SONDE

Your Health Speaks

November 2019

Sonde is unlocking voice as a vital sign and predictor of health

The market lacks objective tools to cue appropriate actions as our health status changes

National Museum of American History & The Medical and Chirurgical Faculty of the State of Maryland

Ability to Understand and Respond to Just A Few Seconds of Speech Is Rapidly Changing Our World

Automatic Speech Recognition
(Focus since ~1952)

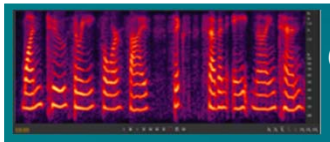
LINGUISTIC FEATURES

VOCAL BIOMARKERS

Voice assistants that automate routine tasks

Voice assistants detect health changes to mediate early and efficient care access

Automatic Health Recognition



OBJECTIVE ACOUSTIC CHANGES

+

CLINICAL GRADE HEALTH DATA

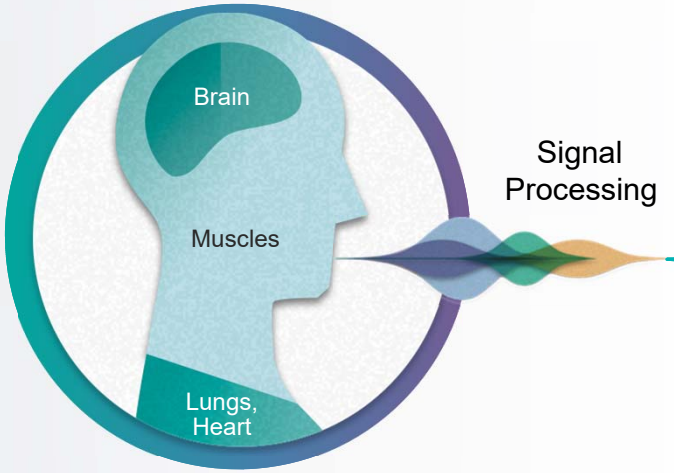
SONDE

Technology Foundation: Objective Characterization of Involuntary Acoustic Changes That Accompany Changes In Health

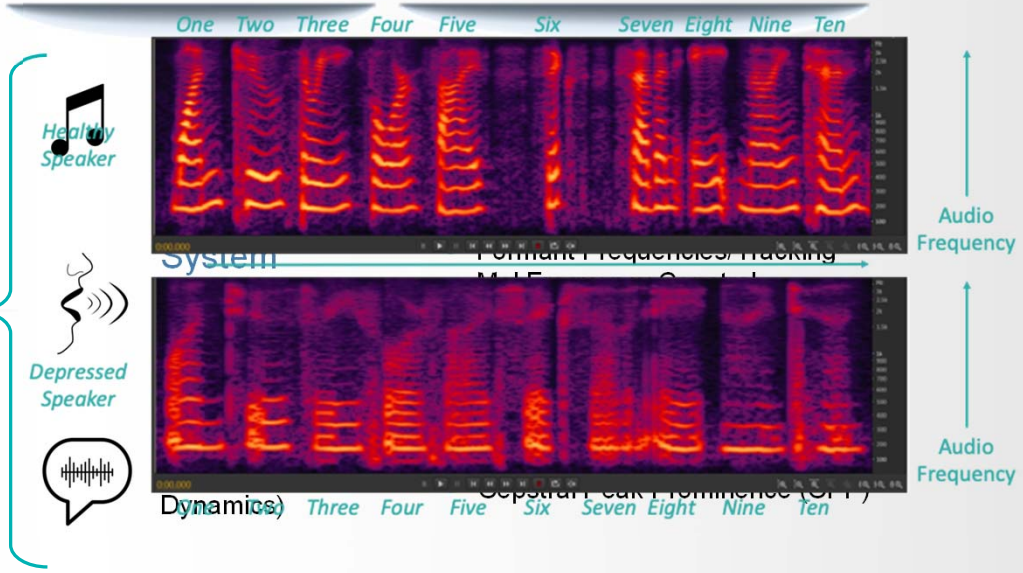
Changing Disease Physiology Alters Acoustic Features & Their Temporal Coordination

Major Acoustic Feature Categories

Individual Feature Examples (from ~thousands)

