THE UNIVERSITY OF SYDNEY

Friedrich-Alexander-Universität Erlangen-Nürnberg

Master thesis

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

Title: Machine learning based software tool for word and vowel identification in acoustic recordings

Background: Determination of voice quality and its characteristics is essential for assessment and management of speech and voice patients, and especially for male to female transgender clients.
Given Data: A set of audio recordings collected from male, female and transgender adult speakers of Australian English with and without voice disorder will be provided. Each speaker produces 19 sentences containing 3-5 target vowels embedded in monosyllabic words. To provide ground-truth measurements, vowels are identified, and formants are estimated manually by experts. To automate vowel identification and formant estimation, a word- and vowel specific guideline for acoustic analysis will be provided.
Research questions: **(1)** Does a novel automated / semi-automated software produce valid and reliable results compared with manual analysis, in calculating a range of key formants across specifically identified vowels in a set of carrier phrases used in a clinical setting? **(2)** Which factors influence validity and reliability of formant measurement using validity and reliability of formant measurement using a novel software analysis tool compared with manual analysis?
Methods & tasks: Develop algorithms for vowel and word identification and implement in software with user friendly GUI. Incorporate following features in software:

1. When given a set of standard sentences, automatically identify individual words and vowels.
2. Then select specific vowels (again a standard set) and automatically identify onset and offset of the vowel - this step would include a manual over-ride in the case where individual variation means the selection is wrong.
3. Then depending on the individual vowel, automatically select the mid 30%/ end 30% etc (as per recommendations for specific vowels) - again this would have a manual over-ride in the case of speaker variation - and produce F0 and up to 5 formant values for this section.
4. Create an F1-F2 vowel plot from the automatically identified vowel formants and overlay it with selected reference plots.

Goal: Develop a suitable vowel and word identification algorithms and implement in a Python based software tool.

The work will be supervised by **Prof. Dr.-Ing. Michael Döllinger** (Member of Department Informatik & AIBE). The thesis is in **cooperation with Prof. Cate Madill** (Discipline of Speech Pathology, School of Health Sciences, Faculty of Medicine and Health, University of Sydney)

We search for a dedicated and motivated student with

- experience in machine learning methods

- experience in Python
- experience in software development

Contact persons:

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

Prof. Cate Madill, PhD (cate.madill@sydney.edu.au) Tel +61 418 278 238.