

Undergraduate Research Academy (URA)

Application for Student Fellowship

SEND TO CAMPUS BOX 1300
BY NOON, WEDNESDAY, MARCH 16, 2004
(Please type)

1. Name Mary Witte 2. ID# _____
3. Local Address _____
4. Local Telephone _____ 5. e-mail _____
6. Academic Major Speech-Language Pathology 7. Hours Completed 82 8. GPA (4-point scale) 3.854
9. Nominated by Dr.Melanie (Richter) Brimer
10. Reviewed by: (Please print)

Faculty Mentor Dr.Melanie (Richter) Brimer Department Chair Dr. Bill Searcy

11. Title of Research Project The Effects of Linear Predictive Coding Analysis/Resynthesis of the Dysarthric Speech for Persons with Amyotrophic Lateral Sclerosis

Please send the original plus 10 copies of the proposal, including budget justification and cover page.

12. Budget Summary Total: \$ 800
Commodities: \$ N/A Contractual Services: \$ 800.00
Travel: \$ N/A Other (specify): \$ N/A

13. Valid nominations must carry all of the following approval signatures with dates:

a. Clearances (as appropriate):
Animal Care N/A Human Subjects X-See Attached Toxic Waste N/A

b. Student Mary Witte's Signature

c. Faculty Mentor Melanie Brimer's Signature

d. Department Chair L. William Searcy's Signature

e. Dean of College/School Curt Lox's Signature

f. Undergraduate Research Academy Interview Cindy Scarsdale's Signature

_____ SIUE Undergraduate Assessment -&- The Undergraduate Research Academy _____
Box 1300 voice: 618/650-2640 e-mail: deder@siue.edu FAX: 618/650-3633

THE UNDERGRADUATE RESEARCH ACADEMY
OFFICE OF UNDERGRADUATE ASSESSMENT & PROGRAM REVIEW
Cover Page *(Please type)*

STUDENT Nellie J. Shaul MENTOR Dr. Melanie (Richter) Brimer

PROJECT TITLE The Effects of Linear Predictive Coding Analysis/Resynthesis of the Dysarthric Speech for Persons with Amyotrophic Lateral Sclerosis

ABSTRACT: The abstract is a brief but comprehensive summary of the contents of the proposal in plain language, approximately 150 words. Readers receive their first impression of the flavor of the topic from this abstract. The information in the abstract needs to be concise, well organized, self contained, and understandable to persons outside the discipline.

Persons with amyotrophic lateral sclerosis (ALS) face a degenerative neuromuscular disease that severely affects their speech. While interventions such as augmentative and alternative communication (e.g. communication devices with synthetic voice output) are helpful, few strategies exist to enhance the speech of individuals with ALS. The overall goal of this line of research is to use computer technology known as linear predictive coding (LPC) to make the speech of persons with ALS more intelligible. The purpose of this study is to investigate the effect that LPC analysis/resynthesis has on speech for persons with ALS. Specifically, the proposed study will determine whether or not significant distortion is added and whether or not specific error sounds are perceived. Since LPC has only been used on disordered speech to a limited degree, it is necessary to see how these distortions affect the speech of persons with ALS. This proposed study will be a critical step in developing speech enhancement technology for persons with ALS based on LPC analysis/resynthesis technology.

Upon submitting this proposal, I verify that this writing is my own and pledge to fulfill all of the expectations of the Undergraduate Research Academy to the best of my abilities. I understand that failure to do so may result in return of fellowship money to the University and forfeiture of academic credit and honors recognition.

Mary Witte's Signature

Signature of the Student

I am able, willing, and committed to providing the necessary facilities and to take the time to mentor this student during this project. I verify that this student is capable of undertaking this proposed project.

Melanie Brimer's Signature

Signature of the Faculty Mentor

This project is within the mission and scope of this department, and the department fully supports the faculty mentor and student during this venture.

L. William Searcy's Signature

Signature of the Department Chairperson

I testify that all necessary research protocols (human, animal, toxic waste) have been fulfilled, and I support this proposed faculty-student scholarly activity as within the mission of the College/School.