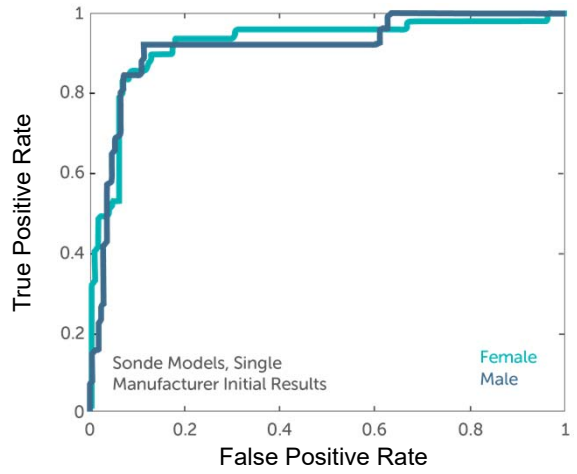


Signal Processing



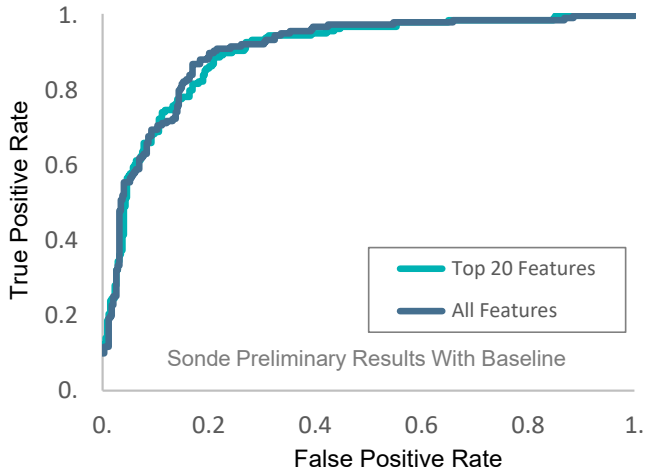
Sonde Has Real-World Proof of Feasibility For Measuring a Range of Health Conditions Using Short Smartphone Voice Samples

Depression Screening Example



Performance on par with PHQ-9 – standard clinical screening instrument (*smartphones, 6 seconds of speech, no baseline, >4K US individuals*)

Sonde Sleepiness Screening Example



With a few training sessions, feasible to accurately predict when KSS ≥ 7, (Best-performing cued sample based on typical voice assistant command)

Sonde Vocal Biomarker Performance Enables Powerful Augmentation of Depression Screening Capabilities

Assuming a population with 10% prevalence for PHQ-9 ≥ 10

	Sensitivity	Specificity	Accuracy	Depression ID Confusion Matrix	
PHQ-9 (Kroenke, et al., 2001)	88%	88%		≤ 9 ≥10	90 [*79* 11 10 1 *9*]
Sonde Vocal Biomarkers (Huang, et al., 2019, in press)	68.5%	71.9%	71.4%	≤ 9 ≥10	90 [*65* 25 10 3 *7*]

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TAFFC.2019.2944380

Investigation of Speech Landmark Patterns for Depression Detection

Zhaocheng Huang, *Member, IEEE,* Julien Epps, *Member, IEEE*, Dale Joachim, *Member, IEEE*

Identify 9 of 10 depressed individuals if *employed and answered honestly*

Identify 7 of 10 depressed individuals correctly *with lower friction and greater objectivity*

