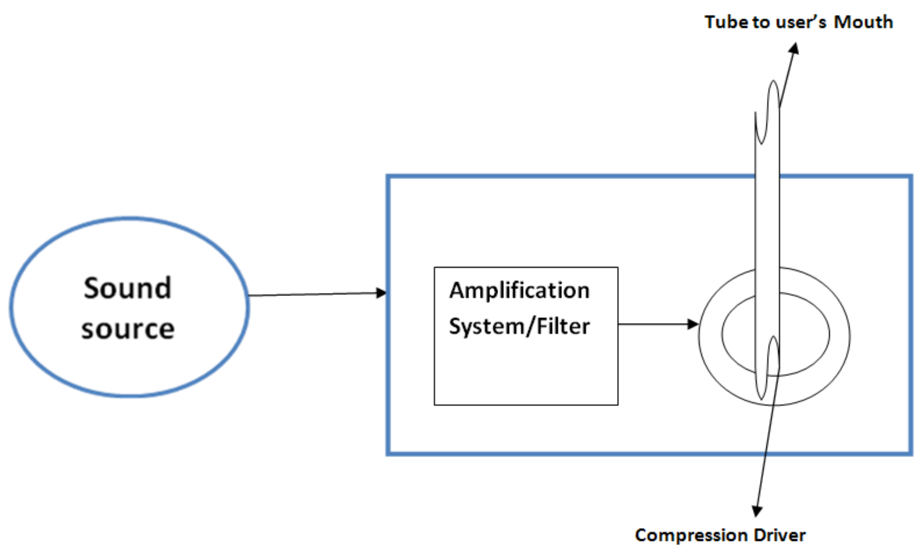

Figure 2: Our hardware comprised of the amplification system and compression driver.

For our search of an amplifier, we looked for something that would not produce too much noise, give us a reasonable gain, and have a good heat sink. Professor Mark Allie of Electrical Engineering works with acoustics. When we consulted him regarding the amplifier he provided us with one which was quite compatible with our compression driver. It had a gain of 21 with a reasonable heat sink.

After testing for different voltage and frequencies, the sound files were saved in an iPod, and this was then fed into the amplifier through headphones. The amplifier was powered using two 6 V batteries which provided a +/- 12 V power supply with ground connection. The set up took about 8 mA of current to run smoothly. The arrangement of the design has been shown in Figure 2. The amplified signal is fed into the compression driver which sends high pressure sounds through the tube into the user's mouth. These sounds are then converted into words via lip sync. This concept is very similar to the way humans produce sound. Using this device, in order to speak one only needs to use his facial muscles. The rest is done by our device.

Testing

When testing our device this semester, we were looking to answer a few specific questions: What frequencies and volumes are best understood? What letters are difficult to understand? How understandable is the device for first time listeners? The first test we did was what frequency and voltage inputs were the best understood. We did this by increasing the voltage by 0.1 volts and found the frequency that was most understandable at each voltage. We found that frequency and voltage have a linear relationship and that the sound was best understood between 2.7 and 4.2 V and 600 Hz and 1.2 kHz. A plot of our testing data can be seen below in Figure 3.

