Preliminary Study Grant: Using speech analysis to explore the clinical effectiveness of stress and depression assessment in adults with aphasia

Principle Investigator: Jacqueline Laures-Gore, PhD, CCC-SLP¹

Co-Investigators: Elliot Moore, PhD²; Scott Russell, MS, CCC-SLP³

¹Georgia State University, ²Georgia Institute of Technology, ³Grady Memorial Hospital

Abstract

Stroke is the leading cause of long-term disability in adults in the United States. Aphasia is a devastating disability among stroke survivors affecting over one million people in the United States. Aphasia is an acquired communication disorder resulting from brain damage impairing an individual’s ability to use, produce, and comprehend language. One critical barrier to developing effective therapies for aphasia is a lack of reliable instruments for diagnosing emotional states and disorders within these adults after they have acquired aphasia. The current study proposes a highly innovative and novel approach to identifying a diagnostic marker for emotional states (stress and depression) that can be captured in the speech signal. The outcome of this line of inquiry will begin to have an impact on healthcare services by providing healthcare professionals with a tool to more accurately diagnose stress and depression in language impaired adults. This project will leverage the work by Dr. Moore's lab in speech signal analysis pertaining to stress and depression with Dr. Laures-Gore's research investigating the emotional states of adults with aphasia to explore transformative research in the development of objective assessments for stress and emotion in adults with aphasia. This one year project will allow us to begin testing the hypothesis that depression and stress can be identified via the speech signal in adults with aphasia and provide foundation data for future grant applications to external funding sources.
1. Introduction

Stroke is the leading cause of long-term disability in adults in the United States [1]. Aphasia is a devastating disability among stroke survivors affecting over one million people in the United States. NIH recognizes aphasia as an "impediment to health literacy and quality of life" (NIH PAR-10-133, PAR-10-134, PAR-10-135). Aphasia is an acquired communication disorder resulting from brain damage impairing an individual’s ability to use, produce, and comprehend language [2]. Aphasia is a language disorder, not a speech disorder. One critical barrier to developing effective therapies for aphasia is a lack of reliable instruments for diagnosing emotional states and disorders within these adults after they have acquired aphasia. The long-term goal of this project is to address this critical barrier and improve healthcare services by identifying a diagnostic marker for emotional states (stress and depression) that can be captured in the speech signal. Quantification of the acoustical measures of the speech signal to emotional markers circumvent the difficulties encountered in assessment of emotional states in adults with aphasia because intact language is not required. This highly innovative study will investigate a quantifiable means of measuring an emotional state while avoiding the methodological problems of assessing perceptions of psychosocial dimensions in adults with aphasia such as the lack of standardized questionnaires or rating scales [3]. Self-report questionnaires sensitive to linguistic burden are too few in number and represent a challenge to assessing emotional states in this population. This project will be a collaboration between clinical speech-language pathology (Dr. Laures-Gore, PI; Scott Russell, Co-I) and computer engineering (Dr. Elliot Moore, Co-I). Combining expertise from these two very different, yet complimentary fields will advance healthcare services.

Aphasia, Stress, and Depression. It is well accepted that depression and stress are closely related in the typical adult population [4-7]. Stressful life events are often precursors to depression [6, 8]. More importantly, perception of stress appears to be a critical variable affecting the development of depression. In fact, greater perceived stress is associated with more severe depression in patients with a clinical diagnosis of depression [9]. As recently discovered in Dr. Laures-Gore’s lab, perceived stress and depression are closely linked in adults with aphasia [10]. Relatedly, in a series of pioneering experiments representing the few studies examining stress in adults with aphasia, Laures-Gore and colleagues have used both stress biomarkers and subjective stress scales to investigate stress in adults with aphasia. Overall, results indicate that adults with aphasia perceive more stress than healthy adults, greater word productivity is linked to heightened levels of salivary cortisol, and they do not demonstrate typical cortisol reactivity to a linguistic stressor [11-13]. Although these results begin to provide a clearer understanding of stress in aphasia, they are fraught with methodological concerns. Most importantly, the use of salivary cortisol as a marker of stress in post-stroke patients has not been validated. The neurologic changes that accompany aphasia may affect the robustness of cortisol as a biomarker of stress in post-stroke patients [14]. Additionally, the use of stress scales with inadequate validation for use with adults with aphasia is problematic (although modifications have been made to some of these questionnaires to make them more aphasia friendly). The use of the speech signal would offer a new technique that can avoid these issues. Because depression is a behavioral response to stress, evaluating stress in adults with aphasia may be helpful in identifying those most at risk for developing depression.

