REAL-TIME CHANGE IN PROSODIC ASPECTS OF TEXT GENERATED SPEECH

by

NURE BIANE KASSAS

ANTHONY BUHR, COMMITTEE CHAIR
MARCUS BROWN
PRISCILLA DAVIS
MARCIA HAY-MCCUTCHEON

A THESIS

Submitted in partial fulfillment of the requirements for the degree of Masters of Science in the Department of Communicative Disorders in the Graduate School of The University of Alabama

TUSCALOOSA, ALABAMA

2014
ABSTRACT

The purpose of this study was to investigate the feasibility and effectiveness of a text-to-speech (TTS) device that allows the user to manipulate pitch and volume as speech is being generated. This device was intended to facilitate the communicative needs of individuals with complex communication needs (CCN) as a result of acquired neurological conditions such as dysarthria. An Android touchscreen tablet with a built-in speech engine was used as the hardware for the TTS device, and a post-audio signal processing approach was utilized to program the TTS device. Results were collected in two separate phases: auditory and use-based. During the auditory phase, participants listened to audio samples from the thesis TTS device, a typical TTS device, and human speech and then rated them based on perceived affect (positive vs. negative) or intent (question vs. statements) categories. During the use-based phase, participants provided feedback about the thesis TTS device after using it to communicate with the study investigator. Although auditory phase results indicated that the thesis device was currently not as effective as human speech when communicating emotion and intent, use-based findings were more promising. Use-based results revealed that the new features the thesis TTS provided (ability to manipulate pitch and volume) were considered beneficial.
DEDICATION

This thesis is dedicated to all of the individuals with complex communication needs who inspired me to complete this project.
# LIST OF ABBREVIATIONS AND SYMBOLS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAC</td>
<td>Augmentative and alternative communication</td>
</tr>
<tr>
<td>ALS</td>
<td>Amyotrophic lateral sclerosis</td>
</tr>
<tr>
<td>CCN</td>
<td>Complex communication needs</td>
</tr>
<tr>
<td>df</td>
<td>Degrees of freedom: values in the final calculation of a statistic that are free to vary</td>
</tr>
<tr>
<td>HMM</td>
<td>Hidden Markov Models</td>
</tr>
<tr>
<td>LMN</td>
<td>Lower motor neuron</td>
</tr>
<tr>
<td>N</td>
<td>Negative</td>
</tr>
<tr>
<td>NS</td>
<td>Non-significant</td>
</tr>
<tr>
<td>P</td>
<td>Positive</td>
</tr>
<tr>
<td>POSM</td>
<td>Patient operated selector mechanism</td>
</tr>
<tr>
<td>Q</td>
<td>Question</td>
</tr>
<tr>
<td>S</td>
<td>Statement</td>
</tr>
<tr>
<td>SAPI</td>
<td>Speech application programming interface</td>
</tr>
<tr>
<td>SGD</td>
<td>Speech generating device</td>
</tr>
<tr>
<td>TBI</td>
<td>Traumatic brain injury</td>
</tr>
<tr>
<td>TTS</td>
<td>Text to speech</td>
</tr>
<tr>
<td>UMN</td>
<td>Upper motor neuron</td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>Chi-square: Test that assesses the goodness of fit between observed and expected values</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

I would like to express my gratitude to Dr. Anthony Buhr for his continued support and invaluable expertise provided throughout this research project. I was privileged to have him as my thesis advisor. I would also like to thank my committee members, Dr. Marcus Brown, Dr. Priscilla Davis, and Dr. Marcia Hay-McCutcheon for their time and encouragement and for being exceedingly nice, as well as the faculty at the University of Alabama who helped get me to this point.

My appreciation also goes to Jess Stough for his collaboration during the development phase of the study and for all the “pro-jamming” sessions we had to produce a working device. Thank you to all the friends and family members who kept me calm and reminded me to have fun by getting me to take study breaks from time to time. I feel blessed to have had so many wonderful people cheer me on while I was completing my thesis. I am also grateful that God showed me what it felt like to struggle to communicate effectively when I was younger and was first learning to speak in Arabic because those experiences helped get me where I am today.
# CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS AND SYMBOLS</td>
<td>iv</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
</tr>
<tr>
<td>I. BACKGROUND</td>
<td>1</td>
</tr>
<tr>
<td>Augmentative and Alternative Communication (AAC)</td>
<td>1</td>
</tr>
<tr>
<td>A. Evolution of AAC</td>
<td>1</td>
</tr>
<tr>
<td>1. 1960s and 1970s</td>
<td>2</td>
</tr>
<tr>
<td>2. Mid 1970s and 1980s</td>
<td>3</td>
</tr>
<tr>
<td>3. 1980s and 1990s</td>
<td>4</td>
</tr>
<tr>
<td>B. Speech Synthesizing Technology</td>
<td>5</td>
</tr>
<tr>
<td>1. Aspects of Speech</td>
<td>5</td>
</tr>
<tr>
<td>2. Speech Synthesis</td>
<td>6</td>
</tr>
<tr>
<td>a. Concatenative Synthesis</td>
<td>6</td>
</tr>
<tr>
<td>b. Formant Synthesis</td>
<td>7</td>
</tr>
<tr>
<td>c. HMM-based Speech Synthesis</td>
<td>7</td>
</tr>
<tr>
<td>C. Disorders Requiring AAC</td>
<td>7</td>
</tr>
<tr>
<td>1. Impact of Dysarthria on the Motor Speech System</td>
<td>9</td>
</tr>
<tr>
<td>a. Amyotrophic Lateral Sclerosis (ALS)</td>
<td>9</td>
</tr>
<tr>
<td>b. Traumatic Brain Injury (TBI)</td>
<td>10</td>
</tr>
<tr>
<td>2. Speech Behavioral Effects of Dysarthria</td>
<td>10</td>
</tr>
<tr>
<td>3. Types of Dysarthria</td>
<td>11</td>
</tr>
<tr>
<td>a. Flaccid Dysarthria</td>
<td>11</td>
</tr>
</tbody>
</table>
4. Overall Usefulness ............................................................. 31

V. DISCUSSION ........................................................................... 32
A. Auditory Data ................................................................. 32
B. Use-based Data ............................................................... 32
C. Prosodic Characteristics of Speech ..................................... 33
D. Future Directions ............................................................. 34

VI. CONCLUSION ...................................................................... 36
REFERENCES ............................................................................ 37
APPENDIX A ............................................................................. 40
APPENDIX B ............................................................................. 42
APPENDIX C ............................................................................. 44
APPENDIX D ............................................................................. 46
APPENDIX E ............................................................................. 47
APPENDIX F ............................................................................. 53
APPENDIX G ............................................................................. 55
LIST OF TABLES

1. Question vs. Statement Results ................................................................. 27
2. Positive vs. Negative Affect Result .......................................................... 28
3. Statistical Analysis Results ........................................................................ 28
# LIST OF FIGURES

1. POSM Device ........................................................................................................... 2
2. AAC Device Operated through Electromuscular Signals ............................... 2
3. Communication Board Attached to a Wheelchair ........................................... 3
4. Primitive SGD ........................................................................................................ 4
5. Upper and Lower Motor Neuron Pathway – Generic Example ................. 8
6. Stick of Gum Analogy ............................................................................................ 20
7. Resampling Analogy ............................................................................................ 21
8. Signal Processing Analogy .................................................................................. 21
9. Signal Processing and Resampling of a Wave – Generic Example ............ 22
10. Use Case Diagram ............................................................................................... 42
I. BACKGROUND

Augmentative and Alternative Communication (AAC)

Individuals with congenital or acquired motor deficits that interfere with speech production have complex communication needs (CCN). Such individuals require methods other than oral speech to help them communicate with other people in their environment. Augmentative and alternative communication (AAC) is any mode of communication aside from oral speech that facilitates purposeful exchanges between two or more people and includes both unaided communication systems (e.g., body language, gestures, and/or sign language) and aided communication systems (e.g., man-made mechanisms such as picture boards or speech generating devices; ASHA, *Augmentative and Alternative Communication*).

A. Evolution of AAC

To appreciate how AAC devices aid non-oral communication for individuals with CCN, it is first helpful to consider how AAC devices have developed over the last few decades to meet those communication needs. The field of AAC emerged due to research on primitive aided communication devices. Such devices were first developed in the (1) 1960s and 1970s and included the patient operated selector mechanism (POSM) and communication boards. More sophisticated devices surfaced in the (2) mid-1970s to the 1980s and included speech generating devices (SGDs). These devices were not always simple for the user to operate, and in the (3) 1980s and 1990s research on ways to optimize these aided communication devices increased (Vanderheiden, 2002).
1. **1960s and 1970s.** The POSM (Figure 1), developed by Reg Maling in the 1960s, was a typewriter controlled through inhaling and exhaling into a mouthpiece (Vanderheiden, 2002). A series of similarly controlled typewriters were developed following the release of the POSM, and these systems eventually transitioned to devices that were controlled by electromuscular signals instead of oral motor movements (Figure 2; Vanderheiden, 2002). Electrodes placed on the surface of the arm and hand of the patient recorded the impulses emitted by the muscles within the arm and hand. A computer connected to both the electrodes and the AAC device translated these impulses into a command that was sent to the device to operate it (Pinheiro et al., 2011; Vanderheiden, 2002). This was especially useful for individuals who were not only paralyzed, but who also had oral motor weaknesses.