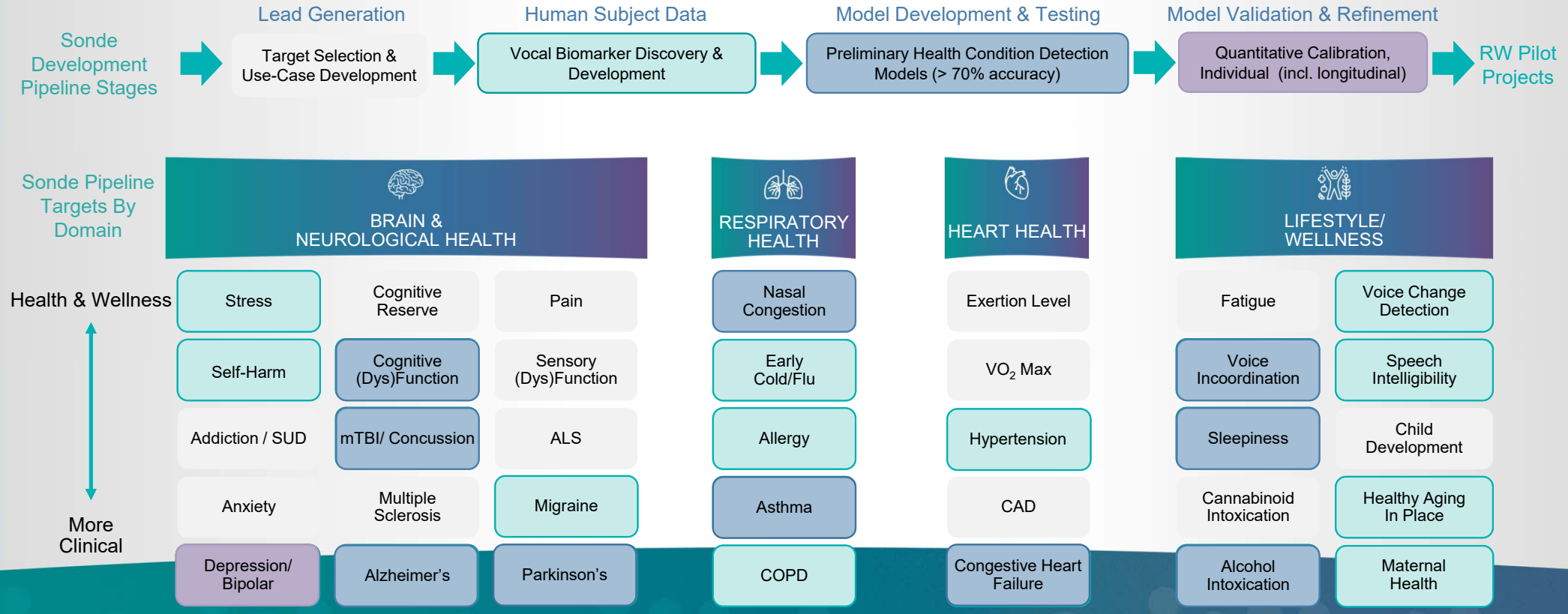
Vocal biomarker performance improving and can be combined with other assessments

Sonde's Depression Risk Level Estimation Based on PHQ-9 Scores Has High Accuracy

Low, Med, High strength corresponding to PHQ-9 ranges			
	Predicted High (17-27)	Predicted Moderate (10-16)	Predicted Low (0-9)
Actual High (17-27)	**9%**	3%	1%
Actual Moderate (10-16)	12%	**18%**	12%
Actual Low (0-9)	7%	8%	**29%**

Deployed initially on-phone testing with 791 individuals. Already generating good initial confusion matrix performance.
Next steps: 0-100 calibration, time-based consensus scoring vs. single file scoring

Potential of Vocal Biomarker Platform Reaches Far Beyond Depression and Sleepiness



Example: Potential for Sonde Vocal Biomarkers to Augment Measurement Capabilites For Monitoring Parkinson's Disease Symptoms & Progression

INTERSPEECH 2015

Segment-dependent dynamics in predicting Parkinson's disease

James R. Williamson[1], Thomas F. Quatieri[1], Brian S. Helfer[1], Joseph Perricone[1], Satrajit S. Ghosh[2], Gregory Ciccarelli[1], Daryush D. Mehta[1]

[1] MIT Lincoln Laboratory, Lexington, Massachusetts, USA
[2] Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

jrw@ll.mit.edu, quatieri@ll.mit.edu, brian.helfer@ll.mit.edu, joey.perricone@ll.mit.edu, satra@mit.edu, gregory.ciccarelli@ll.mit.edu, daryush.mehta@ll.mit.edu

