



**INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
SHORT ABSTRACT OF THESIS**

Name of the Student : Moakala Tzudir
Roll Number : 156102018
Programme of Study : Ph.D.
Thesis Title: Automatic Dialect Identification in Ao, a Low Resource Language
Name of Thesis Supervisor(s) : Prof. S.R.M. Prasanna (IIT Dharwad)
Prof. Priyankoo Sarmah (IIT Guwahati)
Thesis Submitted to the Department/ Center : EEE
Date of completion of Thesis Viva-Voce Exam : 25-01-23
Key words for description of Thesis Work : Ao, Tonal language, Low-resource language, Dialect identification, Tonal features, Excitation source features, Prosodic features

SHORT ABSTRACT

Dialect Identification (DID) is a significant research problem widely explored in major languages like Arabic, Chinese, and Spanish. DID can serve as a frontend for many applications like Automatic Speech Recognition (ASR) that may require special dialect-specific enhancements for improved performance. This thesis proposes an automatic DID system for Ao, an under-resourced language of India. Ao is a Tibeto-Burman language spoken in Nagaland. It is a tonal language with three lexical tones: high, mid, and low. Chungli, Mongsen, and Changki are the three dialects of Ao that differ in their respective tone assignment on lexical words. Four principal contributions are made in this thesis. The first contribution of this thesis is creating a manually collected and annotated novel speech dataset to foster research on the Ao language. The second contribution of the thesis is a detailed acoustic study of the unexplored tone dynamics of the dialects of Ao. Based on the analysis, a tonal feature ($\$F_0\$$) to capture the dialect-specific tone information is proposed. The DID performance improves when the proposed tonal feature is combined with other spectral features. As the third contribution, this thesis explores three excitation source features in the DID task. The source features studied are Residual Mel Frequency Cepstral Coefficient (RMFCC), Integrated Linear Prediction Residual Log Mel Spectrogram (ILPR-LMS), and Linear Prediction (LP)-gammatonegram. A notable performance improvement is observed when the source information is combined with the vocal tract information. The fourth contribution of this thesis is the exploration of prosody-related characteristics of speech signals. The prosodic features are observed to provide significant performance improvements in classifying the dialects of Ao. The thesis work is concluded by combining all the proposed approaches to build an efficient DID system for Ao. Among many hurdles in studying under-resourced languages like Ao, the need for more data is the most prominent. Nevertheless, the contributions of this thesis may bridge some of those gaps and spur future research in this direction.