Communication boards (Figure 3) were another alternative to oral speech that benefited individuals with difficulties controlling muscles involved in speech production.
Communication boards contained letters, words, or symbols that users could point to with their fingers, pointers, head sticks, or other mechanisms to spell out messages. There were a variety of communication boards available that accommodated the needs of different individuals. For example, some communication boards only contained pictures, while others contained hundreds of words. (Vanderheiden, 2002).

2. Mid 1970s and early 1980s. Most of the initial AAC devices such as the POSM or communication boards were keyboards that required the individual to type out messages, but experiments with speech synthesis eventually lead to the development of the first portable (hand-held or wearable) speech generating device (SGD) for individuals with disabilities in the mid-1970s (Vanderheiden, 2002). SGDs require speech synthesizers to synthesize speech. At this point, it would be useful to distinguish between the terms speech synthesis and speech synthesizer. Speech synthesis is the artificial production of human speech, whereas speech synthesizers are computer programs that generate this artificial human speech. Speech synthesizers such as the Votrax and Real-Voice served as platforms, applications or tools that
provide support for developers building speech-generating devices. Early SGDs (Figure 4) such as the HandiVoice and Echo relied on primitive speech synthesizers (i.e. the Votrax and Real-Voice) and both users and listeners of these devices reported that the quality of the generated speech was poor (Higginbotham, 2010). In the 1980s, developers began using a new speech synthesizer known as DECTalk as the platform for their SGDs. Although DECTalk improved intelligibility of output from SGDs in quiet environments, communication breakdowns still occurred in noisy environments because the background noise interfered with the signals produced from the SGD (Vanderheiden, 2002; Higginbotham, 2010; Drager, Clark-Serpentine, Johnson, & Roeser, 2006). Unlike with natural speech, visual cues such as lip reading and facial expressions could not be used to supplement these breakdowns that occurred in noisy environments (Drager et al., 2006).

3. 1980s and 1990s. AAC technologies initially had high physical and cognitive/linguistic demands on its users. This caused the user to place increased focus on operating the device, limiting face-to-face interactions, which ultimately interfered with social interaction (Higginbotham, Shane, Russell, & Caves, 2007). In the 1980s and 1990s,
research to optimize aided communication systems began to increase as improved ways to access and easily use the AAC device to meet complex communication needs were investigated. Various interface options became available for AAC devices. For example, eye tracking features, which allow a user to operate the AAC device through eye gaze, were especially beneficial for patients with very limited to no motor movement. Similarly, motion recognition features tracked and predicted body movements through sensors connected to computer devices to help an individual operate a device. Other features included brain interfaces which used electrodes placed on the scalp or implanted in the brain to allow an individual to control the device through brain waves. Unfortunately, many of these features were costly, and they did not always accurately predict the user’s movements (Higginbotham et al. 2007).

B. Speech Synthesizing Technology

1. Aspects of Speech. The comparison of human speech and synthesized speech highlights important prosodic features lacking in generated speech. Speech conveys information to the listener through both segmental and suprasegmental information. Segmental aspects of speech are the individual sounds that make up words (Gilbert, 2005, p. ix). For example, the sounds represented by the letters “w,” “a,” “t,” “e,” and “r” in the word “water” are segmental aspects of speech. However, the word “water” on its own provides little context to the listener. For example, the speaker might be asking the listener if they would like some water (i.e. water?) or making a statement (i.e. water!). Suprasegmental aspects of speech help convey the speaker’s intended message to the listener (Gilbert, 2005, p. ix). Suprasegmental aspects of speech, also known as vocal prosody, are the physical properties of speech such as amplitude or fundamental frequency that affect how speech is perceived (Bernthal, Bankson, & Flipsen, 2009). Suprasegmental speech cues such as pitch and loudness also convey emotion, making the context of verbal speech easier to interpret.
and understand (e.g., sarcasm vs. sincerity or anger vs. excitement; Pell, Jaywant, Monetta, & Kotz, 2011; Bernthal et al., 2009; Quene & Kager, 1992).

2. Speech Synthesis. Currently, speech synthesizing technology has provided users with the ability to control the segmental aspects of speech by taking input (such as text) and converting this input into output (speech). Speech generating devices (SGDs) featuring speech synthesizers that translate text into speech are known as text-to-speech (TTS) devices. Unfortunately, there is no speech synthesis system that produces both pleasant sounding speech and provides a flexibility to manipulate the generated sound to express emotion. Although prosodic features for SGDs are lacking, such devices are beneficial to patients who are unable to produce vocal speech because they provide the patients with an alternate voice. The primary speech synthesis technologies used in most early SGDs are formant and concatenative synthesizers (Pierre-Yves, 2002; Schroder, 2001). More recently, HMM-based (Hidden Markov Model) speech synthesis has gained popularity due to its ability to provide greater control over stylistic aspects of generated speech such as emotion and speaking style (Tokuda et al., 2013; Fosler-Lussier, E. 1998).

a. Concatenative Synthesis. Concatenative synthesis results from the stringing together of recorded speech sounds. The speech segments that are concatenated are typically monotonous, and the only parameters that can be manipulated with this method are pitch, intensity, and duration of phonemes. Additionally, conveying emotion through concatenative synthesis would require using pre-recorded segments for each emotion, limiting the real-time expression of emotion. Despite this, concatenative synthesis is preferable because it produces more pleasant sounding speech (Pierre-Yves, 2002; Schroder, 2001).

b. Formant Synthesis. Formant synthesis generates a wave based on the manipulation of acoustic properties of speech that are associated with the voice source and vocal tract such
as fundamental frequency and noise levels. The waveform generated is translated into human speech. Although this method allows for great control over the speech signal compared to concatenative synthesis, there are over 20 parameters needed to generate a speech sample, making it difficult to fine tune the sample. This causes the speech produced to sound unnatural (Pierre-Yves, 2002; Schroder, 2001; Styger & Keller, 1994, p.111).

**c. HMM-based Speech Synthesis.** Markov processes are non-deterministic systems that are modeled using probabilities. In a standard Markov Model, the end result is observable. Hidden Markov Models (HMMs) are a variation of standard Markov Models in which only a condition directly related to the end result, not the end result itself, can be directly observed. Hidden Markov Models must rely on probabilistic evaluation to determine the most likely final state (Tokuda et al., 2013; Fosler-Lussier 1998).

In Hidden Markov Model (HMM)-based speech synthesis, fundamental frequency and rate are used as parameters subjected to statistical analysis to generate a sequence of audio samples resembling human speech. Phoneme models dependent on grammar context are formed by the prior analysis of human vocal recordings. These analyses create the statistical models used as input for a sequence of Hidden Markov Models during the synthesis stage. The context-dependent phoneme models are arranged and statistically processed to generate the synthesized speech. This method provides greater control over speech due to its flexibility in changing emotion and speaking styles (Tokuda et al., 2013).

**C. Disorders Requiring AAC**

Individuals with complex communication needs (CCN) are typically unable to produce movements in the vocal tract with the precision required for oral communication. Such individuals tend to have neurological impairments associated with aspects of the nervous system that underlie vocal tract movements. Specifically, acquired neurological
disorders are neurological impairments that occur after birth and can negatively impact a person’s communication abilities (Duffy, 2005). Adults with acquired neurological disorders are typically able to comprehend and produce language, therefore enabling them to participate in social activities such as interacting with friends and family. Thus, although aspects of motor speech control may be impaired, the ability to express communicative intentions via linguistic structures is preserved. Within a social-communication environment, in which language users share a dynamic set of linguistic rules, individuals with neurological impairments affecting speech motor control may be able to communicate by means other than speech (i.e., verbal language through oral motor movements; Owens, Metz, & Haas, 2007).

Figure 5: Upper and Lower Motor Neuron pathway – generic example

Speech is a complex behavior that involves coordination among the respiratory, the phonatory, and the articulatory systems. Sound is produced when air from the respiratory system vibrates the vocal folds at the level of the larynx, which is then shaped through various constrictions within the vocal tract to produce speech sounds. All such movements are mediated by the nervous system. Upper motor neurons (UMN) that originate in the motor
cortex, and lower motor neurons (LMN) that innervate muscle fibers, provide the neural pathway through which speech is produced. Additionally, LMN play a critical role in maintaining autonomic actions such as muscle tone and posture (Murdoch, 2010, p. 245-246; Owens et al., 2007; Duffy, 2005). Figure 5 provides a pictorial representation of the UMN and LMN pathways. Acquired neurological conditions can result in damage to these neuronal pathways, leading to muscle weaknesses. For example, patients with amyotrophic lateral sclerosis (ALS) or traumatic brain injury (TBI) may develop dysarthria due to UMN and LMN damage, and as a result, have difficulty producing intelligible speech. For individuals with impairments to either UMN or LMN, both voluntary behaviors such as speech and autonomic behaviors such as emotional expression would be compromised.