Abstract

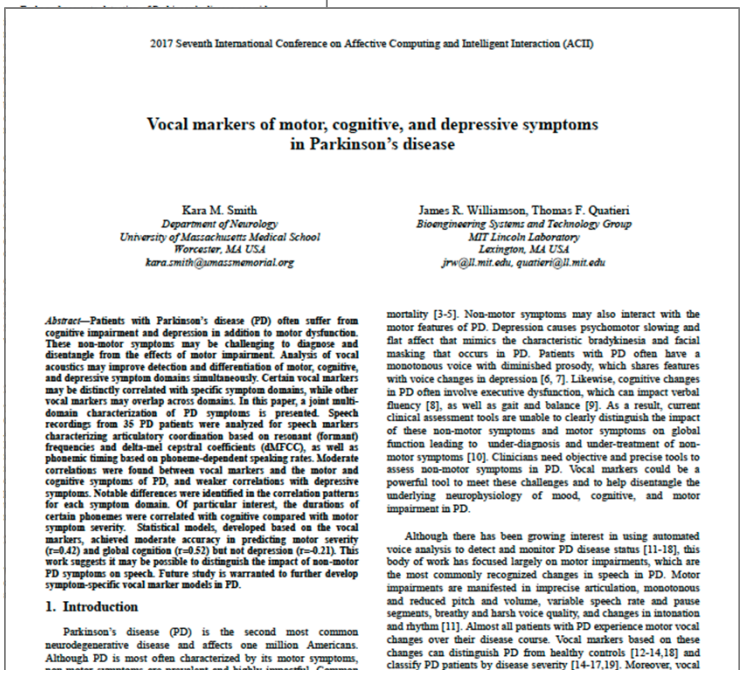
Early, accurate detection of Parkinson's disease may aid in possible intervention and rehabilitation. Thus, simple noninvasive biomarkers are desired for determining severity. In this study, a novel set of acoustic speech biomarkers are introduced and fused with conventional features for predicting clinical assessment of Parkinson's disease. We introduce acoustic biomarkers reflecting the segment dependence of changes in speech production components, motivated by disturbances in underlying neural motor, articulatory, and prosodic brain centers of speech. Such changes occur at phonetic and larger time scales, including multi-scale perturbations in formant frequency and pitch trajectories, in phoneme durations and their frequency of occurrence, and in temporal waveform structure. We also introduce articulatory features based on a neural computational model of speech production, the Directions into Velocities of Articulators (DIVA) model. The database used is from the Interspeech 2015 Computational Paralinguistic Challenge. By fusing conventional and novel speech features, we obtain Spearman correlations between predicted scores and clinical assessments of r = 0.63 on the training set (four-fold cross validation), r = 0.70 on a held-out development set, and r = 0.97 on a held-out test set.

Index Terms: Parkinson's disease, speech biomarkers, phoneme and pause duration, articulatory coordination, neural computational models of motor control

1. Introduction

Parkinson's disease is a neurological disorder with associated progressive decline in motor precision and sensorimotor integration stemming presumably from the basal ganglia. In this disorder, there is a steady loss of cells in the midbrain, leading to speech impairment in nearly 90% of subjects [1]. Speech and voice characteristics of Parkinson's disease include imprecise and incoordinated articulation, monotonous and reduced pitch and loudness, variable speech rate and rushes of breath and pause segments, breathy and harsh voice quality, and changes in intonation and rhythm [2][3][4][5][6].

This work is sponsored by the Assistant Secretary of Defense for Research & Engineering under Air Force contract #FA8721-05-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government. SG was partly supported by a joint collaboration between the McGovern Institute for Brain Research and MIT LL.



2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)

Vocal markers of motor, cognitive, and depressive symptoms in Parkinson's disease

Kara M. Smith
Department of Neurology
University of Massachusetts Medical School
Worcester, MA USA
kara.smith@umassmemorial.org

James R. Williamson, Thomas F. Quatieri
Bioengineering Systems and Technology Group
MIT Lincoln Laboratory
Lexington, MA USA
jrw@ll.mit.edu, quatieri@ll.mit.edu

***Abstract*—Patients with Parkinson's disease (PD) often suffer from cognitive impairment and depression in addition to motor dysfunction. These non-motor symptoms may be challenging to diagnose and disentangle from the effects of motor impairment. Analysis of vocal acoustics may improve detection and differentiation of motor, cognitive, and depressive symptom domains simultaneously. Certain vocal markers may be distinctly correlated with specific symptom domains, while other vocal markers may overlap across domains. In this paper, a joint multi-domain characterization of PD symptoms is presented. Speech recordings from 35 PD patients were analyzed for speech markers characterizing articulatory coordination based on resonant (formant) frequencies and delta-mel cepstral coefficients (ΔMFCC), as well as phonemic timing based on phoneme-dependent speaking rates. Moderate correlations were found between vocal markers and the motor and cognitive symptoms of PD, and weaker correlations with depressive symptoms. Notable differences were identified in the correlation patterns for each symptom domain. Of particular interest, the durations of certain phonemes were correlated with cognitive compared with motor symptom severity. Statistical models, developed based on the vocal markers, achieved moderate accuracy in predicting motor severity (r=0.42) and global cognition (r=0.52) but not depression (r=-0.21). This work suggests it may be possible to distinguish the impact of non-motor PD symptoms on speech. Future study is warranted to further develop symptom-specific vocal marker models in PD.**