Curt Lox's Signature

Signature of the Dean of the College/School

Research Proposal

Introduction and Significance

Amyotrophic lateral sclerosis (ALS), often referred to as "Lou Gehrig's disease," is a degenerative motor neuron disease that attacks nerve cells in the brain and spinal cord. The cause of ALS is unknown, and the average age of onset of the disease is 58 years old (11). ALS affects the function of nerves and muscles of approximately 30,000 Americans at any given time and significantly affects the speech of those living with the disease (1). The speech characteristics of individuals with ALS include disordered articulation of both consonants and vowels, extremely slow production of words in very short phrases, marked hypernasality and a harsh voice, and disruption of prosody (melody) of the voice (10). Consequently, the speech of persons with ALS is often highly unintelligible (difficult to understand), and, thus far, there are no devices available that will improve the intelligibility of their speech. Speech therapy tends to further fatigue the muscles of individuals with ALS and may possibly contribute to acceleration of neurological deterioration. Therefore, direct speech intervention is not deemed a successful treatment (4). While interventions such as augmentative and alternative communication devices (e.g. communication devices with synthetic speech output) are helpful, few strategies exist to enhance the actual speech of individuals with ALS.

The goal of this line of research is to use computer technology to manipulate acoustic properties of recorded speech in a computer in order to make the speech of persons with ALS more intelligible and clear. This proposed device would take advantage of the fact that people with ALS already know how to speak and can already make close approximations of the desired speech sounds. So rather than using a device with synthetic voice, people with ALS would be able to enhance the quality of their own unintelligible speech using future speech enhancement technology with techniques being explored through the present line of research. Theoretically, this proposed device would be able to recognize the disordered speech better than human listeners and would then be able to enhance it. In addition, this speech enhancement device would be faster than communicating with augmentative and alternative communication, because it would use the person's own speech rather than using a synthetic text-to-speech format.

In order to make the speech of persons with ALS more intelligible, the speech must be manipulated to improve the clarity. The speech manipulation may be accomplished by using a computer technology known as linear predictive coding (LPC), which uses mathematical calculations to separate the sound into its two components: the source of the sound (e.g. the vibration of the vocal folds) and the filter that modifies the source (e.g. the vocal tract or the structures above the level of the vocal folds). The filter shapes the source into what we recognize as speech by dampening certain frequencies in the source and intensifying others. Linear predictive coding works through a process of analysis (separation) and resynthesis (recombination) of the two components of a speech signal. It first analyzes the speech signal by estimating the formant frequencies based on an algorithmic calculation. Removing the filter from the speech signal makes it possible to then estimate the source of the speech signal. Then, both the source (vibration or buzzing) and filter (vocal tract characteristics) of the speech signal can be manipulated in order to change the quality or characteristics of the speech signal. When these two components are resynthesized (put back together), a new speech signal exists that is different from the natural speech signal (8).

The proposed study will be investigating the effect that LPC has on the analysis and resynthesis of speech for persons with ALS. Specifically, the study will be comparing listeners' perceptions of the natural speech (i.e. original recordings) of persons with ALS and perceptions of their speech that has gone through the process of LPC analysis/resynthesis. This research is a vital component to the advancement of the overall mission of developing a device to help improve the speech intelligibility for persons with ALS. Previous research by Richter (2002) has explored the effect of manipulating the acoustic properties of speech for persons with ALS by analyzing both natural speech and manipulated speech (i.e. speech where the formant frequencies have been replaced with typical formant frequencies) (8). In order to continue to move forward toward the goal of developing a speech enhancement device for people with ALS, it is necessary to know the effect that LPC alone has on the speech of people with ALS. To do this, speech of people with ALS that has been analyzed and resynthesized without being manipulated must be compared to the natural speech of the people with ALS. In the proposed study, we will examine what types of changes take place in the speech of people with ALS after it has gone through the process of LPC analysis/resynthesis. This study will provide information about the effects of LPC analysis/resynthesis on the perception of the speech of persons with ALS, including, but not limited to, the specific types of sounds that LPC may be more likely to interfere with and/or enhance. It should also identify whether or not certain sounds seem to be omitted or added and any distortion that is added (e.g. popping or clicks).

