ECRONICON OPEN ACCESS

EC PSYCHOLOGY AND PSYCHIATRY
Research Article

Approaching the Clinical Effectiveness of Stress and Depression Assessment in Adults with Aphasia through Speech Waveform Analysis and Medical management

Pushan Kumar Dutta[1]* and Arunima Roy[2]

[1]*School of Engineering and Technology, Amity University Kolkata, India*

[2]*Practioner and Speech Therapist, India*

***Corresponding Author:** Pushan Kumar Dutta, School of Engineering and Technology, Amity University Kolkata, India.

Received: October 13, 2019; **Published:** October 25, 2019

Abstract

Approaching the clinical effectiveness of stress and depression assessment in adults with neurogenic stuttering through speech waveform analysis and medical management is directed to explore the language disorder associated with speech dysfluency in language using automated analysis of speech language impairment. Neurogenic stuttering is an acquired communication disorder resulting from TBI, brain damage impairing an individual's ability to produce speech and language it may or may not be associated with language comprehension. One critical barrier to developing effective therapies for neurogenic stuttering is a lack of reliable instruments for diagnosing emotional states and disorders within these adults after they have acquired neurogenic stuttering.

Keywords: *Stress; Depression; Aphasia; Speech Waveform Analysis*

Abbreviations

SLP: Speech Language Processing; TEO: Teager Energy Operator

Introduction

The aim of this study was to evaluate the available evidence about the possible causal association between vocal tics, Psychological basis of speech and exposure to Speech waves. 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. Speech synthesis is now developed using voice synthesis and similar features. The brain is the coordinating point for speech and associated symptoms. However, we cannot develop a brain mapping system for the purpose of detecting the speech deficiencies as the body system responds on the psychological aspects than the physiological aspects in patients. It is necessary to identify the fuzziness of the speech using clinical analysis and also identify the rate of speech from the voice data spectrum to diagnose the effect of speech dysfluency in neurogenic stuttering. 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 with little emphasis on the acoustic properties of the speech produced. One recent study showed that the prosody of subjects with non-fluent and 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. 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 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 studies related to emotion and stress recognition. The perceptual qualities of depression in the voice have been most commonly studied with regard to prosodic and vocal tract perturbations. Additionally, features related to the glottal waveform and signal profiles extracted from the Teager Energy Operator [1,2] 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. The Teager Energy Operator (TEO) 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.

Materials and Methods

Stuttering has received a variety of interpretations that present a depiction of speech fluency, speed and rhythm dysfunction caused by discoordination of respiratory, phonetic and articulatory muscles (with probable excessive tension) that may be followed by anxiety [3]. This study concerned procedures that count voice symptoms in discrete syllables, so procedures such as time-interval analysis that are not syllable-based are not considered in detail. The voice samples will be classified based on pitch and tonal qualities with the practical help of a medical expert [4] or a clinician [5] who has worked in the speech language therapy. The identification parameters involved will be rate of speech or the number of words spoken over the time window, repetition of the same sound or the prolongation of the speech using zero crossing rate, blocks using the noise detector, and prolongation using the vowel deficiencies. Part of the work includes adapting our own algorithms to measure speech rate and pauses to neurogenetic speech. Embedded in this comprehensive test is a story-description yielding a language sample that will be audio recorded. Audio sample of 2 - 3 minutes captured at a minimum sampling rate of 16 KHz (16-bit). High quality audio equipment has been sought for in this regard. The assessment consists of two sets of pictorial depictions of line drawn figures expressing different emotions, five in each set, the first of which I s a positive-negative scale (valence) and the second an excited-calm scale (arousal). Subjective Measure of Depression. 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. Audio Data is directly imported into MATLAB after classification by researchers and we are going to use a set of different toolboxes and processed and exported to provide rapid feedback to the individual regarding his or her sample parameters through a speech diagnosis test using MATLAB tools. The first process is identification of the feature vectors of the characteristic audio sample. Various audio segmentation for identification of a suitable task in developing an improved algorithm for audio processing applications like speaker identification, speaker

tracking, and automatic speech recognition (ASR) has been done earlier. However, in the present scenario it is to be diagnosed whether speech deficiencies can be identified using MATLAB toolbox. As in this research we will designed our data analysis by comparing with pre recorded normal speech sample, so it is expected that our software able to develop speech analysis much quicker as real time analysis at the same time give more authentic diagnosis. One of the primary objectives of any analysis based on acoustic signal is to identify and remove noises before processing.

Stimuli

The stimuli were audio recordings of 20 sentences spoken by a (50% male and 50% female) native speaker. These recording were chosen from a set of sentences had a mean duration of 2137 ms (2002 - 2306 ms), consisted of 10 syllables (range: 8 - 11) and were matched for number of labially produced phonemes (which are more easily visually discriminated). Method used for signal classification is a challenging phase for researchers in most of the signal processing applications.

WaveSurfer is one of the solutions to carry out speech sample research. It is an open source platform (such as MySQL or OpenOffice) which allows for sound modeling (mainly speech analysis) as well as further processing and manipulation. WaveSurfer can be used as a stand-alone device or as an item of advanced sound processing system that can be accomplished by installing additional custom plugins or embedding their components in other applications.

The most recent version of the programme, i.e. WaveSurfer 1.8.8, can be downloaded free of charge from WaveSurfer home page: http://sourceforge.net/projects/wavesurfer WaveSurfer 1.8.8 features:

- Type: Freeware
- Size: 1.58 MB
- OS: w98 WNT w2000 wxp
- Software language: English.