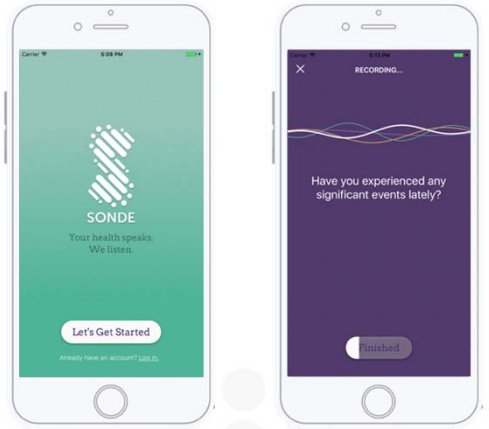
1. Introduction

Parkinson's disease (PD) is the second most common neurodegenerative disease and affects one million Americans. Although PD is most often characterized by its motor symptoms, mortality [3-5]. Non-motor symptoms may also interact with the motor features of PD. Depression causes psychomotor slowing and flat affect that mimics the characteristic bradykinesia and facial masking that occurs in PD. Patients with PD often have a monotonous voice with diminished prosody, which shares features with voice changes in depression [6, 7]. Likewise, cognitive changes in PD often involve executive dysfunction, which can impact verbal fluency [8], as well as gait and balance [9]. As a result, current clinical assessment tools are unable to clearly distinguish the impact of these non-motor symptoms and motor symptoms on global function leading to under-diagnosis and under-treatment of non-motor symptoms [10]. Clinicians need objective and precise tools to assess non-motor symptoms in PD. Vocal markers could be a powerful tool to meet these challenges and to help disentangle the underlying neurophysiology of mood, cognitive, and motor impairment in PD.

Although there has been growing interest in using automated voice analysis to detect and monitor PD disease status [11-18], this body of work has focused largely on motor impairments, which are the most commonly recognized changes in speech in PD. Motor impairments are manifested in imprecise articulation, monotonous and reduced pitch and volume, variable speech rate and pause segments, breathy and harsh voice quality, and changes in intonation and rhythm [11]. Almost all patients with PD experience motor vocal changes over their disease course. Vocal markers based on these changes can distinguish PD from healthy controls [12-14,18] and classify PD patients by disease severity [14-17,19]. Moreover, vocal

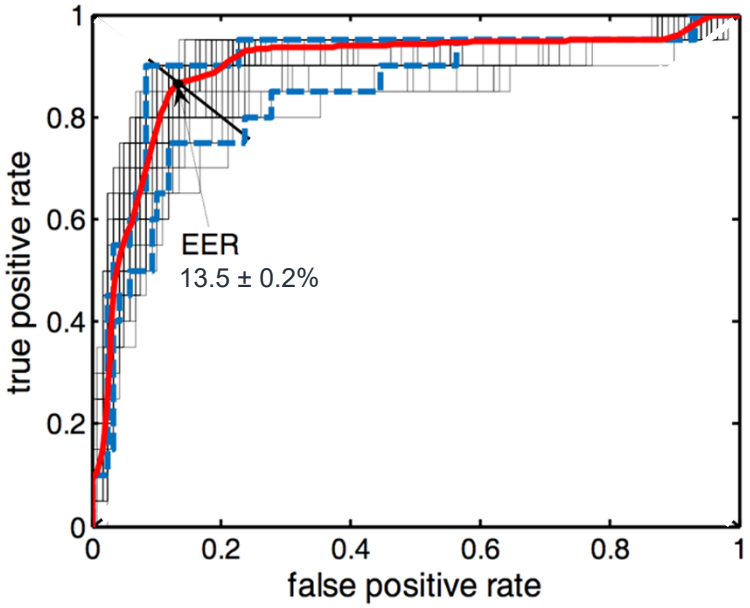
Proofs-of-Concept: show strong correlation factors for UPDRS score prediction with further potential to distinguish impact of non-motor PD symptoms on speech