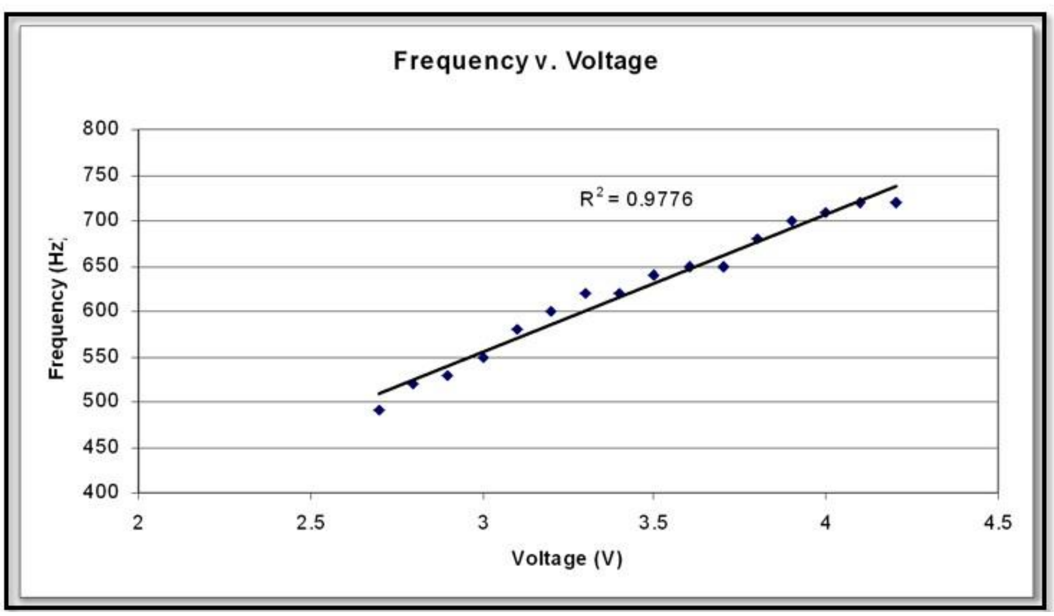

Figure 3: A plot of the frequency and voltage relationship for speech that is best understood.

After determining the best input, we tested the understandability. We first tested which letters of the alphabet were difficult to understand with our device. We sat down with a panel of 10 people who had never listened to our device before, and they listened to a user reciting the alphabet. We asked them to cross out the letters that they found difficult to understand. The letters C, D, and E were found to be difficult to differentiate. Other higher frequency letters were difficult as well, such as T, G, P, and Z. Finally, T and D were difficult to understand because they require the same mouth movement so they come out sounding very similar. The data for this testing can be seen below in Figure 4.

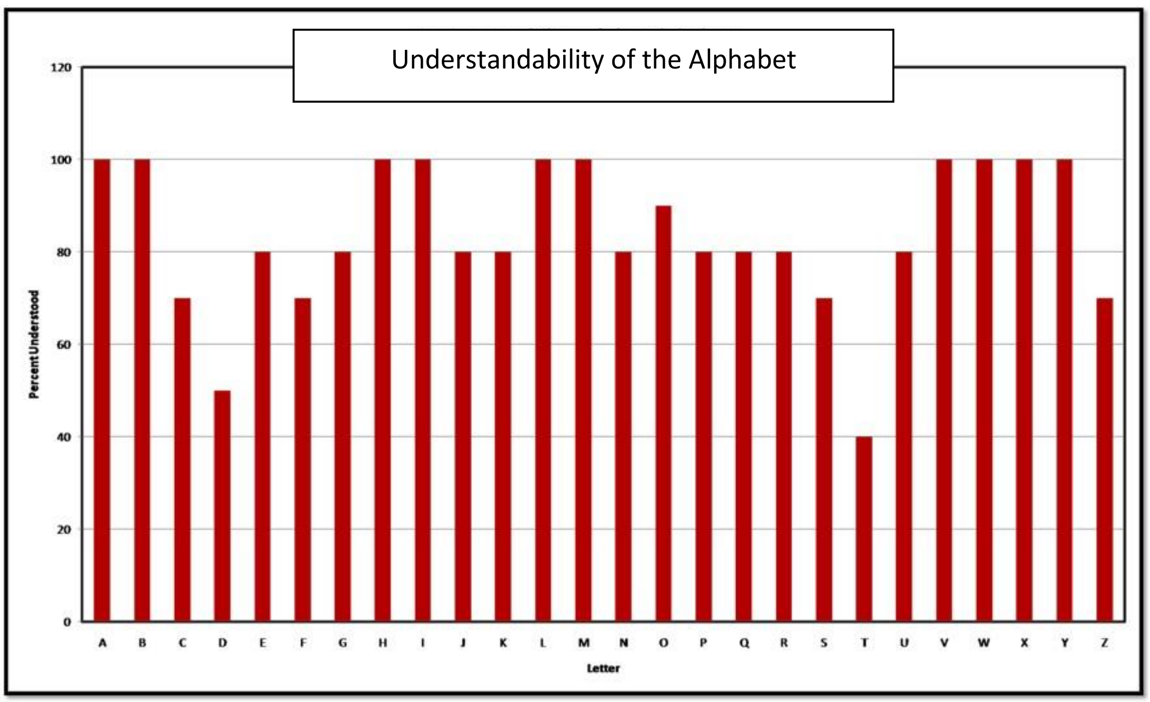

Figure 4: A bar graph of the percentage each letter was understood among the ten listeners.

