

**BBI-ISR Joint Seminar** 



## Speech as a behavioral signal for mental and cognitive health: Some observations and results



## Vikram Mitra Senior Research Scientist Apple, Inc.

Speech is a key signal for human-to-human communication and for human-to-machine communication. It is used to convey our happiness-anger, satisfaction-frustration, excitement-despair and many other emotions. Speech also acts as a behavioral signal for humans to share their state-of-mind. Recognition of speech, speaker and language has been the center stage for speech technology research and development, and with recent advances, the state-of-the-art performance has progressed to admirable heights. Existing speech technologies are capable of deciphering what has been said with a high accuracy

but often fails to decipher how it has been said, limiting the capabilities to detect human behaviors such as satire, frustration, negative thoughts, forgetfulness, etc. Human listeners are capable of comprehending both the meaning and the intent of speech with exceptional accuracies; however, our existing speech processing systems are limited only to the meaning, often failing to comprehend intent. Recent studies have investigated the use of speech to detect a speaker's emotion, mental and cognitive health. This talk will explore speech as a behavioral signal to uncover the speaker's state of mind, particularly their mental health. Psychomotor retardation is a behavioral manifestation of mental health condition, which is an outcome of reduced articulatory motions. This talk will investigate if speech articulatory patterns can be used to uncover the speaker's psychomotor retardation and hence be used as a behavioral signal to detect mental health conditions. In addition, the talk will present how robust speech representations and articulatory information can be used to improve acoustic modeling tasks for speakers with certain clinical conditions.

Vikram Mitra is a Senior Research Scientist who is leading the speech based health efforts at Apple, working on speech processing and machine learning. He worked as an Advanced Research Scientist at SRI International's Speech Technology and Research Laboratory from 2011 to 2017. He is a Research Affiliate of Prof. Carol Espy-Wilson's Speech Communications Lab. He received his PhD in Electrical Engineering from the University of Maryland, College Park in 2010, where he worked with Prof. Espy-Wilson in her Speech Communication Lab on his Ph.D. dissertation titled "Articulatory information for robust speec recognition." His research interests include speech for health applications, robust signal processing for noise/channel/reverberation, speech recognition, production/perception-motivated signal processing, information retrieval, machine learning, and speech analytics. One of his major research contributions is the estimation of speech articulatory information from the acoustic signal, and using such information for recognition of both natural and clinical speech, and the detection of depressive symptoms in adults. He led SRI's STAR lab's efforts on robust acoustic feature research and development, which led to state-of-the-art results in keyword spotting and speech activity detection in DARPA's Robust Automatic Transcription of Speech Program. He has served as the PI/ co-PI of several projects funded by NSF and has worked on research efforts funded by DARPA, IARPA, AFRL, NSF and Sandia National Laboratories. He is a senior member of the IEEE and an affiliate member of the Speech and Language processing technical committee (SLTC), and he has served on the scientific committees of SPASR2013 and MLSLP2012. **Host**: Carol Espy-Wilson (espy@umd.edu)

## 1146 A. V. Williams Building | February 25 | 2 p.m.