Post-stroke depression is a controversial issue, especially in regards to diagnosis [15]. Because of this controversy, the percentage of stroke patients identified as depressed presents a rather wide range of 10% to 60% of stroke patients [16-17], although a recent meta-analysis indicates that one third of stroke patients have depression [18-19]. The range of depression diagnoses is
indicative of the complexity of this disorder and potentially its under- or over-diagnosis in this population. Post-stroke depression evolves as patients enter the chronic phase and reflects their reaction to the impairments associated with their stroke [20-21]. Degree of functional disability and presence of motor, cognitive, or communication impairments have been identified as contributors to post-stroke depression [22-23], although there appears to be a great deal of debate regarding causes of depression [24]. As Laures-Gore and DeFife recently discussed, most studies of post stroke depression exclude adults living with aphasia; thus, an understanding of the contribution of aphasia to post-stroke depression is limited [25]. Only a handful of studies have included patients with aphasia in their sample with mixed results as to the effect of aphasia on developing post-stroke depression [26-28]. Adults with aphasia are excluded due to comprehension and expression disabilities that many questionnaires cannot accommodate.

The current study aims to identify diagnostic markers for emotional states (stress and depression) that can be captured in the speech signal. In previous work, Moore and colleagues examined acoustic markers of depression in non-neurologically impaired adults [29-34]. The investigators believe that analysis of the speech signal should provide an opportunity for identifying objective measures of stress and depression in adults with aphasia. Because aphasia is a language disorder, many commonly used questionnaires cannot be completed by adults with aphasia. The SADQ-10 was developed for diagnosing depressive symptoms in adults with aphasia, but this instrument must be completed by someone close to the adult with aphasia. Concerns regarding the validity of biomarkers of stress and stress questionnaires remain. Because depression occurs in adults with aphasia and depression is closely linked to stress, development of tools to diagnose depression are needed. In turn, it is predicted this will improve therapy outcome and compliance. Because prolonged, chronic stress can lead to mental and physical disease it is important to develop a better understanding of stress in this population. If greater knowledge about this topic is acquired, then better treatments can be developed to help those living with aphasia cope with stress and avoid responses such as depression.

Speech Signal Processing. While there is an extensive body of literature on the feature analysis and computational models for the recognition of stress and depression in persons without aphasia, there is an analogous lack of such research for persons with aphasia. This lack of research is not surprising considering that aphasia affects the articulation (non-fluent aphasia) and/or language coherency (fluent aphasia) of persons who suffer from it. Therefore, a lot of the analysis related to speech in adults with aphasia involves the choice of words and phonemes in naming tasks or connected speech [35-40] with little emphasis on the acoustic properties of the speech produced. One recent study showed that the prosody of subjects with Broca (non-fluent) and Wernicke (fluent) aphasia showed statistically significant differences in prosody stability (i.e., jitter, shimmer, etc.) during the phonation of a sustained vowel compared to speech from a non-aphasic subject [41]. While this study was not correlated with any analysis of the affective state of the subject, it does highlight that there are measurable acoustic differences that can exist between different types of aphasia that should be considered in any analysis of the speech.

The detection of stress and depression in the voice of adults without aphasia embodies a large volume of work on the study of affect (emotion) in the voice. Several literature surveys over the years [42-48] have provided overviews and updates on work related to the creation and analysis of speech emotion databases. Each of these surveys have highlighted state-of-the-art techniques and features used for the detection of emotion in the voice using prosodics (e.g., pitch, speaking rate) and spectral features related to speech content (i.e., formants, cepstral coefficients). Measurements of the fundamental frequency (i.e., the quasi-periodic rate of vocal
fold vibration), speaking rate and energy in the voice and have found great use in a variety of studies related to emotion and stress recognition [49-56]. The perceptual qualities of depression in the voice have been most commonly studied with regard to prosodic and vocal tract perturbations [29, 57-66].

Additionally, features related to the glottal waveform and signal profiles extracted from the Teager Energy Operator have also provided additional insight into the correlation of acoustic measures of speech and the affective state of the speaker. The glottal waveform (or voice source) relates to the profile of the volume-velocity of airflow through the glottis from the lungs. Features related to the glottal source are often excluded from voice analysis due to the difficulty in extracting accurate representations directly from the acoustic speech signal without concurrent voice recordings with auxiliary devices (e.g., electroglottographs (EGGs)). Without these auxiliary devices, information on the motion of the vocal fold dynamics can only be measured through the evaluation of the glottal waveform itself which must be extracted directly from the speech signal using glottal waveform extraction algorithms [67-79]. Studies by Moore and others have noted the impact of including glottal features in the analysis of depression in the voice of adults without aphasia [32, 80-82]. The Teager Energy Operator (TEO) [80, 83-88] has also provided a usable base for features in the detection of stress and emotion in individuals without aphasia. While traditional speech analysis is based on linear discrete model (i.e., each component is represented with separable Linear Time-Invariant (LTI) filters), the TEO is based on equations that allow for non-linear airflow during speech production. The features based on TEO have generally been shown to improve the performance of traditional linear features of speech production when combined for analysis of stress. A significant part of this grant will be to build on the type of transformative research that will provide for the acoustic characterization of speech from adults with aphasia in relation to emotion and stress that currently does not exist.

2. Specific Aims

The outcome of this line of inquiry will begin to have an impact on healthcare services by providing healthcare professionals with a tool to more accurately diagnose stress and depression in language impaired adults.

