Deterioration of Speech as an Indicator of Physiological Degeneration (DESIPHER)

Bonnie J. Dorr[1], Ian Perera[1], Samuel Phillips[2], Jan Jasiewicz[2]

[1] Florida Institute for Human and Machine Cognition, Ocala, FL

[2] JAH Veterans' Hospital, HSR&D/RR&D Center of Innovation on Disability and Rehabilitation Research (CINDRR), Tampa, FL

Abstract

We approach the problem of understanding the disease progression and speech of patients suffering from ALS as a language divergence identification problem. We summarize the promises and challenges of using speech-related biomarkers to adapt speech recognition to ALS patients and correlate language divergence with disease progression.

Introduction

Most physiological assessments commonly used to determine the functional status of patients with Amyotrophic lateral sclerosis (ALS) require trained clinical personnel to administer and interpret the results. Speech impairments eventually affect 80-95% of patients with ALS (Beukelman, 2011). Initial impairments include reduced speaking rate, imprecise articulation, and changes in phonatory quality, and limit activities in which ALS patients are able to participate, leading to social isolation (Tomik, 2010).

We investigate the use of automatic speech recognition and machine learning software to: identify speech pathologies and use them to predict other aspects of physiological degeneration associated with ALS (e.g., respiratory difficulty or inability to swallow), improve speech recognition for those with speech impairments, and improve the quality of life for patients through conversation with a computer.

The DESIPHER project (DEterioration of Speech as an Indicator of PHysiological DegenERation) involves the analysis of speech-related symptoms through the processes used to analyze language accent and dialectal divergence. The greater the changes in speech, i.e. stronger divergence, the greater the progression of the disease. We believe this can ultimately provide improved speech recognition in this population, and serve as a marker of disease progression.

Background

Amyotrophic Lateral Sclerosis (ALS) is a rapidly progressive, neurodegenerative disease that involves upper and lower motor neurons and is the most common form of motor neuron disease. A serous healthcare crisis is emerging from the lack of capacity to support ALS sufferers, who typically must rely on a caregiver to manage their daily activities and to assist in relaying information between medical professionals and the patient.

Our speech research focuses on the detection of dialectal variations by identifying speech language *divergences* along a range of different dimensions. We borrow the notion of divergence from the study of cross-linguistic variations (Dorr, 1993) and apply it towards developing an assessment of bulbar function in patients with ALS, to improve upon existing assessments (Green et al., 2013).

A significant aspect of this work is the identification of possible speech-related biomarkers to be examined for correlations with other aspects of physiological function. Some that have been published in the literature (e.g., Duffy et al., 2012; Rudzicz at al., 2014) that are relevant to dialectal variations—but have not heretofore been measured objectively through dialectal analysis—include imprecise consonants, distorted vowels, slow rate, short phrases, continuous voicing, and distorted substitutions or additions.

The closest speech processing study to the divergence approach described above is that of (Biadsy et al., 2011), where the variation of speech properties under intoxicated and sober conditions were investigated. We will leverage an important discovery from this earlier work: since specific speech properties systematically vary in impaired speech, we can correlate the divergence from baseline English to different stages of ALS (i.e., the degree of biophysical degeneration) and therefore go beyond recognition of the presence of the disease.

Additional approaches that have examined recognition of impaired speech to answer yes/no questions include detection of Parkinson's (Little, 2012) and Alzheimer's (Rudzicz et al., 2014). Work on dialect and accent adaptation and other research applicable to the problem of recognizing impaired speech (Livescu and Glass, 2004) could

also be used to measure the degree of deterioration over time.

Approach

Borrowing from the field of machine translation (Dorr, 1993), we adopt a paradigm in which the notion of divergence is central. To illustrate the concept of divergence across languages, consider the properties in Table 1 shared between Spanish and four other languages:

Spanish	Italian	French	English	Chinese
vocabulary	X	X	X	
syntax	X	X		
pronunciation	X			

Table 1: Linguistic Divergence across Languages

In Table 2, articulatory properties for "Baseline" English (i.e. pre-symptomatic) are compared to those exhibited during different stages of ALS (early, mid, late), thus enabling a framework for detecting the degree of divergence applied to stages of ALS rather than different languages.