Literature Review

Linear predictive coding is one type of technology that makes the analysis and resynthesis of speech possible. Ideally, LPC would compute speech signals precisely without any measurable reduction in intelligibility or naturalness. However, LPC uses algorithms that cannot precisely compute the aspects of a speech signal. Additionally, it is important to note that this type of technology uses algorithms based on typical, native English speakers. It has also been evaluated almost exclusively with respect to its effects on the speech of speakers with typical English speech production abilities (2).

Only recently has evaluation begun with languages having characteristics that are very different from typical spoken English. Previously, few, if any, evaluations have been available on how analysis/resynthesis technology affects the speech of persons who have a disorder of voice and/or speech (2). Previous research by Qi (1990) demonstrated that LPC can be used separate the features of the vocal tract filter and source of vowels produced by alaryngeal speakers (i.e. people who have had their voice box removed)(6). Qi (1995) expanded the research to show that the overall quality of words spoken by alaryngeal speakers can be enhanced using LPC analysis/resynthesis (7).

Another researcher studying the manipulation of disordered speech has looked at the effects of speech manipulation on the speech of people who are deaf. In this study, the acoustic property of fundamental frequency was manipulated in an attempt to make the speech more intelligible (3). Further research on the effect of LPC manipulation was done by Shuster (1996) in a study on children with incorrect productions of "r". In this study, researchers analyzed the incorrect speech productions and then manipulated the filter of the speech to create more typical sounding "r's". After resynthesis, listeners judged the edited sound to be a correctly produced /r/, thus indicating that LPC manipulation and synthesis can be a useful tool. However, Shuster found that it was difficult to obtain high quality synthesis of nasalized speech (e.g. hypernasal or sounds produced in error through the nasal cavity), suggesting a possible limitation of LPC (9).

In a study by Monsen and Engebretson (1983) on the accuracy of LPC to predict formant frequencies, they found that overall LPC analysis was equal to or better than human analysis of speech. However, the researchers discovered that when the speech was nasalized, LPC identification of formant frequencies was less effective (5). While some studies have used LPC on disordered speech, few studies have used LPC on the speech of people with ALS.

In a study by Richter (2002), LPC analysis/resynthesis was used to manipulate acoustic properties in the speech of people with ALS. Specifically, vowel formant frequencies used by listeners to distinguish vowels were manipulated to the formant frequency values of typical male speakers. Results indicated that speech manipulation using LPC analysis/resynthesis improved vowel intelligibility for one of the five subjects. The vowel intelligibility for this speaker improved from 39% intelligibility to 71% intelligibility. Distinguishing characteristics for the speaker were that he had the lowest vowel intelligibility and that his second formant frequency was more severely affected than for any of the other subjects (8). These results demonstrate the potential for this line of speech manipulation research. In order for this line of speech manipulation research to continue, it is necessary for the effects of LPC analysis/resynthesis on the speech of persons with ALS to be thoroughly understood, which is the primary purpose of this proposed study.

Hypothesis

This proposed research study will examine the effects of LPC analysis/resynthesis on the dysarthric speech of persons with ALS. It will provide information regarding how people with typical hearing perceive the speech of people with ALS after the speech has gone through the process of LPC analysis/resynthesis. Due to their neurological condition, people with ALS are likely to have hypernasal speech, and the literature has reported difficulty using LPC with nasal consonants. Therefore, we hypothesize that there will be confusion for the stop consonants (i.e. “p”, “b”, “t”, “d”, “k”, “g”) with the nasal consonants (i.e. “m”, “n”, and “ng”). Additionally, since LPC analysis/resynthesis is based on mathematical calculations, and based on previous research, we further hypothesize that LPC analysis/resynthesis will add distortion to the speech (e.g. popping and clicking). Since LPC has only been used on disordered speech to a limited degree, it is necessary to see how these distortions affect the speech of persons with ALS. This proposed study will be a critical step in developing speech enhancement technology for persons with ALS based on LPC analysis/resynthesis technology.

