

A multilingual approach towards improving the linguistic module of a TTS system: Case Navarro-Lapurdian dialect

María Andrea Cruz Blandón

Erasmus Mundus Master Program
in Language and Communication Technologies

Supervisors:

Prof. Inma Hernaez Rioja¹, Prof. Eva Navas Cordón¹, Prof. Denis Jouvét²

¹ University of the Basque Country

² University of Lorraine

Abstract

The Navarro-Lapurdian dialect is a Basque dialect spoken in the French side of the Basque country. This dialect differs from the standard Basque in terms of its phonology, as well as at the grammatical and lexical levels. Additionally, passages in this dialect are code-switched texts with French. TTS systems for this dialect need to handle both Navarro-Lapurdian and French phonemes repertoire. Inaccurate processing of the French words can result in using the Basque phonology to transcribe them or even in a wrong verbalisation. Previous TTS system has shown that failing to identify and correctly preprocess the French words cause a drop in the quality of the system.

In this work, we propose a multilingual approach for the linguistic module of the system to improve the phonetic transcription of French words. We included a language identification (LID) task at the first stage of the process and a multilingual Grapheme-to-Phoneme (G2P) model at the last stage. A Max-Entropy classifier and a Conditional Random Field (CRF) classifier are used to identify the language at the word-level. Besides, the Transformer architecture, a deep neural network, is used to train the multilingual G2P model. CRF outperforms the Max-Entropy classifier achieving a 0.828 F1-measure for the French words in the LID task, showing an improvement of 0.126 over the Max-Entropy classifier. The best G2P model trained on monolingual and code-switched sentences and tested on the code-switched corpus achieves a PER of 6.96% and a WER of 14.13%.

Keywords: Code-switching, Multilingual G2P, Language Identification, TTS systems, Deep Neural Networks, CRF classifier.