"Baseline"	ALS (early)	ALS (mid)	ALS (late)
consonants	imprecise	imprecise	imprecise
vowels	~baseline	distorted	distorted
bilabials (b,m)	~baseline	~baseline	hypernasal

Table 2: Example of Speech-related Divergence

Two hypotheses central to this work are the following:

H1: Automatic speech recognition can be achieved for ALS patients using speech-model adaptation.

H2: Divergence of speech from baseline will serve as a marker for disease progression, i.e., specific changes in physiological functioning such as difficulty swallowing or breathing are associated with different types of language/speech errors.

Experimental Design

The Florida Institute for Human and Machine Cognition and James A. Haley Veteran's Administration (Tampa VA) have partnered to investigate the hypotheses above in a longitudinal study of ALS patients. Veterans who come in at least quarterly as part of the Tampa ALS program will be followed over a period of two years, with passages recorded and transcribed for use in the study.

Annotation of the training data will be undertaken to identify speech characteristics in the transcribed text (e.g., hypernasality) and to use these for the development of speech adaptation methods.

We summarize the experiments described above in terms of the tasks associated with our two hypotheses:

H1: The speech recognition output will be compared to the annotated transcript, and recognition is achieved if the accuracy was greater than 80%.

H2: We will apply Pearson or Spearman to test the correlation of each functional measure with divergence. Measurements will be compared from baseline to each subsequent point. i.e. T1/T2, T1/T3, etc. Secondarily, a times series analysis will be performed.

Impact

At the Tampa VA, since 2007, there has been a consistent rise in the number of veterans diagnosed and treated with ALS. By understanding how speech impairment correlates with other types of physiological degeneration, it may be possible to apply a new, non-invasive measure for assessing the functionality of an ALS patient. Results of the study are expected to improve our ability to appropriately identify and intervene when veterans with ALS are at risk of serious adverse medical issues such as respiratory failure and aspiration, while speech understanding through adaptation could allow us to improve their quality of life through conversation with a computer. Future work will explore whether this understanding could be used as part of a natural language interface in a Smart Home setting to enable greater levels of personal autonomy and safety while allowing monitoring for changes in divergence that signal a need for intervention.

Acknowledgements

This material is based upon work supported by ONR grant N00014-12-1-0547 and by the Nuance Foundation.

References

Beukelman, D., Fager, S. and Nordness, A. 2011. Communication Support for People with ALS. *Neurology Research International*.

Biadsy, F, Wang, W. Y., Rosenberg, A. 2011. Intoxication Detection using Phonetic, Phonotactic and Prosodic Cues. *Proceedings of Interspeech*.

Dorr, Bonnie J. 1993. *Machine Translation*: A View from the Lexicon. MIT Press, Cambridge, MA.

Duffy, J. 2012. *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management*, 3rd edition.

Green, J. R., Yunusova, Y., Kuruvilla, M. S., Wang, J., Pattee, G. L., Synhorst, L., Zinman, L., Berry, J. D. 2013. Bulbar and speech motor assessment in ALS: Challenges and future directions. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*. 14: 494–500.

Little, M. 2012. A Test for Parkinson's with a Phone Call. *TED Talk*.

Livescu, K. and Glass, J. 2000. Lexical Modeling of Non-Native Speech for Automatic Speech Recognition. In *Proceedings of ICASSP*: 1683-1686.

Rudzicz, F., Chan-Currie, L., Danks, A., Mehta, T., Zhao, S. 2014. Automatically identifying trouble-indicating speech behaviors in Alzheimer's disease. In *Proceedings of ACM ASSETS*, Rochester NY.

Tomik, B. and Guiloff, R.J. 2010. Dysarthria in amyotrophic lateral sclerosis: A review. *Amyotroph Lateral Sclerosis,* 11(1-2): 4-15.