

## VOICE AND SPEECH ANALYSIS FOR TECHNOLOGY

### STATEMENT OF RESEARCH PROJECT

Commercial off-the-shelf speech recognition systems have not yet been proven suitable for Singapore English. Leading commercial companies with speech recognition products such as Apple's SIRI, Google's speech API, Microsoft's Cortana, Amazon's Echo, amongst others, have not been able to understand the Singaporean accent with a reasonable level of accuracy. Their rate of accuracy remains poor as they lack specific domain knowledge related to lexical, syntactic, and semantic items. More importantly, these products have not been built using the local accent corpus and pronunciation dictionary. In order to make automatic speech recognition work for Singapore, it requires knowledge and research in several areas. For one, it requires a comprehensive analysis of the full phonetic system of Singapore English. It also demands for an understanding of how Singaporeans of different ethnic groups, age, educational and socioeconomic backgrounds sound. For the system to work in everyday contexts, it is also necessary to account for how the speech patterns of Singaporeans in interaction in different contexts, and to reasonably predict and decode codeswitching across different languages. Automatic speech recognition in the Singaporean context is therefore challenging because we know very little about how Singaporeans speak.

This proposed research therefore aims to understand the ways Singaporeans speak framed by the following broad questions:

1. What are the key phonetic features that make up the prototypical "Singaporean" English speech?
2. What are the phonetic features that differentiate the speech of Singaporeans along the dimensions of education, age, ethnicity, and social class?
3. How can one create innovations in speech technology that are suitable for use in Singapore?

Corresponding to the above questions, this project therefore has the following specific aims:

- Aim 1: To build a comprehensive model of the Singapore English sound inventory;
- Aim 2: To identify speech patterns that correlate to Singaporeans along the dimensions of education, age, ethnicity and social class;
- Aim 3: To build an automatic speech recognition system that can automatically understand and transcribe Singaporean speech, including codeswitched interactions, as a technological innovation for Singapore.

### SCOPE OF WORK FOR SELECTED PHD STUDENT

We are not expecting a single PhD student to carry out the entire project, but this student should be able to provide analyses of a few of the following:

1. The consonantal and vowel features of Singapore English;
2. The prosodic system of Singaporean speakers;
3. The phonetic and prosodic system of Singaporean speakers along dimensions of class, education, ethnic group, and age; and then
4. Build a speech/voice recognition software based on the above.

The project itself is large enough to be able to support a couple of students, and our idea is to have the PhD student work out which component in the project would suit him/her best.