1. Impact of Dysarthria on the Motor Speech System. Dysarthria is a motor speech disorder that results from the impairment of neuromuscular control. Individuals with moderate dysarthrias require AAC devices to supplement their speech to help repair communication breakdowns, whereas individuals with severe to profound dysarthrias require AAC devices in most circumstances (Higginbotham, Shane, Russell, & Caves, 2007). To better understand how AAC devices benefit patients with acquired neurological disorders, it is important to discuss why these disorders not only impair communication, but also prevent intelligible speech from being fully recovered.

a. Amyotrophic lateral sclerosis (ALS). Amyotrophic lateral sclerosis, also known as Lou Gehrig’s disease, is a progressive neurodegenerative condition that affects upper and lower motor neurons in the brain and spinal cord resulting in the loss of voluntary muscle control at the beginning stages of the disease and the eventual death of the patient at the end stages of the disease (Bukelman, Fager, Ball, & Dietz, 2007; ALS Association, 2013; ALS Information Page, 2013). As the disease progresses, the muscles in the diaphragm and chest begin to fail and patients eventually require ventilation (ALS Information Page, 2013).
Patients with ALS eventually experience speech difficulties with around 95% of patients becoming unable to speak at some point prior to death (Bukelman, Fager et al., 2007).

**b. Traumatic Brain Injury (TBI).** Traumatic brain injuries result from either penetrating or closed head injuries and can result in speech difficulties (ASHA, *Traumatic Brain Injury*). Communication difficulties that persist during the middle stages of recovery from TBIs typically result from impairments to the respiratory and phonatory system as a result of chronic, severe dysarthria (Bukelman, Doyle, & Karantounis, 2007; Bukelman, Fager et al., 2007). The abnormal speech characteristics that impair communication in patients with dysarthria are discussed below.

2. **Speech Behavioral Effects of Dysarthria.** Abnormal speech characteristics of patients with dysarthria may include (Duffy, 2005):

- **Prosodic excess:** excess and equal stress and a slower rate of speech
- **Articulatory-resonatory incompetence:** hypernasal speech, articulation errors, nasal emissions, or the use of shorter phrases
- **Prosodic insufficiency:** a monotonous or harsh vocal quality
- **Phonatory incompetence:** breathy, harsh, or strained-strangled speech that may also consist of pitch breaks
- **Inability to sustain phonation**

These characteristics may impair speech by changing the affected person’s vocal quality to a harsh or breathy texture. Motor speech disorders may also interfere with the affected person’s ability to sustain speech for extended periods of time or to produce changes in intonation or loudness associated with emotional speech.
3. **Types of Dysarthria.** Both ALS and TBIs can result in multiple types of dysarthrias including flaccid, spastic, and mixed dysarthrias. Patients with ALS and TBIs will present with different speech characteristic based on the type of dysarthria they acquire.

   a. **Flaccid Dysarthria.** Flaccid dysarthria results from damage to the lower motor neuron system and causes muscle weakness, reduced muscle tone (hypotonia), muscle atrophy, and involuntary muscle twitches and contractions. Around 32% of all flaccid dysarthrias result from TBIs while 13% of cases result from degenerative diseases such as ALS. Speech difficulties include phonatory and resonatory incompetence as well as phonatory-prosodic insufficiency (Duffy, 2005). Patients with flaccid dysarthria may have reduced loudness along with a breathy or harsh vocal quality. They may also exhibit hypernasal or robotic speech (McCraffrey, 2013).

   b. **Spastic Dysarthria.** Spastic dysarthria results from bilateral damage to the upper motor neuron (UMN) system and causes a wide variety of effects. Damage to UMN causes a loss of fine, skilled muscle movement and increased muscle tone (hypertonia), hyperactive reflexes, and spastic muscle movements. Approximately 40% of all spastic dysarthria cases result from degenerative disorders while around 8% are caused by TBIs. Speech difficulties include prosodic excess, articulatory-resonatory incompetence, prosodic insufficiency, and phonatory stenosis (Duffy, 2005). Patients with spastic dysarthria may have a strained or harsh vocal quality combined with occasional bursts of loudness, and may also show signs of hypernasal speech (McCraffrey, 2013).

   c. **Mixed Dysarthria.** Mixed dysarthrias are the most common type of dysarthria speech-language pathologists encounter and are a mixture of any two or more dysarthrias. Sixty-six percent of all mixed dysarthria cases result from degenerative diseases and 5% of all cases are caused by TBIs. The most common type of mixed dysarthria is flaccid-spastic,
and it accounts for 42% of all mixed dysarthria cases. Characteristics of mixed dysarthrias depend on the types of dysarthrias the patient acquired. For example, a patient with flaccid-spastic dysarthria would exhibit speech and motor characteristics common to both flaccid and spastic dysarthria (Duffy, 2005). Depending on the location of the damage (UMN vs. LMN), patients with mixed dysarthria would have either a breathy or a harsh vocal quality. LMN damage can result in a weakening at the laryngeal valve, specifically within a structure known as the glottis. This weakness leads to inadequate opening and closing of the vocal folds, causing patients to have a breathy vocal quality (McCraffrey, 2013; Duffy, 2005). On the other hand, increased tension in the laryngeal muscles as a result of UMN or decreased muscle tension in the laryngeal muscles as a result of LMN can cause patients to have a harsh vocal quality (McCraffrey, 2013; Duffy, 2005).

D. Benefit and Limitations of AAC Devices

1. Benefits of AAC. Text to speech (TTS) devices can provide individuals with complex communication needs (CCN) with the ability to become active participants in their environment. TTS devices would be particularly advantageous to patients with motor speech disorders because they reduce the physical (motor) demands of natural speech production while providing users with a way to produce immediate acoustic output and increasing the number of conversational turns (Blischak, Lombardino, & Dyson, 2003).

Prior to the 1980s, it was still considered taboo for speech pathologists to use AAC devices with their clients, even if their client was making little verbal progress after a year of therapy (Vanderheiden, 2002). Growing evidence suggests not only that AAC intervention does not harm speech development or recovery, but AAC also has positive benefits as well, such as increasing the frequency of vocalizations and intelligibility of the user (Vanderheiden, 2002; Light & McNaughton, 2012). AAC intervention improves
communicative effectiveness by promoting communication, enhancing language and literacy, and helping some patients recover natural speech (Blischak et al., 2003).

Prior studies have investigated experiences of patients using AAC devices. Patients reported their AAC devices helped them become more active in society by giving them the ability to communicate with unfamiliar people and learn new skills (McNaughton & Bryen, 2007). In addition, patients were also able to find and maintain jobs using their AAC device. Employers reported that AAC devices increased the quality of work of their employees who required AAC devices as well as had a positive impact on their coworkers (McNaughton et al., 2007).

Not only have patients reported positive benefits from using AAC devices, but the adoption rate of AAC devices has increased among patients required to use such devices. For example, AAC acceptance among both male and female patients with ALS increased from between 72-74% prior to 1996 to 96% in 2004 (Bukelman, Fager et al., 2007). Studies also indicate that 94% of patients with TBIs that were advised to use high-technology AAC devices such as TTS devices began regular device usage. Three years later, 81% of these individuals continued to utilize their device (the patients who discontinued use of their device did so due to lack of necessary funds or facilitator support; Bukelman, Fager, et al., 2007).

2. Limitations of AAC Devices. Although AAC devices provide many positive benefits, patients continue to prefer naturally produced speech as opposed to synthetically produced speech (Ratcliff, Coughlin, & Lehman, 2002). However, most devices have not been designed to facilitate the naturalness of real-time social interactions and cause generated speech to be perceived as flat and lacking emotion. Speech generated from TTS devices is computerized, and as a result sounds monotonous and robotic. There are currently no available devices that provide the user with a way to personalize their own voice. Although
text to speech devices with prerecorded messages that do convey emotion exist, such devices still limit the user’s ability to express emotion during speech in real-time (Higginbotham, 2010; Pierre-Yves, 2002). This limitation prevents users from being able to convey emotion in speech. Facial expressions, prosody, and the spoken word collectively contribute to a person’s ability to convey emotion, which in turn facilitates a person’s social interactions and relationships (Regenbogen et al., 2012; Mauss et al., 2011). In addition, individuals with motor speech disorders typically have facial muscle weakness, preventing them from using facial expressions to compensate for emotionless speech. Thus, the inability to convey emotion limits the affected person’s ability to fully participate in an important activity for social interaction: conversation.
II. PURPOSE AND AIMS

A. Purpose

The purpose of this thesis was to develop a text to speech (TTS) device in which users were able to manipulate suprasegmental aspects of computer generated speech in real-time. Specifically, the user was able to manipulate pitch and loudness using a touchscreen. The user was able to increase and decrease loudness by sliding their finger upwards and downwards across the touchscreen and increase and decrease pitch by sliding their finger right and left across the touchscreen. Such a development would potentially enhance communicative experience by enabling users to convey emotion.

This thesis project investigated the communicative effectiveness of a text to speech AAC device that allowed the user to control pitch and volume as speech was generated. There were two phases to the study. During the first phase of the study, the participants listened to and identified recordings based on affect (positive vs. negative) and linguistic (question vs. statement) categories.

During the second phase of the study, the participant(s) in the study used the device over a period of 2-4 weeks, and data was gathered upon initial and final use to assess 1) how well the device met the communication needs of the participant, and 2) whether these extra features were beneficial. The participants were trained on how to use the device during the first data collection session.
B. Emotional and Prosodic Aspects of TTS Devices

Previous experiments conducted using computer-based sound manipulation techniques indicated that pitch, intensity, and timing were the most crucial aspects related to emotional speech (Pierre-Yves, 2002). The device was built for the purposes of this thesis was intended to allow the user to manipulate pitch (by changing frequency) and volume as speech was being generated. A touchscreen was chosen to accommodate these manipulations because it detects and translates even limited finger movements while also providing optical feedback. The purpose of the device is to provide an individual with complex communication needs with a new voice of their own.