Our final test sought to determine how understandable phrases were to people who had never heard the device before (Figure 5). We had the same 10 people sit down and listen to the user recite 5 phrases. We found that the worst phrase, phrase 3, was still understood 70%

of the time. The overall average understandability of the device was 88%. This result is very encouraging. One factor that will be easily fixed with a touch pad sound source is that when the sound ceases in between words, they will be easier to distinguish. We were using a continuous sound source so it was difficult at times to distinguish when one word ended and another began. We also hope that understandability will be improved as we improve the sound source because there will be less noise. In addition, we hope to be able to correct problems producing percussive constant sounds using this device with the input source we develop.

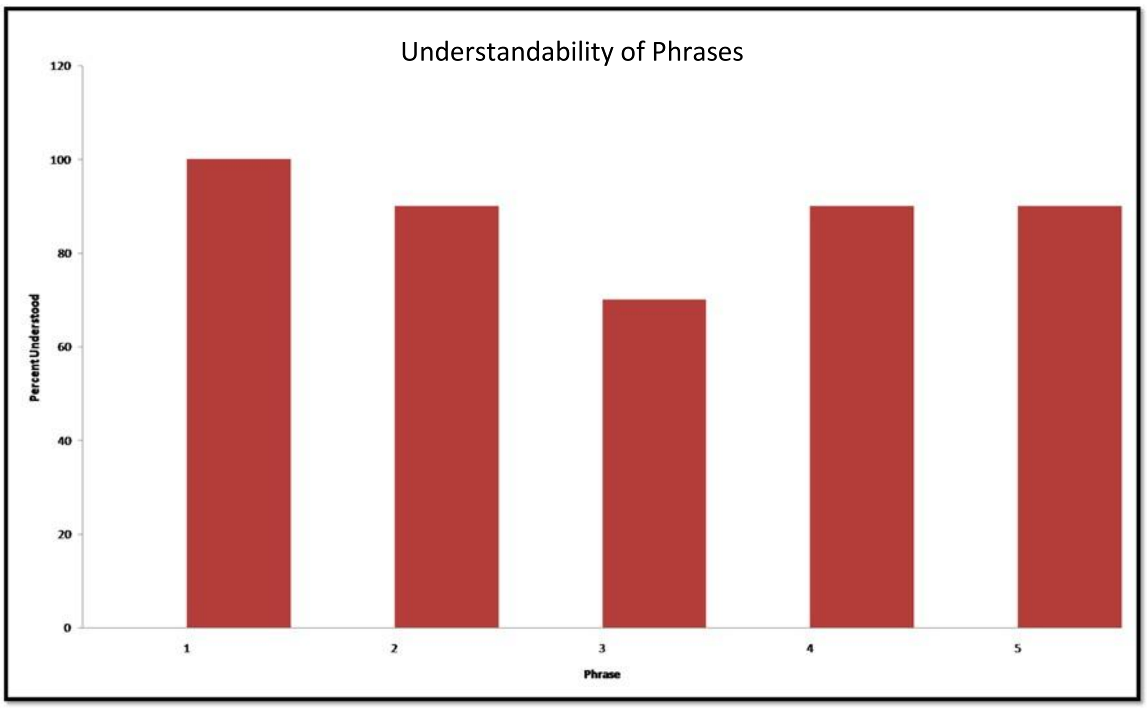

Figure 5. The phrases were understood 88% of the time by new listeners.

