Master Thesis title / MAL izenburua

Adaptation of FastSpeech 2 text to speech architecture to Spanish and Basque

Proposer(s) / Proposatzailea(k): names / izenak

Eva Navas, Inma Hernáez

Contact / Kontaktua: email

eva.navas@ehu.eus

inma.hernaez@ehu.eus

Description / Deskribapena

FastSpeeh 2 is a deep neural network based text to speech system that is able to generate speech from text faster than other TTS architectures with comparable quality. Several free implementations of this architecture have been proposed (https://paperswithcode.com/method/fastspeech-2), mostly trained for English using LJSpeech dataset. The goal of this work is to adapt this architecture for Spanish and Basque. Speech data of enough quality and quantity will be located and curated and then used to adapt FasSpeech 2 architecture. The resulting synthetic voices will be evaluated and compared with the ones currently available at the laboratory.

Goals / Helburuak

Selecting the most suitable FastSpeech 2 implementation

Selecting and preparing the speech data

Adapting and training FastSpeech 2 for Spanish and Basque

Evaluating the quality of the resulting synthetic speech

Requirements / Betebeharrak

knowledge of Scripting languages

experience with programming in Python

Framework / Esparrua

Framework here if needed

Tasks and plan / Atazak eta plana

Tasks and plan here if needed