C. Aims of this Study

1. To adapt existing speech synthesizing technology to meet the needs of individuals with complex communication needs.

2. To assess whether the device improves the communication abilities of people who use it.

3. To assess whether users feel the device improves their communication abilities.
III. METHOD

A. Participation

Twelve adult individuals (6 male and 6 female) with no prior knowledge of the study or the thesis device were recruited to provide auditory feedback regarding the efficacy of the device in conveying suprasegmental information. These participants were administered a pure tone hearing test at the following frequencies- 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz – to confirm that their hearing was within normal limits. The participants then listened to utterances prerecorded with (1) the unmodified TTS device, (2) the thesis TTS device, and (3) human speech. These utterances were then identified by each participant based on (1) perceived meaning (question vs. statement) and (2) affect (positive vs. negative). Investigators did not identify which recording belonged to the thesis device, the unmodified device, or human speech, and the recordings were randomized to avoid potential response bias. The hearing test and the recordings were presented in a sound booth.

In addition, 2 adult individuals with mild to moderate dysarthria (e.g., as a result of ALS or a TBI) participated in this study. The individuals had preserved language abilities and a level of physical impairment in which they were able to operate such a device. These participants utilized and provided use-based feedback regarding the efficacy of the device. Data was collected following a training session/initial use of device and then once more after a 1 week interim. The participants provided feedback through open-ended questions as well as based on a rating scale of 1 to 5 (1- lowest and 5- highest).
A description of the hardware (physical components) is provided first. This is followed by a discussion of the software (computer program). Finally, testing procedures of the AAC device is explained according to both auditory and use-based performance measures.

1. The device. The hardware for the text to speech device is an Android tablet (i.e. Samsung Galaxy and Asus Nexus, similar to an iPad) that takes user input via touchscreen so that the user can manipulate pitch and volume of generated speech. If the user slides his or her finger to the left, pitch decreases, and if the user slides his or her finger to the right, pitch increases. On the other hand, if the user slides his or her finger up, volume increase, and if the user slides his or her finger down, volume decreases.

The touchscreen tablet was chosen because it is portable which gives the patient the ability to use the device in multiple settings. The touchscreen also provides a user with even the most limited hand or arm movements the ability to use the device. The user can use the touchscreen to build sentences either through a keyboard component or by selecting built-in vocabulary words. The vocabulary words are from the AAC Institute’s AAC Performance Report (“AAC Institute,” n.d.). Refer to Appendix A for a list of the core vocabulary words that were included in the device (Hill & Romich, 2001). Combining all the components of the text to speech device (i.e. touchscreen, built-in keyboard and vocabulary words) into one instrument provides users with increased opportunities for face-to-face interactions, enhancing social communication.

2. Programming the device. Free text to speech (TTS) engines such as NeoSpeech (NeoSpeech™, 2011) or Acapela (Acapela Group, 2013) are available for download and use. TTS engines are speech synthesis engines which read text and convert it into speech
(Schneiderman & Plaisant, 2010, p. 338-9). These TTS vendors translate text into speech in both male and female voices, and some engines automatically attempt to include natural-sounding inflections. Most engines, however, either require complex commands to produce natural-sounding inflections or only have the ability to generate monotonous speech.

Two approaches were considered for the programming of the device: (a) markup-based synthesizer instructions [SAPI tags] and (b) post-synthesis audio signal processing.

a. Markup-based synthesizer instructions. The TTS vendors mentioned above (i.e. NeoSpeech and Acapella) run on the speech application programming interface (SAPI) standard. The SAPI standard is a programming interface created by Microsoft to allow the use of speech recognition and synthesis. The voices are modulated in pitch, speed, and volume by inserting the speech application programming interface (SAPI) command tags into the text. SAPI tags are built-in commands that run on the speech application programming interface. Tags are included for each type of modulation such as pitch, rate, and volume. The speech engine changes the output based on the tag it encounters. For example, if the speech engine encounters a pitch tag, it changes the pitch of the generated voice from that point forward. The SAPI tags are analogous to neurotransmitters and the speech engine is analogous to a neuron (nerve cell). Depending on the type of neurotransmitter (SAPI tag) the neuron (speech engine) encounters, the neuron reacts differently. For example, the neurotransmitter glutamate has an excitatory effect on the neuron (Webb, W. & Adler, R., 2008).

b. Post-synthesis audio signal processing. TTS vendors synthesize text into speech, and the generated complex audio wave can be processed to achieve modulation of pitch, volume, and rate. Pitch and rate are modulated through a process known as resampling. Resampling is simply the stretching or compressing of the audio wave. Pitch and rate are
inversely related. As a result, additional processing is required to provide custom modulation of both pitch and rate independently. This is achieved by taking overlapping splices of the input (audio) wave known as frames and then either increasing or decreasing the overlap between the frames. This artificially shortens or extends the wave, respectively. The frame edges are then smoothed to produce a higher quality output. Finally, the frames are recombined to generate an output (audio) wave. Volume is modulated by sending instructions to the device’s hardware speakers to either increase or decrease playback volume.

To help clarify audio resampling and signal processing, consider the audio wave as a stick of chewing gum (Figure 6). If you were to stretch the stick of gum to make it longer, it would become thinner. This is comparable to a decrease in pitch and an increase in duration. Likewise, if you were to compress the gum to make it shorter, it would become thicker. Figure 7 provides a pictorial representation of this process. This corresponds to an increase in pitch and a decrease in duration. No matter how hard you stretch or compress the stick of gum, there is no way to affect either thickness or length individually. This is where signal processing comes into play.
Audio waves can be stretched and compressed (similar to how a stick of gum can) through a process known as resampling.

To overcome the limitation of inversely related properties, an alternative approach must be considered. Instead of stretching or compressing the stick of gum, an optimal strategy would be to simply generate more of it. To accomplish this without discernable repetition, the stick of gum is sliced into small, regularly sized pieces. These pieces are then analyzed to determine their fundamental ingredients, and then a new slice of the appropriate size is generated. After all the new slices are complete, they are then combined to form a new stick closely resembling the original, with the same pitch, but with a different duration. This new piece can then be stretched or compressed to achieve the desired attributes. Refer to Figure 8 for a pictorial representation of this process.

Audio waves are comprised of different frequencies, similar to how a stick of chewing gum is composed of different ingredients. Identifying the frequencies that make up an audio wave allows the signal to be recreated.
Similar to a stick of chewing gum, audio waves are comprised of different “ingredients” or properties. Complex audio waves such as those that are generated by speech are a combination of individual sine waves at various frequencies (i.e., a Fourier series). Analyzing and identifying the properties of a generated audio wave allow the signal to be recreated at a variable length while maintaining pitch and signal resolution. For example, to increase the pitch, the properties of the analyzed audio wave would be used to create a longer version of the original wave. The newly created wave is then resampled back to the original duration for no net difference, while a net increase in pitch is observed (Figure 9). Likewise, to decrease pitch, the properties of the original wave would be replicated to produce a shorter version of the wave which would then be stretched back to the original duration via resampling.

Figure 9: Signal Processing and Resampling of a Wave – Generic Example

The program for this thesis acted as a bridge between the TTS generator and the device speakers using an event-driven model. An event-driven model is a software development architecture based on the concept of significant changes causing an event to occur. The event is then passed to an event handler for further processing. An event in the scope of this application is an action or occurrence detected by a program such as a keystroke or mouse click. An event handler is a function containing lines of code that tell the program what to do in response to an event. The event handler then translates that input (e.g., touchscreen movements) into the appropriate modulation of speech. The resulting output is then played by the hardware speakers (Deitel & Deitel, 2012, p. 561).

Pitch and loudness constraints were added to the program to prevent the user from increasing or decreasing the fundamental frequency (for pitch) and amplitude (for loudness) to an uncomfortable level. The ceilings and floors were based on natural human speech. The range for pitch is 100-8000 Hz and for loudness is no more than 85 dB (Kent and Read, 2002).

To further explain how the AAC system works, (c) a description of the system workflow is provided along with (d) a system workflow example. For a more detailed definition of how the AAC device and user interact to allow the user to convey emotion refer to the use case in Appendix B.

c. Description of System Workflow.

1. The speech engine (program that converts text into speech) began synthesis with default parameters.
2. The hardware detected finger movements and translated them into events based on the direction and intensity of the movements.
3. The event handler in the program translated the events into the appropriate volume, rate, and pitch parameters, updating the previous parameters in the process.

**d. System Workflow Example.**

1. The user intended to increase the pitch of the speech the device is generating.
2. The user slid his (or her) finger along the touchscreen in the upward direction.
3. The hardware detected the movement from the touchscreen and invoked the program with an event.
4. The program received the event and translated the vertical axis of motion into a request for change in pitch.
5. The program modified the target pitch parameter to match the updated value received from the event.
6. The speech engine increased the pitch of the generated speech.

**3. Testing the Device.** Auditory performance measures assessed whether the device (a) conveys emotion and (b) convey different meanings, while use-based performance measures assessed (c) increased efficacy of communication of individuals with complex communication needs. Below is a more detailed explanation of the empirical data that was collected.