Future Work

The focus of next semester will be the design of a new input sound source. This will likely be in the form of an iPhone/iTouch application that will create a more personalized sound. It will use a pre-recorded voice at different pitches that the user can play by pressing different areas on the touch pad. With this, the user will be able to stop the sound simply by removing his/her finger. In this manner, it will be possible to have better control over when the device plays the sound, and this leads to clearer output since the words will be more enunciated. Additionally, the varying pitches and volumes will allow the user to express his/her emotions, as well as making it clear whether the phrase is intended to be a statement, question, or exclamation. This feature will allow the user to engage in more natural conversation.

We also need to refine the hardware. A new connection between the compression driver and the tube may prevent any air from leaking out, providing a greater pressure to push the signal to the user's mouth. We also need to test various tube lengths to determine the best combination of comfort and function. In addition, a new amplifier will need to be constructed in a similar manner to the one we used this semester. All these pieces will need to be connected in the most concise manner possible, and this will all be packaged in one portable container, such as a PCB case. The end product will be compact and commercially usable.

In order to ensure the product is functional for individuals with actual speaking disorders, it will be important to set up a study where a group of people with applicable speaking disorders each test the device. We have begun the process of gaining IRB approval for this, and Dr. Kaplan has a number of willing patients who would try the device. This will give us great insight into the limitations our device may have, as well as the disorders for which it works the best. Finally, if this goes well, we will complete the patent process and present our product to various corporations.

Limitations

Since the user is required to lip sync in order to effectively use this device, damaged facial muscles and nerves that are involved in speech would inhibit the user from using our device.

The orbicularis oris muscle which surrounds the lip area is the main muscle involved in speech. Any damage to that muscle would prevent the user from using the device. Damage to other muscles like the buccinator and zygomaticus major could also limit one from effectively using this device. However, some words could still be produced in spite of damage to these two muscles.

Out of the 12 cranial nerves that are present in a human, the hypoglossal nerve is involved in articulation of speech and innervates the muscles associated with it. Hence, any damage to it could also prevent the user from using the device.

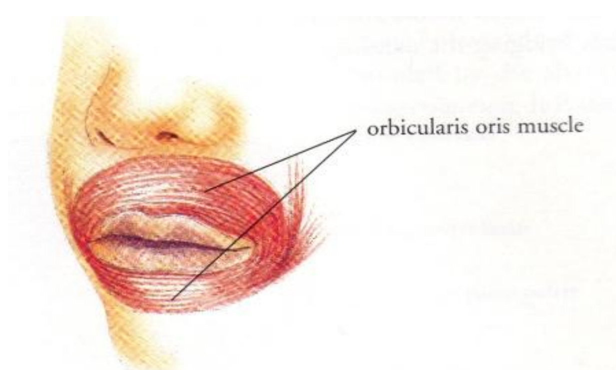

Figure 4. The orbicularis oris muscle[5].

Works Cited

[1]EMedicine. "Cerebral Palsy." http://emedicine.medscape.com/article/1179555-overview

[2]National Institute of Neurological Disorders and Stroke. "Cerebral Palsy: Hope Through Resarch." http://www.ninds.nihngov/disorders/cerebral_palsy/detail_cerebral_palsy.htm#88033104

[3]Gait Analysis Laboratory. "Cerebral Palsy Program/Guide." http://gait.aidi.udel.edu/gaitlab/cpGuide.html

[4] National Cancer Institute. "Larnygeal Cancer Treatment." http://www.nci.nih.gov/cancertopics/pdq/treatment/laryngeal/patient/

[5] Image: http://72.233.24.94/~worldcf/wp-content/uploads/2008/12/f3.jpg

[6] R. Patel. Phonatory control in adults with cerebral palsy and severe dysarthria. *Augmentative and Alternative Communication*; 2002, Vol. 18, No. 1, Pages 2-10.

[7] A Waller, F Dennis, J Brodie, A Y Cairns. "Evaluating the use of TalksBac, a predictive communication device for nonfluent adults with aphasia." *International Journal of Language & Communication Disorders*; 1998, Vol. 33, No. 1, Pages 45-70.