A Platform for Automated Acoustic Analysis for Assistive Technology

Harriet Fell  
Northeastern University

Lorin Wilde  
S.T.A.R. Corp.

Suzanne Boyce  
University of Cincinnati

Keshi Dai  
Northeastern University

Joel MacAuslan  
S.T.A.R. Corp.

Objective
While physical, neurological, oral/motor, and cognitive impairments can all significantly impact speech, people with disabilities may still be best able to communicate with computers through vocalization.

Aspects of vocal articulation are highly sensitive markers for many neurological conditions. As a source of data, recordings are
- non-invasive,
- inexpensive to collect, and
- easily integrated into existing research and clinical protocols.

Both users with disabilities and neuroscience researchers could therefore benefit if there were convenient, automated tools for the acoustic analysis of speech and other vocal productions.

Statement of the Problem
To date, the use of vocal production data has been limited by a steep learning curve and the need for laborious hand measurement.

Our Solution
We are building a tool set for measures designed by clinicians to screen, diagnose or provide training to patients. This objective will be achieved by extending existing shareware software platforms with “plug-ins” that produce specific measures and report results to the user.

Past Work
We have built several successful therapeutic software devices and diagnostic tools that exploit non-verbal information in utterances
- To create assistive devices for individuals with disabilities
  - Treatment monitoring of Deep Brain Stimulators (DBS) for patients with Parkinson’s Disease (PD)
  - Evaluating progress in response to speech therapies
  - Encouraging young children to produce multi-syllable prospeech utterances using a teaching toy, i.e., visiBabble
- To analyze progress or decline for clinical or research studies
  - Acoustic analysis of cough to help diagnose disease
  - Determining degree of fatigue from clarity of utterances
  - Automatically measuring syllable complexity as a marker for clinically-relevant changes in vocalizations

Approach
Our Wavesurfer landmark plug-in includes user controls to automate:
- grouping of landmarks into syllable-like clusters
- detection of voicing and voice-onset time (VOT)
- removal of stray sounds, i.e., non-speech
- rapid computation of the vowel space.

and can export results about a single recording or a directory of recordings.

Evaluation
- Structured interviews and web surveys with speech scientists, neuroscientists, and clinicians.
- Ten graduate and postdoc speech scientists will use and evaluate the Wavesurfer tool on a corpus of recording of children with diagnosed autism.
- Pre- and post-test questionnaires

Future Work
We have started developing a plug-in for R to automate statistical analysis and reporting of landmark information.

Background
Speech landmarks (Ken Stevens et al. 1992, Liu 1995) are particularly useful when vocalizations are ill-formed.
- to compute clinically relevant measures for changes in speech production
- to provide real-time processing of vocalizations that are unsuitable for speech recognition.

Evaluation
This graph shows laryngeal vibration, or voicing and five frequency bands’ energy envelopes. Red ovals highlight the following:
(a) only two bands show abrupt change — real landmark is detected
(b) a well-formed pattern of abrupt changes simultaneous across all frequency bands just before laryngeal vibration onset — typical for a word-initial syllable such as “it”
(c) abrupt change simultaneous across all frequency bands during laryngeal vibration — typical of the release of a sonorant such as “it”

This work was sponsored in part by NIH grants R43-DC010104, R44-DC005534, R42-NS047959, R44-DC002925.