**a. Auditory data #1 (Intonation).** Five statements/utterances that convey emotion (positive and negative affect) were prerecorded using (1) human speech, (2) the thesis TTS device, and (3) an unmodified TTS device. Twelve individuals listened to audio samples presented via a computer monitor, and responded to changes in pitch. Emotion can be categorized into positive and negative affect (Watson, Clark, & Tellegen, 1988). The individuals were asked to select whether they perceived positive or negative affect in each statement/utterance for each of the recordings. For example, the statement “It is raining” can
be expressed with positive affect by raising intonation on the word *raining* (as when the garden finally gets the rain it needs). This same statement can also be conveyed with negative affect by decreasing intonation on the word *raining* (as in the wedding is spoiled). Refer to Appendix C for more details.

- **Control group**: Natural speech
- **Experimental group #1**: the TTS device designed for the purposes of this project
- **Experimental group #2**: a TTS device that does not allow for manipulation of pitch and volume.

**b. Auditory data #2 (Stress).** Five statements/utterances that convey different meanings based on syllable or word stress were prerecorded using (1) human speech, (2) the thesis TTS device, and (3) an unmodified TTS device. Twelve individuals listened to audio samples presented via a computer monitor, and responded according changes in loudness. They individuals were asked to select whether they perceived a question or a statement. For example, the utterance “She can dance” is conveyed as a question if the word *dance* is stressed. On the other hand, that same utterance is conveyed as a statement/comment if the word *can* is stressed. Refer to Appendix C for more details.

- **Control group**: Natural speech
- **Experimental group #1**: the TTS device designed for the purposes of this project
- **Experimental group #2**: a TTS device that does not allow for manipulation of pitch and volume.

**c. Use-based data #1.** Two participants with complex communication needs (patients with dysarthria as a result of ALS or a TBI) provided subjective feedback regarding whether they feel the device was beneficial and made a difference in their communicative interactions. The participants had the opportunity to voice their opinions both after initial use of the device and following a 1-week interval. Refer to Appendix D to see the questionnaire
IV. RESULTS

Results collected during the development phase as well as auditory and use-based measures obtained during the testing phase are presented below.

A. Development Phase

The post-synthesis audio signal processing approach was utilized during the text to speech (TTS) software development cycle. Initial research during the development phase revealed that this approach was preferable compared to the use of markup-based synthesizer instruction (SAPI tags) method. The TTS engine selected for this thesis is Hidden Markov Model (HMM)-based.

The Android platform was chosen because of its dominant market share (Tyson, 2014) as well as its inclusion of text-to-speech within the core API. It also uses Java, a programming language with a large development community that is also able to run on different devices without modification (Flanagan, 2005). The Android platform also has an open source development status, which allows anyone to make improvements and modifications to its software.

The main benefit of using signal processing over SAPI tags is that signal processing allows the audio to be manipulated in real-time. The TTS engine is allowed to fully generate the entire phrase, enabling context to be read and implemented. This eliminates the possibility of repeated audio signal dropout caused by TTS engine inefficiencies. After the complete phrase has been generated, the finished audio wave is queued for playback within the device. This normally involves reading small chunks of data representing several
milliseconds of audio from memory, and then queuing the chunks for playback from the audio hardware. Processing the audio signal between these two steps allows for real-time manipulation of the audio signal in the immediate step before it is played audibly.

The core advantage of the algorithm designed for this thesis is that it can be applied to the final output of any TTS engine, regardless of synthesis type, and still generate similar results. This allows the device used for speech generation to be interchanged as long as it provides an audio signal containing the generated speech. HMM-based synthesis was selected for the purposes of this thesis because of its ability to provide greater control over emotion and speaking style.

B. Auditory Data

Twelve participants with typical hearing were asked to listen to recordings made using the unmodified TTS device, the thesis TTS device, and human speech. The tables below summarize the ability of these participants to accurately identify the prerecorded utterances as question vs. statement (Table 1) or as positive vs. negative affect (Table 2) for the unmodified TTS device, the thesis TTS device, and human speech.

![Accuracy of Utterance Identification - Question vs. Statement](image.png)

Table 1: Question vs. Statement Results

<table>
<thead>
<tr>
<th></th>
<th>Total Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Question</strong></td>
<td>100%</td>
</tr>
<tr>
<td><strong>Statement</strong></td>
<td>80%</td>
</tr>
</tbody>
</table>

Accuracy of question vs statement utterance identification based on type of each of the recording methods (unmodified TTS device, thesis TTS device, and human speech).
Accuracy of positive vs. negative utterance identification based on type of each of the recording methods (unmodified TTS device, thesis TTS device, and human speech).

<table>
<thead>
<tr>
<th></th>
<th>Unmodified Device</th>
<th>Thesis Device</th>
<th>Human Speech</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \chi^2 )</td>
<td>df</td>
<td>p-value</td>
</tr>
<tr>
<td>Q</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>S</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>P</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>N</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Q</td>
<td>0</td>
<td>1</td>
<td>NS</td>
</tr>
<tr>
<td>S</td>
<td>.2344</td>
<td>1</td>
<td>NS</td>
</tr>
<tr>
<td>P</td>
<td>2.0108</td>
<td>1</td>
<td>NS</td>
</tr>
<tr>
<td>N</td>
<td>1.454</td>
<td>1</td>
<td>NS</td>
</tr>
<tr>
<td>Q</td>
<td>1130.284</td>
<td>1</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>S</td>
<td>.817</td>
<td>1</td>
<td>NS</td>
</tr>
<tr>
<td>P</td>
<td>35.5171</td>
<td>1</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>N</td>
<td>46.9386</td>
<td>1</td>
<td>&lt;0.05</td>
</tr>
</tbody>
</table>

Q – Question, S – Statement, P – Positive, N – Negative. \( \chi^2 \) – chi squared value, df – degrees of freedom, p-value – probability of rejecting the null hypothesis, NS – Non-significant. The following table summarizes the significance of relationships between the different recording methods (i.e., unmodified TTS device, thesis TTS device, and human speech).

Non-parametric statistical analysis of the data was performed to determine the significance of the relationships between the thesis TTS device and human speech, the thesis
TTS device and the unmodified TTS device, and the unmodified TTS device and human speech. The results are presented in Table 3. An alpha level of 0.05 was used for all statistical tests. Significance was determined by comparing how accurately participants identified each utterance type (question vs. statement and positive affect vs. negative affect) based on recording method (unmodified TTS device, thesis TTS device, and human speech).

A proportion test was used to compare how accurately participants identified questions vs. statements with each recording type. A significant difference was found between the participants’ ability to identify questions presented via the thesis TTS device and human speech. A significant difference was also revealed between the participants’ ability to identify questions presented via the unmodified TTS device and human speech. Results showed participants were able to identify questions most accurately in utterances recorded by human speech. There was no significant difference between the participants’ ability to accurately identify questions presented via the unmodified TTS device and the thesis TTS device and. Results showed that participants were able to identify statements more accurately than questions in utterances presented via the unmodified TTS device and the thesis TTS device. There was no significant difference in the participants’ ability to accurately identify statements presented via all recording methods. As mentioned earlier, typical (unmodified) TTS devices generate monotonous speech which is perceived as flat. As a result, all speech generated from such devices is perceived as a statement. This is consistent with the results that show both the unmodified TTS device and thesis TTS device conveyed statements as effectively as human speech.

A proportion test was also used to compare how accurately participants identified positive emotion vs. negative emotion with each recording method. A significant difference was found between the participants’ ability to identify positive and negative emotion in utterances presented via the thesis TTS device and human speech. A significant difference
was also found between the participants’ ability to identify positive and negative affect in utterances presented via the unmodified TTS device and human speech. Results revealed participants were able to identify positive and negative emotion most accurately in utterances recorded by human speech. Results also showed participants were able to identify negative emotion more easily than positive emotion for all recording methods. Although participants were able to identify positive and negative emotion more easily with the thesis TTS device than with the unmodified TTS device, a proportioned test showed no significant difference existed between the two recording methods.

These results indicate that the thesis device is not as effective as human speech when communicating emotion and intent.

C. Use-based Data:

Two individuals with mild to moderate dysarthria utilized and provided feedback about the thesis TTS device in the following areas: (1) device interface, (2) audio output, (3) new features, and (4) overall usefulness. Data was collected following the training session/initial use of device and then once more after a 1 week interim. The participants provided feedback on 14 different features through open-ended questions as well as based on a rating scale of 1 to 5 (1- lowest and 5- highest).

1. Device Interface. The participants provided positive feedback about the ease of learnability, the layout, and the color scheme of the device. The individuals noted that the device was easy to manipulate and navigate and provided an overall rating of 4 for this feature. The participants also indicated that the layout and color scheme were consistent and clear, allowing for easy readability of text, and provided a combined rating of 5 for these features. The users also reported that it was easy to reverse mistakes made when using the device and gave this feature an overall rating of 5.
Finally, the participants voiced negative feedback regarding button size and provided an overall rating of 2.5 for this feature.

2. Audio Output. The participants provided positive feedback about the intelligibility of speech and provided an overall rating of 5 for this feature.

3. New Features. The participants provided positive feedback about the new features of the device, including the ability to manipulate pitch and volume of text generated speech in real-time. The individuals expressed that the new features were beneficial. The users noted that it was easy to understand and convey the emotions and intent they wanted to express and provided an overall rating of 5 for these features.