Aim 1: Test the hypothesis that depression and stress can be identified via the speech signal in adults with aphasia. Our hypothesis predicts that markers of depression and speech are contained within the extracted acoustic profiles of the speech signal and can be used to detect elements of depression and stress in adults with communication disorders. Speech samples will be collected from subject interviews during a picture description task. These samples will then be subjected to an extensive set of state-of-the-art feature extraction algorithms that will provide acoustic profiles for machine learning tasks in the classification of various levels of depression and stress that may exist in the subject. The levels of stress and depression in each subject will be evaluated through self-reports by the subject and reports by their caretakers. Machine learning will be applied to correlate the objectively measured acoustic features of the subject's speech with the reported levels of stress and depression to establish a baseline of markers to be examined in future studies.

Aim 2: Submitting grants for external funding. This foundation data will support a grant application to NSF and NIH. This project is the first to apply signal processing to the speech of language impaired adults with the final goal of identifying stress and depression. This novel and innovative approach unites computer engineering and speech-language pathology making the project attractive to either NSF or NIH. We recently submitted a similar project to NSF through
the Smart-Health mechanism and believe that NIH-NIMH or NIH-NIDCD will be another funding source.

3. Research Methods
One group of stroke patients with aphasia will be included (N = 20). The group will be comprised of left hemisphere stroke patients with aphasia greater than 1 month post-onset of stroke. This time post-onset will allow for medical stabilization of the participant. All types and severities of stroke will be included as there is currently no evidence that stress responses differ between aphasia types or severity. Participants will be over 18 years of age and report English as their native language. Participants will be recruited from Grady Memorial Hospital and the Georgia State University Speech & Hearing Clinic. Refer to Human Subjects for inclusion/exclusion criteria.

All participants and their caretakers will be required to come to the Georgia State University Aphasia and Motor Speech Disorders Laboratory (Dr. Laures-Gore’s lab) to participate in assessments of aphasia severity as well as self-reports on stress and depression.

**Experimental Manipulations/Measures and Outcomes.** All stroke patients will be assessed in the Georgia State University Aphasia and Motor Speech Disorders Laboratory with the Western Aphasia Battery-R (WAB-R) [89] to classify for type and severity of aphasia. The WAB is a standardized, comprehensive language assessment of aphasia, which has high reliability and validity [90]. Embedded in this comprehensive test is a story-description yielding a language sample that will be audio recorded. It is expected that at least 2-3 minutes of speech will be collected per participant. Audio will be captured at a minimum sampling rate of 16 KHz (16-bit). High quality audio equipment will be provided by Dr. Moore’s lab. Audio recording will be designed to minimize environmental noise and to compensate for any unique acoustical aspects of Dr. Laures-Gore’s lab.

**Subjective Measures of Stress, All Groups.** Subjective Measures of affect will be measured using the Self-Assessment Manikin (SAM) [91], a non-verbal assessment technique designed to measure valence and arousal associated with a person’s affective reaction to different kinds of stimuli. The assessment consists of two sets of pictorial depictions of line drawn figures expressing different emotions, five in each set, the first of which is a positive-negative scale (valence) and the second an excited-calm scale (arousal). Participants will complete a SAM assessment before and after administration of the WAB-R.

**Subjective Measure of Depression.** The Community Stroke Aphasic Depression Questionnaire-10 [92] will be completed by a caregiver/friend/spouse of each aphasic patient. This 10-item observer-rated questionnaire was developed specifically to assess depressed mood in individuals with aphasia. Observers are asked to rate ten different depression-associated behaviors (e.g., crying, displays of anger, restlessness) according to frequency (0 = Never, 3 = Often), with higher scores indicating lower mood.

**Feature Extraction.** Figure 1 provides an overview of the proposed study indicating how the assessments will be combined with state-of-the-art speech feature extraction, correlation analysis and machine learning. Each participant will provide speech recordings that will be segmented and divided based on the voiced and unvoiced sections of the speech (unvoiced will not be analyzed in this study as the information of interest is primarily in the voiced portions of speech production). The assessments will be used to create categories for correlation analysis and the construction of machine learning algorithms.
4. Anticipated Results

We predict that the discrete speech markers of stress and depression found in non-neurologically involved adults can easily be extracted from the speech signal of adults with aphasia. The identification of these markers will lead to more accurate diagnosis of stress and depression in these patients by circumventing the pitfalls of questionnaires. This is a first-step in development of a computer-based tool that can eventually be used by healthcare professionals. We anticipate collecting enough foundation data for grant submissions to NSF and NIH.

5. Budget and Justification

Two masters level GRAs will be supervised by Dr. Laures-Gore and will aid in recruitment, data collection, and data analysis ($5,070 Per GRA for one year). Participants will be paid $100 for their time (20 participants, $2000). Parking at GSU will be provided ($10 per participant, one visit to campus, $200). One GRA will be supervised by Dr. Moore and will conduct the speech analysis ($10,787 for one PhD level GRA funded over a 3-month period). Scott Russell will aid in recruitment of participants and will not be requesting funds. Both Dr. Laures-Gore’s and Dr. Moore’s labs own all necessary equipment (aphasia assessments, signal processor) to conduct this study with the exception of audio recording equipment ($1873). Total costs: $25,000
References