Materials, Procedures, and Time Line

In order for this line of speech manipulation research to continue, it is necessary for this study to be conducted. This research project is an essential component of the overall research goals of developing a speech enhancement device for people with ALS. My tasks include further developing my literature review, running subjects, transcribing words phonetically, analyzing data, and meeting with my mentor on a weekly basis. My mentor and I plan to submit this project jointly and present together for the American Speech Language Hearing Association (ASHA) Convention in 2005.

In this study, subjects will be no more than 100 males and females of varying race/ethnicity who are between the ages of 19 to 45 and who have typical hearing. For the study, we will use data from 50 subjects, but we may require more subjects (up to 100) due to the fact that things may happen outside of our control that will require us to have to throw out the data. For example, if one a sound

file does not play or a subject does not pass a hearing screening, we would have to throw out the data from that particular subject for our study. The subjects will not have extensive experience or be familiar with persons with amyotrophic lateral sclerosis (ALS or Lou Gehrig's Disease), and they will not be familiar with persons who have similar speech disorders causing dysarthria (e.g. cerebral palsy). The subjects will be recruited through written announcements posted around campus. Additionally, announcements will be made in classes other than the faculty member's classes.

Subjects will be seated in an audiology soundproof booth. They will be given a brief hearing screening. They will then be asked to fill out demographic information. The demographic information will include: name, birth date, native language(s), ethnicity, gender, and other questions related to their experiences with speech disorders. The subjects will be asked to listen to ten training words. After each word, they will be asked to type the word that they heard on a computer, even if they thought it sounded like a nonsense word. They will use a guide indicating what symbols to type for different vowel sounds. The subjects will be asked to press enter when they are done typing the word so that their responses can be timed. They will then be asked to indicate how confident they are that they got the correct answer on a scale from zero to seven. Upon selecting a confidence level, the next word will be played. If they complete the training activity and do not have any questions, they will then follow the same procedures for 140 more words. If they have questions about the procedures, those will be answered at that time. If they need more practice, the training portion will be repeated. The subjects will be given a break after 5 words during the training portion and after every 35 words (for a total of 3 breaks) during the experimental portion. The experiment should last approximately one hour.

Materials required include a laptop computer, the Alvin computer program, high quality headphones, audiometer, an SPSS statistical package to analyze the results, previously recorded word lists for natural and unmanipulated speech from persons with ALS, and a method of extracting data from the computer to distribute to the transcription team for analysis (e.g. flash key USB pen drive). The laptop computer will be used to play sounds for the subjects. These sounds are part of the previously-recorded word lists spoken by persons with ALS. The Alvin computer program will be in a format that enables the subjects to type in their responses and click on their confidence level (0-7). The program plays the stimuli in the desired order and times the subjects' responses. The subjects will use the high quality headphones to listen to the word lists. We will use the audiometer for hearing screenings for the subjects to ensure they meet the criteria for typical hearing. The flash key USB pen drive will allow us to remove the data on a frequent basis from the research laptop in the laboratory to the computers used to distribute and analyze data (i.e. computer for compiling data to give to the transcription team, computer with SPSS). Additionally, the flash key USB pen drive will be much cheaper over the long run as compared to recordable cds. The cds each cost one dollar and are single use. The flash key USB pen drive will be completely reusable and will be a one-time purchase. Additionally, it will be very portable and easy to use. It will also be used for projects in the future.

The data from the subjects will be in a typed format, which we will transcribe along with our transcription team. We will transcribe the words phonetically with the International Phonetic Alphabet commonly used in speech-language pathology. We will use the SPSS statistical package to see statistically significant differences. For example, we will compare the natural and unmanipulated speech intelligibility. We will also compare the confidence levels for each and the response times for each. We will also look at differences between vowels and consonants across subjects. In addition, we will utilize error matrices, which have phonemes (speech sounds) across the top and down the side.

Based on our results, we can fill out the matrices by putting the actual sound that was said in the recording and mark in the other column the sound that was perceived by the listeners. This will allow us to distinguish which sounds are being confused and help us to classify the speech sounds of people with ALS that are most affected by LPC analysis/resynthesis.