4. Overall Usefulness. The participants noted the word bank and keyboard features were useful and provided an overall rating of 5 for these features. The users indicated they found the device to be beneficial for people with complex communication needs and strongly agreed that they would recommend the device to someone who needed it.

Analysis of the feedback obtained in the use-based phase of the study indicated that thesis TTS device was received favorably by the participants with complex communication needs (CCN). The new features (ability to manipulate pitch and volume) were found to be valuable by providing individuals with CCN a new way to express themselves. Examination of the results also highlighted several ways the thesis TTS device could be improved upon that would allow the device to better meet the needs of its users.
V. DISCUSSION

This thesis adapted existing speech synthesizing technology to allow for real-time manipulation of pitch and volume of text generated speech and examined whether these features improved the communicative effectiveness of individuals with complex communication needs. Feedback gathered throughout the development phase and testing phase of the thesis study highlighted several components for enhancement of overall user experience in the areas of audio manipulation and interface design, which are discussed in more detail below.

A. Auditory Data

Samples recorded from the thesis TTS device were compared to samples recorded from an unmodified TTS device and human speech to determine the overall efficacy of the thesis TTS device. A significant difference was found between the thesis TTS device and human speech for both the linguistic and affect categories. This indicated that the thesis TTS device was not as effective as human speech for verbal communication. A non-significant difference was found between the thesis TTS device and the unmodified TTS device for both the linguistic and affect categories. These results suggest that the thesis TTS device was not an improvement over a typical TTS device in terms of communicative effectiveness.

B. Use-based Data

Participants with CCN were provided with the opportunity to use the thesis TTS device during the course of this study. The participants noted features of the device that
limited their ability to fully interact with the device. Specifically, they indicated that larger buttons for the built-in vocabulary words and on the keyboard would allow for better motor control. The individuals also stated that it would be beneficial if the rate of the output could be controlled. Overall, however, the users reported a positive experience with the device. The device was easy to manipulate, and the individuals were excited about the device providing them with the ability to manipulate pitch and volume during speech generation.

C. Prosodic Characteristics of Speech

A discrepancy existed between the results obtained through auditory and use-based measures which could have resulted from several factors, including participant bias and sound quality of the text generated speech. Informal interviews with participants following the auditory data collection revealed that the participants sometimes were unsure of the emotion or intent expressed in the recordings and as a result, provided a guess based on context. For example, for the utterance “It is raining,” participants who disliked rain categorized this utterance as negative affect. Similarly, because the utterance “He walked to the mall” was not phrased as a question, participants mentioned that they identified the utterance as a statement. This could have affected the overall accuracy of the results.

Pitch and volume constraints could be another reason why it was difficult to convey emotion and intent effectively. When producing human speech, pitch and volume changes occur naturally and automatically; however, when using the thesis TTS device, the user is required to think about how to manipulate pitch and volume in order to convey their desired emotion or intent. The users had to learn how to speak all over again. Similarly to how the acquisition of speech and language takes time, learning whether to increase or decrease pitch and volume and by how much will also take time. When producing the recordings, identifying the appropriate pitch and volume constraints was difficult and may have
influenced the ability for affect and intent to be conveyed accurately. Changes to implement to improve learning the system would be an important step in the development of this device.

A third factor is that while the device allowed for manipulation of pitch and volume, it did not allow the user to change rate and as a result, the recordings sounded distorted. This may have impacted the overall quality the recordings. This issue can be resolved by modifying the algorithm used when programming the device to take into account the distortion by allowing the user to manipulate rate in addition to pitch.

The individuals who participated in the use-based data collection phase were aware that the investigator helped develop the thesis TTS device and expressed their appreciation of the investigator’s efforts to develop a device that could enhance their communicative on several occasions. The investigator was responsible for the data collection, and it is possible because of these circumstances the participants were less than upfront about voicing criticism of the device. A combination of all these factors could account for the discrepancy between the auditory and use-based results.

**D. Future Direction**

A number of future developments can be implemented to improve both the design and user experience of the thesis TTS device. First, to closer mimic human vocal qualities of the text generated speech, the algorithm could be modified to account for partial resampling to allow for the simultaneous rate and pitch manipulation present in human speech. Additionally, in order to benefit a broader range of individuals, alternate methods for pitch and volume manipulation could be implemented to improve the process of learning how to use the device. User feedback and suggestions from Schneiderman’s 8 Golden Rules of Interface Design (Schneiderman & Plaisant, 2005) propose the following interface improvements: increased button size, protection against accidental button presses, improved
scrolling options, and safe zone. To mitigate the effort required to scroll through large collections of words, lists will provide an index to automatically and immediately scroll to a particular entry. A safe zone can better accommodate users with limited motor control by restricting the area on the screen that displays input.
VI. CONCLUSION

The ability to interact and communicate with others is an integral part of social interaction and overall quality of life. Some individuals have complex communication needs (CNN) as a result of neurological disorders such as dysarthria, which interfere with speech production and thus prevent them from expressing themselves effectively. Dysarthria can result from degenerative conditions such as ALS or from TBIs sustained in accidents. Such individuals require augmentative and alternative methods (AAC) such as text to speech (TTS) devices to supplement oral speech. Current AAC devices do not provide users with the ability to produce suprasegmental aspect of speech and are therefore limited compared to human speech. This thesis examined whether real-time prosodic aspects of text generated speech enhanced the communicative abilities of individuals with CCN required to use TTS devices. During the study, attempts to adapt existing speech synthesizing technology to allow AAC users to manipulate pitch and volume as speech was generated were successful. Although use-based feedback was positive overall, auditory measures indicated that a significant difference exists between accurate identification of emotions and intent from human speech and from the speech generated by the thesis device. The idea presented in this thesis is groundwork for future alterations such as the addition of rate manipulation and enhancing the quality of generated speech output.
REFERENCES


low  lower  lowest
made  make  makes  making
many  may  me  mean  meaner  meanest
meaning  means  meant
might  mine  money
more  most  mostly
much  my  myself
near  nearer  nearest
nearly  need  needing
needs  never
new  newer  newest
next  nice  nicer  nicest
night  no  nobody
not  nothing
now  nowhere
of  off  oh  old  older  oldest
on  one  only
or  other  others
our  ours  ourselves
ourselves
out  outside  over  overly
own  owned  owner
owning  owns  people
place  placed  places
placing  prettier  prettiest
pretty  probable
probably
put  puts
putting
quite
real  really
right
said  same
saw
say  saying
says
see
seen
she  she's
short  shorter  shortest
should  shouldn't
slow  slower  slowest
small  smaller  smallest
so  some  somebody
somebody
some day
some day
some how
someone
some thing
sometime
sometimes
somewhat
somewhere
sort  spell
spelled  spelling
spells  start
started  starting
starts  stop
stopped  stopping
stops  stuff
such
take
taken
takes
taking
talk
talked
talking
talks
tell
telling
tells
than
that
that's
the
their
their's
them
then
there
there's
these
they
they're
ting
think
thinking
thinks
this
those
though
thought
thousand
three
through
time
times
to
told
too
took
two
under
up
us  use
used
uses
using
very
wait
waited
waiting
waits
want
wanted
wanting
wants
was
way
ways
we
week
weeks
well
went
were
what
when
where
where's
which
who
whose
why
wife
will
with
within
without
wives
won't
work
worked
working
works
worse
worst
would
wouldn't
yeah
year
years
yes
you
you're
you've
your
yours
yourself
yourselves
Appendix B

Use case diagram of how the AAC user interacts with the AAC system. The AAC user inputs text using a keyboard/touch screen/touch pad. The AAC system converts this text into speech. The AAC user hears the output. As the output is being generated, if the user wishes to manipulate pitch or volume, the user slides their finger across the touch screen/touch pad. The AAC system interprets the movements and modifies the output accordingly.

Use Case #1:

Title: Speak with Emotion

Primary actor: Patient

Goal in Context: To produce emotional speech

Precondition: Patient needs to have the cognitive abilities to use this device

Minimal Guarantees: Patient is able to produce emotional speech using this AAC program.
Success Guarantees: Patient is able to produce emotional speech using this AAC program as speech is being generated using a touchscreen.

Trigger: Patient wants to convey emotion when speaking with their communicative partner.

Scenario:

1. The user constructs the message/sentence they want generated
2. The user presses the “speak” button to generate the message
3. The AAC system generates the message
4. The user slides their finger across the touchscreen in an upward direction
5. The AAC system detects and interprets these movements
6. The AAC system increases the volume of the generated speech output

Extensions:

1. The user desires to decrease the volume of the generated speech output so the user slides their finger downward across the touchscreen.
2. The user desires to increase the pitch of the generated speech output so the user slides their finger to the right across the touchscreen.
3. The user desires to decrease the pitch of the generated speech output so the user slides their finger to the left across the touchscreen.
Appendix C

I. Questions vs. Statements

Note: The syllables/words that were stressed to convey each utterance as either a question or statement are italicized.

1. He walked to the mall? vs. He walked to the mall.
3. Well, what are you going to do? vs. Well, what are you going to do?
4. This is where we are staying? vs. This is where we are staying.
5. She can dance? vs. She can dance.

Sample of the answer sheet

Directions: Please select the most appropriate based on the utterances that are played in each recording:

1. a. question  b. statement
2. a. question  b. statement
3. a. question  b. statement
4. a. question  b. statement
5. a. question  b. statement

II. Positive vs. Negative Affect

Note: The syllables/words that are italicized indicate the intonation was raised to convey positive affect while the syllables/words that are in bold indicate the intonation was decreased to convey negative.