One of the many features of the WaveForms system is the development of spectrograms by which a speech therapist or any other researcher concerned with speech analysis is presented with an objective record of specific sounds in articulation and can therefore analyze it and equate it with template patterns. The photo can prove valuable to the client as well as to the therapist. The wave and its spectrogram clearly show multiple repetitions of “p” sound as well as “po” and “do” syllables in non-fluent utterance, all characteristic of the anlaut sound or syllable stutter. In fact, the spectrogram shows the amount of energy present in repetitions by increased saturation. All pronouncements often vary in the overall utterance duration. Throughout traditional logopedics, computer-aided speech analysis plays an important role. The introduction of cutting-edge speech analysis techniques helps the therapist to provide a more accurate diagnosis for the client, and it also facilitates the process of speech therapy [4]. The direct comparison of the preceding figures does not necessarily reflect the time span of both statements (different time scale applied), contributing to the need for visualization. These classification methods are either application specific or not efficient for all types of signal. Signal classification techniques play an important role in stuttering speech recognition and analysis. Support vector machine (SVM) is a simple and primarily used to solve nonlinear problems including stuttering speech signal characterization and categorization [5]. Voiced and unvoiced segments of speech signal can be effectively classified using SVM. A basic phenomenon of SVM is to visualize the problem in to high dimensional space so as to classify the data accurately. Stuttering signal consist of voiced, unvoiced and redundant segments of speech. ANN will also be used to classify fluent and non-fluent portion of stuttering acoustic signal. Recognition of prolongation and repetition in stuttering speech are interesting applications of ANN.

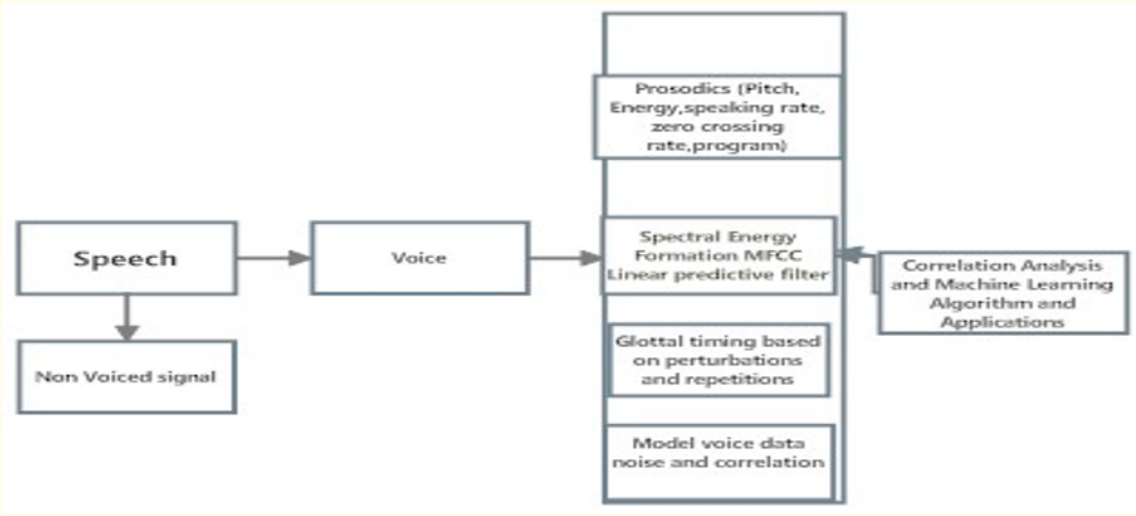

Figure 1: *Application of speech processing analysis prototype#1.*

Results and Discussion

We conducted a preliminary study with five participants to examine the relationships between the various parameters of this analysis system and their effects on the management of Aphasic patient. Stutterers ' speech acoustic analysis has developed techniques for identifying phonation breaks. The method of identification of these speech anomalies in stutterers includes: segmentation of speech signals, equalization of higher frequency portion energy levels and measurement of average signal energy (as in repetition according to the formula). The software then calculates the energy state's upper and lower value for that component. These values allow the preceding phonation break to be determined by the trailing edge. The edge of the trailing describes the rate of energy decline. Prototype #1 was used in this study and the parameters were controlled on the host PC. Based on the lessons learned from the study, in the following sections we present the parameters that should be examined in detail in future formal studies

Conclusion

So far no single method has proven to work for everyone, despite reports of "cures" through new treatments, drugs or devices. For these reasons, any Medical Foundation does not endorse any single approach to Aphasia treatment. Stutterers ' computer acoustic speech analysis enables notable and objective evaluation of changes in the structure of the speech signal, and increasingly accurate tests provide insight into mechanisms that accompany the stutter phenomenon.

Acknowledgements

We acknowledge the contribution of the West Bengal speech therapy organization for providing us data and also to interact with some patients with speech therapy based self help groups.

Conflict of Interest

No conflict in interest exists in our research model.

Bibliography

1. Van Borsel J and Eeckhout H. "The speech naturalness of people who stutter speaking under delayed auditory feedback as perceived by different groups of listeners". *Journal of Fluency Disorders* 33.3 (2008): 241-251.

2. Lincoln M., *et al*. "Altered auditory feedback and the treatment of stuttering: a review". *Journal of Fluency Disorders* 31.2 (2006): 71-89.

3. Tarkowski Z. "Jąkanie". Wydawnictwo Naukowe PWN, Warszawa (2002).

4. Bakat B., *et al*. "Does Voice Therapy Cure all Vocal Fold Nodules?" *International Journal of Phonosurgery and Laryngology* 4.2 (2014): 55-59.

5. Y Sumi., *et al*. "Analysis environment of conversational structure with nonverbal multimodal data". In Proceedings of ICMI-MLMI'10 (2010): 44.

Volume 8 Issue 11 November 2019