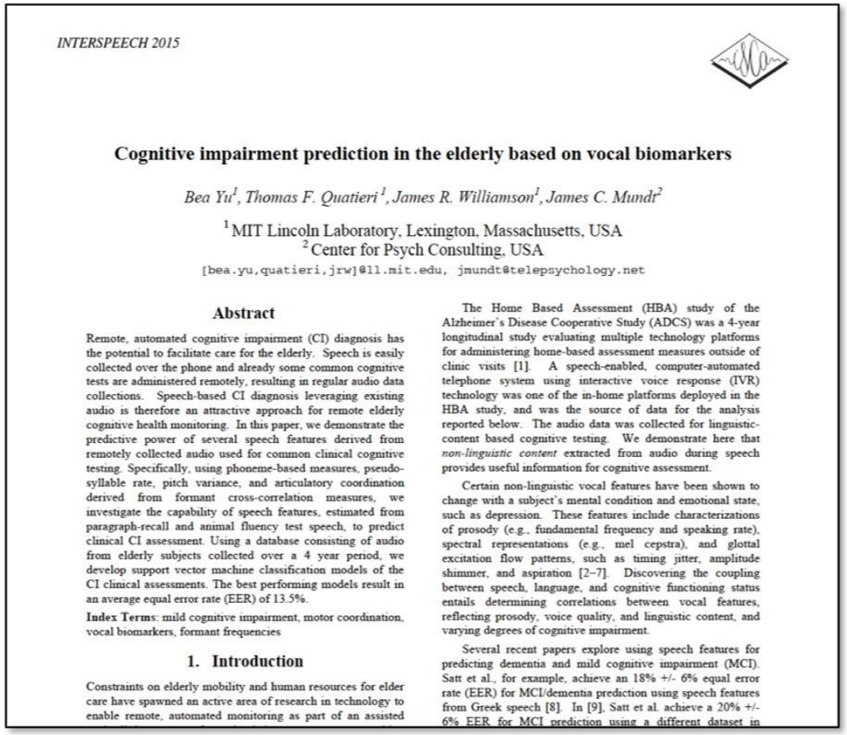
Opportunity for rapid scaling, validation, and commercialization using Sonde's platform



Example – Potential for Sonde Vocal Biomarkers to Augment Screening for Mild Cognitive Impairment & Early AD Dementia

Voice-Only ML Prediction Of Cognitive Impairment vs. Clinical Screening (No Healthy Baseline Needed)





INTERSPEECH 2015

Cognitive impairment prediction in the elderly based on vocal biomarkers

Bea Yu[1], Thomas F. Quatieri[1], James R. Williamson[1], James C. Mundt[2]

[1] MIT Lincoln Laboratory, Lexington, Massachusetts, USA
[2] Center for Psych Consulting, USA

[bea.yu,quatieri,jrw]@ll.mit.edu, jmundt@telepsychology.net

Abstract

Remote, automated cognitive impairment (CI) diagnosis has the potential to facilitate care for the elderly. Speech is easily collected over the phone and already some common cognitive tests are administered remotely, resulting in regular audio data collections. Speech-based CI diagnosis leveraging existing audio is therefore an attractive approach for remote elderly cognitive health monitoring. In this paper, we demonstrate the predictive power of several speech features derived from remotely collected audio used for common clinical cognitive testing. Specifically, using phoneme-based measures, pseudo-syllable rate, pitch variance, and articulatory coordination derived from formant cross-correlation measures, we investigate the capability of speech features, estimated from paragraph-recall and animal fluency test speech, to predict clinical CI assessment. Using a database consisting of audio from elderly subjects collected over a 4 year period, we develop support vector machine classification models of the CI clinical assessments. The best performing models result in an average equal error rate (EER) of 13.5%.

Index Terms: mild cognitive impairment, motor coordination, vocal biomarkers, formant frequencies

1. Introduction

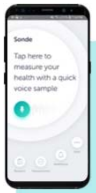
Constraints on elderly mobility and human resources for elder care have spawned an active area of research in technology to enable remote, automated monitoring as part of an assisted

The Home Based Assessment (HBA) study of the Alzheimer's Disease Cooperative Study (ADCS) was a 4-year longitudinal study evaluating multiple technology platforms for administering home-based assessment measures outside of clinic visits [1]. A speech-enabled, computer-automated telephone system using interactive voice response (IVR) technology was one of the in-home platforms deployed in the HBA study, and was the source of data for the analysis reported below. The audio data was collected for linguistic-content based cognitive testing. We demonstrate here that *non-linguistic content* extracted from audio during speech provides useful information for cognitive assessment.