1. It is raining. vs. It is raining.
2. It is almost done. vs. It is almost done.
4. We’re going home. vs. We’re going home.
5. It’s time to go. vs. It’s time to go.
Sample of the answer sheet

Directions: Please select the most appropriate based on the utterances that are played in each recording:

1. a. positive affect  b. negative affect
2. a. positive affect  b. negative affect
3. a. positive affect  b. negative affect
4. a. positive affect  b. negative affect
5. a. positive affect  b. negative affect
Appendix D

AAC Participant Questionnaire:

1. The text to speech device was easy to manipulate and navigate.
   1- Strongly disagree  2- Disagree  3- Neutral  4- Agree  5- Strongly agree

2. The text to speech device was easy to learn how to use.
   1- Strongly disagree  2- Disagree  3- Neutral  4- Agree  5- Strongly agree

3. The speech generated by the text to speech device was intelligible.
   1- Strongly disagree  2- Disagree  3- Neutral  4- Agree  5- Strongly agree

4. It was easy to convey and understand the emotions I wanted to express.
   1- Strongly disagree  2- Disagree  3- Neutral  4- Agree  5- Strongly agree

5. It was easy to convey and understand the intent I wanted to express.
   1- Strongly disagree  2- Disagree  3- Neutral  4- Agree  5- Strongly agree

6. I found the text to speech device to be useful.
   1- Strongly disagree  2- Disagree  3- Neutral  4- Agree  5- Strongly agree

7. I was satisfied with the size of the buttons in the text to speech device
   1- Strongly disagree  2- Disagree  3- Neutral  4- Agree  5- Strongly agree

8. The color scheme was consistent and allowed for easy readability of text.
   1- Strongly disagree  2- Disagree  3- Neutral  4- Agree  5- Strongly agree

9. The layout was consistent and clear enough to see and read.
   1- Strongly disagree  2- Disagree  3- Neutral  4- Agree  5- Strongly agree

10. The word bank was beneficial and easy to use.
    1- Strongly disagree  2- Disagree  3- Neutral  4- Agree  5- Strongly agree

11. They keyboard feature was beneficial easy to use.
    1- Strongly disagree  2- Disagree  3- Neutral  4- Agree  5- Strongly agree

12. I found the new features (being able to manipulate pitch and volume) to be beneficial.
    1- Strongly disagree  2- Disagree  3- Neutral  4- Agree  5- Strongly agree

13. It was easy to reverse/fix my mistakes when using the device.
    1- Strongly disagree  2- Disagree  3- Neutral  4- Agree  5- Strongly agree

14. I would recommend this text to speech device to someone who needed it.
    1- Strongly disagree  2- Disagree  3- Neutral  4- Agree  5- Strongly agree
Appendix E Consent Forms

E.1: Consent to Participate in a Research Study: Auditory Phase

WHY ARE YOU BEING INVITED TO TAKE PART IN THIS RESEARCH?
You are being invited to take part in a research study about perceptions of meaning/intent and emotion in recordings of various statements/utterances. You are being invited to take part in this research study because you are over the age of 19 and your hearing is within normal limits.

WHO IS DOING THE STUDY?
The person in charge of this study is Nure Biane Kassas of University of Alabama Department of Communicative Disorders. She is being guided in this research by Dr. Anthony Buhr.

WHAT IS THE PURPOSE OF THIS STUDY?
By doing this study, we hope to discover how individuals perceive emotion and meaning in speech.

ARE THERE REASONS WHY YOU SHOULD NOT TAKE PART IN THIS STUDY?
If you are under the age of 19 or having any degree of hearing loss, you should not take part in this study.

WHERE IS THE STUDY GOING TO TAKE PLACE AND HOW LONG WILL IT LAST?
The research procedures will be conducted at The University of Alabama Speech and Hearing Center. Your participation should require no more than 2 hours of your time.

WHAT WILL YOU BE ASKED TO DO?
Before participating in this study, you will be administered a hearing screening. The results of the hearing screening will not be recorded and no identifying information will be recorded or revealed. You will be required to listen to and rate your perception of various recordings of statements/utterances on a computer at the University of Alabama Speech and Hearing Center. After you hear a statement/utterance, you will be required to rate either the meaning or emotion you perceive in the statement/utterance from the options provided for you. You will be required to select an answer before moving on to the next recording.
WHAT ARE THE POSSIBLE RISKS AND DISCOMFORTS?

To the best of our knowledge, the things you will be doing have no more risk of harm than you would experience in everyday life.

WILL YOU BENEFIT FROM TAKING PART IN THIS STUDY?

You will not get any personal benefit from taking part in this study. Your willingness to take part, however, may, in the future, help society as a whole better understand this research topic.

DO YOU HAVE TO TAKE PART IN THE STUDY?

If you decide to take part in the study, it will be a voluntary basis. You will not lose any benefits or rights you would normally have if you choose not to volunteer. You can stop at any time during the study and still keep the benefits and rights you had before volunteering.

IF YOU DON’T WANT TO TAKE PART IN THE STUDY, ARE THERE OTHER CHOICES?

There are no other choices available if you choose not to be in the study.

WHAT WILL IT COST YOU TO PARTICIPATE?

The only cost to participate is your time.

WILL YOU RECEIVE ANY REWARDS FOR TAKING PART IN THIS STUDY?

You will not receive any rewards or payment for taking part in the study.

WHO WILL SEE THE INFORMATION THAT YOU GIVE?

Your consent form will be kept in a secure cabinet in a locked room at the University of Alabama Speech and Hearing Center. Any personal information gathered such as your gender, age, and type and severity of dysarthria will not be linked to your names (so all data collected will be not be identifiable). The only people who will have access to your personal information (age, gender, type and severity of dysarthria) are the research investigator and her thesis advisor.

CONFIDENTIALITY

You will be assigned an id number at the start of the study. This id number will be linked to the answers you provide while listening to the recordings. Your real name and any personally identifiable information will not be associated with the id number. The results of the hearing screening will not be recorded as within normal limits hearing is a requirement for you to be a participant in this study. Your consent form will be locked in a room at the University of Alabama Speech and Hearing Center and will be in a separate file from any data collected.
CAN YOUR TAKING PART IN THE STUDY END EARLY?

If you decide to take part in the study you still have the right to decide at any time that you no longer want to continue. You will not be treated differently if you decide to stop taking part in the study.

The individuals conducting the study may need to withdraw you from the study. This may occur if you are not able to follow the directions they give you or if you do not pass the hearing screening.

WHAT IF YOU HAVE QUESTIONS, SUGGESTIONS, CONCERNS, OR COMPLAINTS?

Before you decide whether to accept this invitation to take part in the study, please ask any questions that might come to mind now. Later, if you have questions, suggestions, concerns, or complaints about the study, you can contact the investigator, Nure Biane Kassas at 205-393-1484. If you have any questions about your rights as a volunteer in this research, contact Mrs. Tanta Myles, the Research Compliance Officer of the University of Alabama at 205-348-8461 or toll free at 1-877-820-3066. We will give you a signed copy of this consent form to take with you.

You may also ask questions, make suggestions, or file complaints and concerns through the IRB Outreach website at http://osp.ua.edu/site/PRCO_Welcome.html or email the Research Compliance office at participantoutreach@bama.ua.edu.

After you participate, you are encouraged to complete the survey for research participants that is online at the outreach website, or you may ask the investigator for a copy of it and mail it to the University Office for Research Compliance, Box 870127, 358 Rose Administration Building, Tuscaloosa, AL 35487-0127.

I have read this consent form. I have had a chance to ask questions. I agree to take part in it. I will receive a copy of this consent form to keep.

_________________________________________    ____________
Signature of person agreeing to take part in the study   Date

_________________________________________
Printed name of person agreeing to take part in the study

______________________________________________   ____________
Name of (authorized) person obtaining informed consent   Date

E.2: Consent to Participate in a Research Study: Use-based Phase

WHY ARE YOU BEING INVITED TO TAKE PART IN THIS RESEARCH?
You are being invited to take part in a research study about manipulating pitch and volume to convey emotion using a text to speech device. You are being invited to take part in this research study because you currently use an augmentative and alternative communication
device and already have a diagnosis of dysarthria. If you volunteer to take part in this study, you will be one of two people to do so.

WHO IS DOING THE STUDY?
The person in charge of this study is Nure Biane Kassas of University of Alabama Department of Communicative Disorders. She is being guided in this research by Dr. Anthony Buhr.

WHAT IS THE PURPOSE OF THIS STUDY?
By doing this study, we hope to learn whether we can 1) adapt existing speech synthesizing technology to meet the communication needs of individuals who require augmentative and alternative communication devices to supplement speech, 2) the device designed for this study improves the communication abilities of people who use it, and 3) whether users feel the devices improves their communication abilities.

ARE THERE REASONS WHY YOU SHOULD NOT TAKE PART IN THIS STUDY?
If you are under the age of 19 or do not currently use a text to speech device to help you communicate, you should not participate in this study.