Beginning in early September, we plan to start this project by creating a phonetic transcription team of a few students in the speech-language pathology department. The purpose of this team will be to transcribe the speech and evaluate the reliability of the results. Specifically, they will be looking at the subjects' typed responses and writing the sounds using the International Phonetic Alphabet. Additionally, the transcription team will allow other undergraduate students to be involved in research and to use phonetic transcription in a practical application. So, they will benefit from the experience of being involved in research and using phonetic transcription in a practical application, but they will not be compensated otherwise. We will also develop subject folders, prepare the computer program, and plan to test out equipment by running our pilot subjects. In the process, I will be further developing my literature review. In October we will begin recruiting and running subjects. We will also transcribe their responses with the transcription team. This will be ongoing until approximately January. In November 2004, I plan to attend the American-Speech Language Hearing Association (ASHA) national convention. This will give me the opportunity to see other researchers present and learn more about current research in my field of interest. In January, we will submit the URA progress report, as well as begin analyzing results. We will draw conclusions and work on writing the final report between the months of February and May. March and April are set aside for creating our poster and getting ready to present in April. In April, I will also apply to present at the ASHA 2005 national speech-language pathology and audiology convention.

Time Line 2004-2005									
	September	October	November	Decemeber	January	February	March	April	May
Create Transcription Team	X								
Develop Subject Folders	X								
Prepare Computer Program	X	X							
Background Research	X	X							
Test Equipment/Run Pilot Subjects	X	X							
Recruit Subjects		X	X	X	X				
Run Subjects		X	X	X	X				
Transcribe Responses		X	X	X	X				
Attend ASHA 2004			X						
Progress Report					X				
Analyze Results					X	X			
Draw Conclusions/Write Report						X	X	X	X
Create Poster							X	X	
Present								X	
Apply for ASHA 2005								X	
Submit Final Report									X

References

1. ALS Association. Internet. March, 2004. <http://www.alsa.org>.
2. Jamieson, D. G, et. al, 2002. Interaction of Speech Coders and Atypical Speech II: Effects on Speech Quality. *Journal of Speech, Language & Hearing Research*, 45 (4): 689-700.
3. Maassen, B. & Povel, D., 1984. The Effect of Correcting Fundamental Frequency on the Intelligibility of Deaf Speech and its Interaction with Temporal Aspects. *Journal of the Acoustical Society of America*, 76 (6), 1673-1681.
4. Mathy, P., 2000. Amyotrophic Lateral Sclerosis. *ASHA Leader*, 5 (2): 6-9.
5. Monsen, R. B. & Engebretson, A. M., 1983. The Accuracy of Formant Frequency Measurements: A Comparison of Spectrographic Analysis and Linear Prediction. *Journal of Speech and Hearing Research*, 26, 89-97.
6. Qi, Y., 1990. Replacing Tracheosophageal Voicing Sources Using LPC Synthesis. *Journal of the Acoustical Society of America*, 88 (3): 1228-1235.
7. Qi, Y., Weinberg, B. & Bi, N., 1995. Enhancement of Female Esophageal and Tracheoesophageal Speech. *Journal of the Acoustical Society of America*, 98 (5): 2461-2465.
8. Richter, M. (2002). *The Digital Manipulation of Dysarthric Speech for Persons with Amyotrophic Lateral Sclerosis*. A doctoral dissertation. University of Nebraska-Lincoln.
9. Shuster, L. I., 1996. Linear Predictive Coding Parameter Analysis/Resynthesis of incorrectly produced /r/. *Journal of Speech & Hearing Research*, 39 (4), 827-832.
10. Yorkston, K. M., Miller, R. M., & Strand, E. A., 1995. *Management of Speech and Swallowing in Degenerative Diseases*. Tuscan, AZ: Communication Skill Builders, 1-55.
11. Yorkston, K. M., Beukelman, D. R., Strand, E. A., & Bell, K. R., 1999. *Management of Motor Speech Disorders in Children and Adults (Second Edition)*. Austin, TX: Pro-Ed, 156-167.