Certain non-linguistic vocal features have been shown to change with a subject's mental condition and emotional state, such as depression. These features include characterizations of prosody (e.g., fundamental frequency and speaking rate), spectral representations (e.g., mel cepstra), and glottal excitation flow patterns, such as timing jitter, amplitude shimmer, and aspiration [2–7]. Discovering the coupling between speech, language, and cognitive functioning status entails determining correlations between vocal features, reflecting prosody, voice quality, and linguistic content, and varying degrees of cognitive impairment.

Several recent papers explore using speech features for predicting dementia and mild cognitive impairment (MCI). Satt et al., for example, achieve an 18% +/- 6% equal error rate (EER) for MCI/dementia prediction using speech features from Greek speech [8]. In [9], Satt et al. achieve a 20% +/- 6% EER for MCI prediction using a different dataset in

Careful Technology Progression to Manage Security, Privacy, and ELSI



Cued Mode

- Direct Sonde access to device audio
- Familiar voice assistant skill UX
- Tailored elicitations for enhanced performance

When a user repeats words requested by the appwords for six seconds, for example, “Pataka, pataka, pataka…”

Screening, personal exploration, novelty & social, etc

Command Mode

- Opt-in
- All voice commands available for analysis
- Special controls, e.g. pause/mute timer, etc.

When the user mentions common commands like, “Hi Bixby, What’s the weather like today?”

Tracking, monitoring, health cues & alerts, patient management

Always Sensing Mode

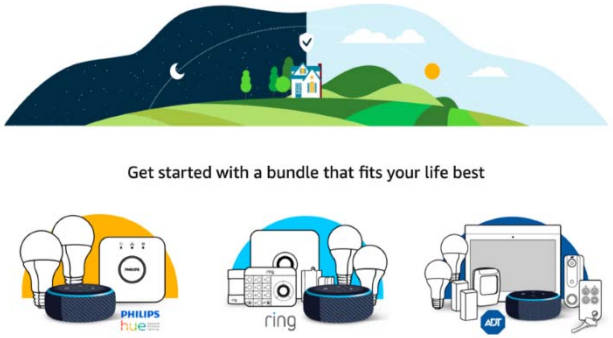
- Opt-in or dedicated/premium device
- Embedded-mode-only option (e.g. all results local)
- New “wake-to-health” UX opportunities

The app is listening for situations that require real-time tracking

Passive, persistent, and proactive health experiences (and health data volume/velocity)

As Voice Assistants Begin Moving Beyond Speech Recognition, Sonde Health is Building a Unique Platform For Addressing Broad Home Health Needs

Amazon's initial foray beyond speech recognition and into acoustic event detection is driving synergistic voice product integration opportunities for home security systems (~$34 billion global market)



Sonde's opportunity: Create voice enabled automatic alerting systems for important health conditions (global health spend $7.7 trillion+)

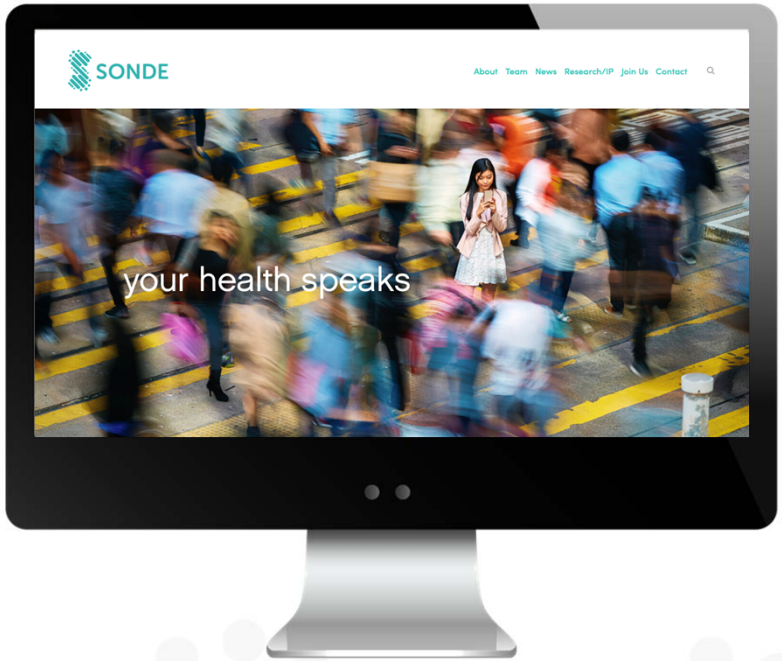
Sonde Acoustic Health Event Detection Can Transform the Way Mental and Physical Health is Monitored and Managed

Health Guard – Alerts to check in on family members/patients based on out-of-range voice health measures

Depression/anxiety/suicide risk detection and management

Real-world evidence

Thank You



sondehealth.com

PURETECH HEALTH

M. VENTURES

neoteny

MP Healthcare Venture Management, Inc.