WHERE IS THE STUDY GOING TO TAKE PLACE AND HOW LONG WILL IT LAST?
The research procedures will be conducted at The University of Alabama Speech and Hearing Center. You will need to come to the University of Alabama for an initial training session which will last about 2 hours. During this time, you will be instructed on how to use the device to manipulate pitch and volume. You will also be required to attend at least 2 additional sessions during which you will use the device to communicate. Each of those visits will take about 1 hour. The total amount of time you will be asked to volunteer for this study is at least 4 hours over the next 3 months. If you require additional assistance learning how to manipulate the device, you may be required to attend more sessions.

WHAT WILL YOU BE ASKED TO DO?
During the course of this study, you will be asked to use a text to speech device with features that allow you to manipulate pitch and volume to convey emotion. You will be required to attend a training session on how to use the device. After the training session, you will be required to arrive at the Speech and Hearing Center for at least 2 additional sessions during which you will use the device and its features to communicate. If you are still having difficulty manipulating the device, and you request additional assistance, you will be granted additional sessions. The investigator will discuss and arrange times for all sessions with you in advance. You will be asked to both self-report the type and severity of your dysarthria. You will also be asked to voluntarily provide medical records confirming the severity and type of your dysarthria. You may decline if you wish to do so. If you choose to provide the medical records related to your dysarthria diagnosis, the id number will be recorded on the medical records and all identifying information will be blacked out with a permanent marker to protect your privacy. This information will also be stored in a secure location, accessible
only to the study investigator and thesis advisor. Existing medical records or other protected health information (PHI) that is part of your client file at the University of Alabama Speech and Hearing Center will not be accessed or used for any purpose. No other medical records will be requested from you.

**WHAT ARE THE POSSIBLE RISKS AND DISCOMFORTS?**

To the best of our knowledge, the things you will be doing have no more risk of harm than you would experience in everyday life.

**WILL YOU BENEFIT FROM TAKING PART IN THIS STUDY?**

You will not get any personal benefit from taking part in this study. Your willingness to take part, however, may, in the future, help society as a whole better understand this research topic.

**DO YOU HAVE TO TAKE PART IN THE STUDY?**

You will not lose any benefits or rights you would normally have if you choose not to volunteer.

You can stop at any time during the study and still keep the benefits and rights you had before volunteering. Should you choose to stop participating in the study, please inform the investigator as early as possible. You may reach the researcher, Nure Biane Kassas, at (205) 393-1484.

**IF YOU DON’T WANT TO TAKE PART IN THE STUDY, ARE THERE OTHER CHOICES?**

There are no other choices available if you choose not to be in the study.

**WHAT WILL IT COST YOU TO PARTICIPATE?**

The only cost to participate is your time.

**WILL YOU RECEIVE ANY REWARDS FOR TAKING PART IN THIS STUDY?**

You will not receive any rewards or payment for taking part in the study.

**WHO WILL SEE THE INFORMATION THAT YOU GIVE?**

This study is anonymous. That means that no one, not even members of the research team, will know that the information you give came from you.

**CONFIDENTIALITY**

You will be assigned an id number at the start of the study. This id number will be linked to the information you provide throughout the study including the type and severity of dysarthria you have acquired and the answers from your questionnaire. Your real name and
any personally identifiable information will not be associated with the id number. Your consent form will be locked in a room at the University of Alabama Speech and Hearing Center and will be in a separate file from any data collected.

CAN YOUR TAKING PART IN THE STUDY END EARLY?

If you decide to take part in the study you still have the right to decide at any time that you no longer want to continue. You will not be treated differently if you decide to stop taking part in the study.

The individuals conducting the study may need to withdraw you from the study. This may occur if you are not able to follow the directions they give you, if they find that your being in the study is more risk than benefit to you, or if the agency funding the study decides to stop the study early for a variety of scientific reasons.

WHAT IF YOU HAVE QUESTIONS, SUGGESTIONS, CONCERNS, OR COMPLAINTS?

Before you decide whether to accept this invitation to take part in the study, please ask any questions that might come to mind now. Later, if you have questions, suggestions, concerns, or complaints about the study, you can contact the investigator, Nure Biane Kassas at 205-393-1484. If you have any questions about your rights as a volunteer in this research, contact Mrs. Tanta Myles, the Research Compliance Officer of the University of Alabama at 205-348-8461 or toll free at 1-877-820-3066. We will give you a signed copy of this consent form to take with you.

You may also ask questions, make suggestions, or file complaints and concerns through the IRB Outreach website at http://osp.ua.edu/site/PRCO_Welcome.html or email the Research Compliance office at participantoutreach@bama.ua.edu.

After you participate, you are encouraged to complete the survey for research participants that is online at the outreach website, or you may ask the investigator for a copy of it and mail it to the University Office for Research Compliance, Box 870127, 358 Rose Administration Building, Tuscaloosa, AL 35487-0127.

I have read this consent form. I have had a chance to ask questions. I agree to take part in it. I will receive a copy of this consent form to keep.

_________________________________________    ____________
Signature of person agreeing to take part in the study   Date

_________________________________________  
Printed name of person agreeing to take part in the study

______________________________________________   ____________
Name of (authorized) person obtaining informed consent   Date
Appendix F Advertisements for Test Subjects

F.1: Advertisement for Auditory Phase of Study:

CALL FOR VOLUNTEERS

“Real-time Change in Prosodic Aspects of Text Generated Speech”

Nure Kassas, a graduate student at the University of Alabama, is recruiting volunteers to participate in a study about perceptions of intent and emotions in speech. The study will take place at the Speech and Hearing Center at the University of Alabama.

Details of the Study:

- It will require about 10 minutes of your time.
- You will be administered a hearing screening.
- You will listen to and rate your perception of a series of recorded statements/utterances

NOTE: No identifying data will be collected

If you have

1) Hearing within normal limits
2) Are 19 years of age or older

and you would like to participate in the study, please contact Nure Kassas at (205) 393-1484, or send an email to nbkassas@crimson.ua.edu.”

F.1: Advertisement for Use-based Phase of Study:

CALL FOR VOLUNTEERS

“Real-time Change in Prosodic Aspects of Text Generated Speech”

Nure Kassas, a graduate student, is currently recruiting individuals to participate in a study using a newly developed text to speech device that will allow people to communicate. You will be asked to arrive at the Speech and Hearing Clinic at the University of Alabama on three different occasions over the course of four months. The initial session will last about 2 hours, and the remaining two sessions will last about an hour each.
Details of the Study:

- You will be asked to use a text to speech device developed for the purposes of this project to communicate with the study investigator and/or your family members.
- You will be asked to fill out a questionnaire regarding the effectiveness of the text to speech device

NOTE: No identify data will be collected

If you have

3) A diagnosis of dysarthria
4) Have already been using a text to speech device
5) Are 19 years of age or older

and you would like to participate in the study, please contact Nure Kassas at (205) 393-1484, or send an email to nbkassas@crimson.ua.edu.”
Appendix G: IRB Approval Letter

October 23, 2013

Nure Kasses
Department of Communicative Disorders
College of Arts and Sciences
Box 879242

Re: IRB # 13-OR-220-ME, “Real time Change in Prosodic Aspects of Text Generated Speech”

Dear Ms. Kasses:

The University of Alabama Institutional Review Board has granted approval for your proposed research.

Your application has been given expedited approval according to 45 CFR part 46. Approval has been given under expedited review category 6 as outlined below:

6) Collection of data from voice, video, digital, or image recordings made for research purposes.

Your application will expire on October 22, 2014. If your research will continue beyond this date, please complete the relevant portions of the IRB Renewal Application. If you wish to modify the application, please complete the Modification of an Approved Protocol Form. Changes in this study cannot be initiated without IRB approval, except when necessary to eliminate apparent immediate hazards to participants. When the study closes, please complete the Request for Study Closure Form.

Please use reproductions of the IRB approved stamped consent forms to obtain consent from your participants.

Should you need to submit any further correspondence regarding this proposal, please include the above application number.

Good luck with your research.

Sincerely,

[Redacted Name]

Stuart Usdan, Ph.D.
Chair, Non-Medical IRB
The University of Alabama
UNIVERSITY OF ALABAMA
INSTITUTIONAL REVIEW BOARD FOR THE PROTECTION OF HUMAN SUBJECTS
REQUEST FOR APPROVAL OF RESEARCH INVOLVING HUMAN SUBJECTS

I. Identifying Information

Principal Investigator: Nan Blain Konen
Second Investigator: Dr. Anthony Baker
Third Investigator: 
Department: Communication Disorders
College: Arts and Sciences
University: University of Alabama
Address: 901 Hargrove Rd, Apt 6A
Tuscaloosa, AL 35401
Fax: 205-348-1645
Telephone: 205-348-1643
E-mail: rblaine@crimson.ua.edu

Title of Research Project: Real-time Change in Prosodic Aspects of Text Generated Speech

Date Submitted: 09/16/2013
Funding Source: N/A

Type of Project: □ New □ Revision □ Renewal □ Completed □ Exempt

Title or Investigator's name: [Redacted]

II. NOTIFICATION OF IRB ACTION (to be completed by IRB):
Type of Review: □ Full board □ Expedited

IRB Action:
□ Rejected Date: [Redacted]
□ Titled Pending Revisions Date: [Redacted]
□ Approved Pending Revisions Date: [Redacted]
□ Approved this proposal complies with University and federal regulations for the protection of human subjects

Approval is effective until the following date: 10-23-13

Research protocol (dated [Redacted])
Informed consent (dated [Redacted])
Recruitment materials (dated [Redacted])

Approvalsignature: [Redacted] Date 10-23-13