



MASTER THESIS

To students within
Medical Engineering / Medizintechnik,
Computer Science/ Informatik / Artificial Intelligence,
Data Science

**Hals-Nasen-Ohren-Klinik
Kopf- und Halschirurgie**
Direktorin: Prof. Dr. med. Sarina Müller, MHBA
**Abteilung für
Phoniatrie und Pädaudiologie**
Prof. Dr.-Ing. Michael Döllinger

Telefon: 09131 85-33814
Fax: 09131 85-32687
Michael.doellinger@uk-erlangen.de
Waldstr. 1, 91054 Erlangen
CICERO-Gebäude: Raumerstr. 1a

19.03.2026

Title: Machine Learning-Based Signal Typing for Adaptive Voice Onset Analysis

Background: Voice onset refers to the initiation of voice production and is broadly classified into three types: **modal, breathy, and pressed onset**. Clinically, identifying onset type is an important component of voice quality assessment and targeted voice therapy, particularly in the treatment of disorders such as muscle tension dysphonia. An objective acoustic measure that captures onset behaviour is **Vocal Rise Time (VRT)**, which quantifies the time from phonation onset to the point at which a stable amplitude is reached during sustained phonation [1]. Reliable extraction of VRT depends on the **analysis window size**, as different onset types produce distinct rise-time characteristics. Using a fixed analysis window can therefore introduce measurement errors. Automated signal typing as a preprocessing step for adaptive VRT analysis has been only minimally explored.

Therefore, the **overall goal** of this work is to automatically identify voice onset type prior to VRT analysis, enabling adaptive parameter selection and improving the accuracy of VRT estimation.

Current state: The Voice Onset Analysis Tool (VOAT) is a Python-based software tool that enables automated VRT extraction from sustained vowel recordings [2]. VOAT supports both manual and automated batch processing of voice datasets. Preliminary testing across multiple datasets shows that the VRT output is sensitive to the selected analysis window size, and its optimal value depends on the underlying phonation type. To address this limitation, a signal typing dataset containing recordings representing breathy, modal, and pressed phonation has been made available. Additional annotated datasets with manually measured VRT values are available for validation of the classification and subsequent VRT analysis.

Research questions: (1) Can acoustic recordings be automatically classified into breathy, modal, and pressed phonation types using machine learning techniques? (2) Which acoustic features differentiate best between the three voice onset types? (3) Can incorporating automated signal typing reduce VRT estimation error compared to a fixed-window analysis?

Methods & tasks:

1. Review the literature on laryngeal biomechanics, acoustics, and phonation onset analysis;
2. Train different machine learning models using acoustic features (e.g. spectral, noise-related, harmonicity) for multi-class signal typing classification;

3. Evaluate model performance using standard classification metrics and analyze the relationship between signal typing and VRT measurements.
4. Implement the algorithms in a Python framework and integrate them as a preprocessing module in the VOAT application pipeline.

Goal: Develop an ML-based automated signal typing module and integrate it into the VOAT pipeline to improve the accuracy and robustness of VRT measurement across different voice onset types.

The work will be jointly supervised by **Dr.-Ing. Tomas Arias Vergara** (Chair of Computer Science – Pattern Recognition Lab at FAU) and **Prof. Dr.-Ing. Michael Döllinger** (Member of Department Informatics & AIBE at FAU). The thesis is in close co-operation with Prof. Cate Madill at the University of Sydney, Australia.

We search for a dedicated and motivated student with

- Good programming skills (Python) and knowledge of GUI framework (PyQt)
- Experience in developing machine learning models
- Strong mathematical and signal processing background
- Good written and verbal communication skills in English

Contact person:

Dr.-Ing. Tomas Arias Vergara (tomas.arias@fau.de)

Prof. Dr.-Ing. Michael Döllinger (michael.doellinger@uk-erlangen.de)

References:

[1] Bi MS, Nguyen DD, Arias-Vergara T, Döllinger M, Holik J, Madill CJ. Effects of Instructed Laryngeal Manipulation on Vocal Rise Time. *Journal of Voice*. 2024 Nov 12.

[2] Arias-Vergara T, Madill C, Nguyen D, Holik J, Döllinger M. VOAT: Voice Onset Analysis Tool. *SoftwareX*. 2024 Sep 1;27:101802.