

Speech Recognition for Lightly Code-Switched Speech

Summary

School of Information Technologies, Department of Software Sciences offers a 4-year PhD position in language technology.

Research field:	Information and communication technology
Supervisor:	Tanel Alumäe
Availability:	This position is available.
Offered by:	School of Information Technologies Department of Software Science
Application deadline:	Applications are accepted between June 01, 2020 00:00 and July 03, 2020 23:59 (Europe/Zurich)

Description

In recent years, automatic speech recognition has enjoyed huge success. Mostly due to improvements in using deep neural networks, accurate speech recognition systems have become widely available for a variety of languages. In some cases, the accuracy of speech recognition system have been claimed to reach human parity.

Code-switching is defined as the fusion of two or more distinct languages within the same utterance by a speaker. Code-switching is a phenomenon that happens more commonly in spoken form than in written form. Code-switching has been mostly investigated in the context of multilingual communities, such as Cantonese-English and Spanish-English. Code-switching is also very common way of communication in India, where English, Hindi and other Indian languages can be used inside a single speech utterance.

However, minor code-switching also happens in many other languages, especially in certain domains. For example, in most languages it is common to frequently use English terms and multi-word expressions in technical talks and communication. Similarly, in the medical domain, Latin words could be used interchangeably with native words. This poses a problem for most speech recognition systems: typically, foreign language expressions cannot be properly decoded by single-language speech recognition models, since the phonemic and grammatical rules that are valid for the native language cannot be applied to the foreign languages words. This results in speech recognition errors when code-switching is used in the utterance. The problem is amplified by the fact that foreign language words carry usually important linguistic content that is highly valuable for understanding the meaning of the sentence. The inability of the speech recognition system to decode code-switched words reduces the value and usability of such systems for many potential practical use cases, such as transcribing technical lectures and meeting recordings.

This goal of this topic is to improve speech recognition systems to better handle code-switched speech. Code-switching is usually not characterized as the random mixing of the words or phrases from two or more languages. The switching between the languages appears to follow some broad syntactic rules. The topic will investigate improving both acoustic and language models, based on both monolingual corpora and multilingual data. The work will focus on light code-switching (i.e., when foreign words and expressions are infrequently inserted into the native language speech), which has been relatively less researched than heavy code-switching that happens in multilingual contexts. The topic investigates linguistically motivated methods and data-driven approaches, as well as their hybrids.



To get more information or to apply online, visit <https://taltech.glowbase.com/positions/79> or scan the the code on the left with your smartphone.