Master thesis

April, 2025

To students at the Technical Faculty
(Medical Engineering / Medizintechnik, Computer Science/ Informatik / Artificial Intelligence)

Title: Machine learning based software tool for identification of key vocal fold dynamics in Adductor Spasmodic Dysphonia during connected speech from high-speed videoendoscopy and acoustic recordings

Background: Determination of signature biomarkers of vocal fold dynamics from connected speech that differentiate adductor spasmodic dysphonia (also known as adductor laryngeal dystonia) from vocally healthy subjects.
Given Data: A set of simultaneous high-speed videoendoscopy and audio recordings collected from 5 adults who have adductor spasmodic dysphonia (adductor laryngeal dystonia - ALD) and 5 vocally healthy without voice disorder will be provided. Each speaker will produce three signal types: one sustained vowel 'e', and 1 sentence (3 repetitions) in the English language that are known to elicit key voice breaks in patients with adductor spasmodic dysphonia, see figures on page 2. Results of auditory-perceptual ratings (i.e. gold standard) of voice quality from expert listeners using the Consensus Auditory-Perceptual Rating Scale will be provided; i.e. amount of oscillatory/voice breaks, micromotions (visible right before voice breaks in the GAW), motion irregularities will be labeled.
Research questions: **(1)** Do novel automated / semi-automated algorithms (implemented in a software) produce valid and reliable results compared with auditory perceptual rating in a clinical setting? **(2)** Does machine learning methods using (1) only sentences versus a (2) combination of sentences in combination with vowel provide better accuracy for identification of adductor spasmodic dysphonia; i.e. what data is best for judgement?
Goal: Develop a suitable software tool (Python) that detects and counts the four parameters of irregularities (micro motions, oscillatory breaks, motion irregularities, tremor rate in Hz) in ALD patients.
Methods & tasks: (1) Preprocessing of bimodal data: image processing of HSV and edit length of acoustic and HSV intervals. (2) Augment existing data (e.g. noise, frequencies, etc) to generate a larger data set for next step (3) Develop and investigate DL based models to automatically identify the looked after four parameters (micro motions, oscillatory breaks, motion irregularities, tremor) – see page 2. (4) implement algorithms in a usable application.

The work will be supervised by **Prof. Dr.-Ing. Michael Döllinger** (Member of Department Informatik & AIBE)**.** The thesis is in **cooperation with Prof. Rita Patel** (Dept. of Speech Language Pathology and Audiology, College of Health Sciences, Indiana University, USA)

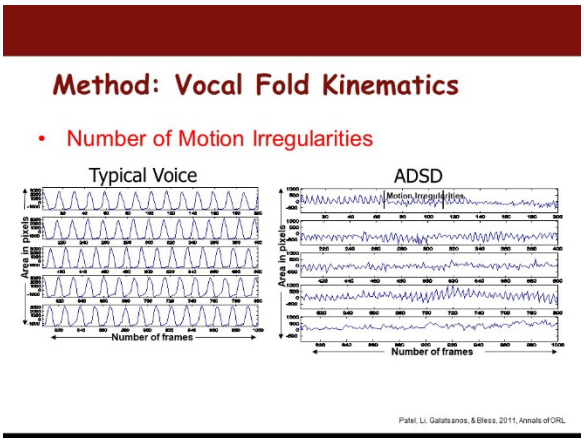
We search for a dedicated and motivated student with

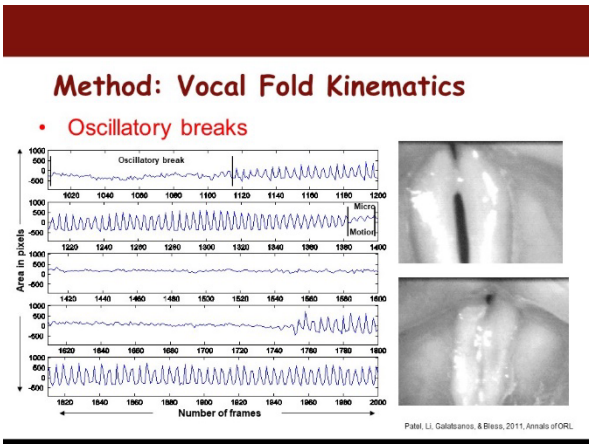
- experience in machine learning methods
- experience in Python

- good English skills (cooperation partner Prof. Patel is in the USA)

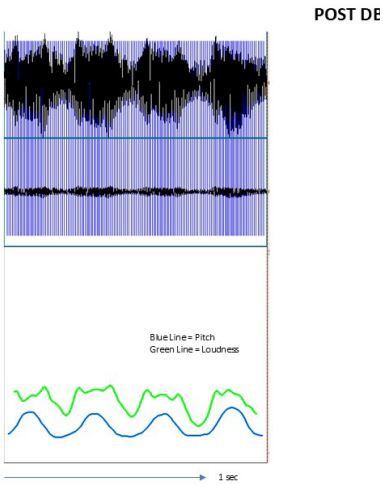
Contact persons:

Prof. Dr.-Ing. Michael Döllinger (michael.doellinger@uk-erlangen.de / Tel. 09131-85 33814)





Sustained Vowel /a/



Tremor