CANEPA HEALTHCARE

The New York Times

The Boston Globe

The Telegraph

THE WALL STREET JOURNAL.

THE SUNDAY TIMES

STAT

SCIENTIFIC AMERICAN

FierceMedical Devices

Xconomy

+MASS DEVICE

MIT Technology Review

BOSTON BUSINESS JOURNAL

MedCity News

BostInno

MEDTECH BOSTON

90.9 wbur BOSTON'S NPR® NEWS STATION

Boston MAGAZINE

CRAIN'S BOSTON

mobihealthnews

AARP Real Possibilities

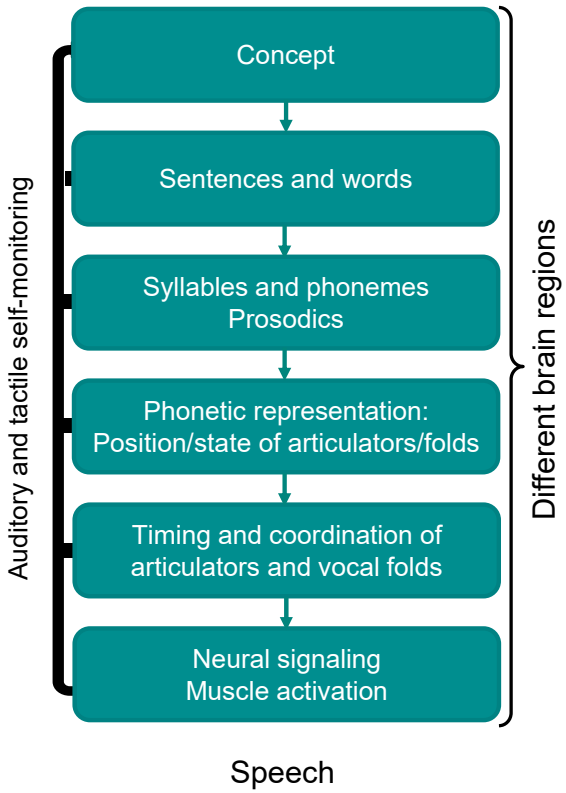
Futurism

QUARTZ

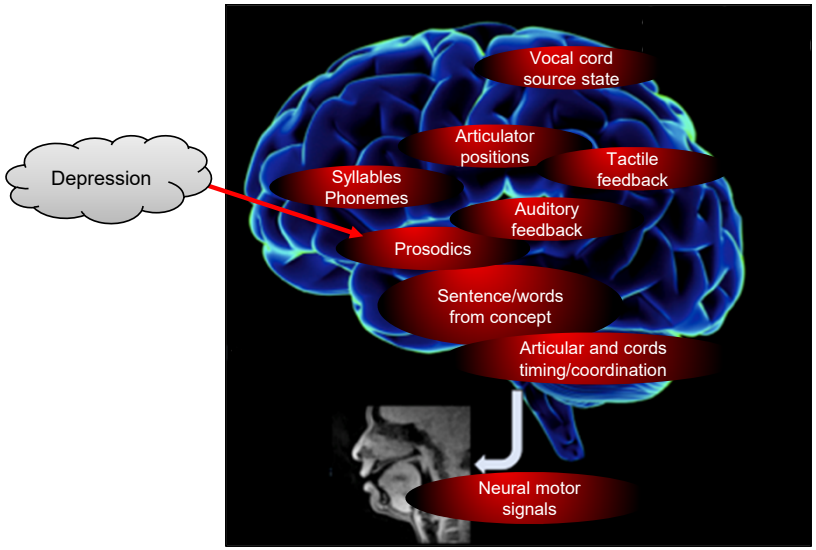
PARKINSON'S NEWS TODAY

Appendix: Sonde's Technology Enables Objective Characterization of Involuntary Acoustic Changes That Accompany Changes in Health

Approximate flow of complex speech production process



Broad Distribution of Core Speech Network



Disease mechanisms produce distinctive patterns of physiologic and neurologic disruption, produce speech modulations that may be similarly distinctive