Automated Speech Segmentation: Example of an African Language

UMR 7309 • CNRS • AMU
PAROLE ET LANGAGE

Brigitte Bigi
brigitte.bigi@lpl-aix.fr

Speech segmentation is the process of identifying boundaries between speech units in the speech signal and determining where in time they occur.

Linguistic resources of the target language should be defined:
- a lexicon (the words to be recognized),
- a word dictionary (their pronunciations as a sequence of phonemes),
- an acoustic model (a stochastic representation of input waveform patterns per phoneme).

SPPAS software tool implements language-and-task-independent algorithms. This multilingual approach was applied to the african langage Naija (Nigerian pidgin).

We developped language resources for a tokenizer, an automatic speech system for predicting the pronunciation of the words and their segmentation.

SPPAS: the automatic annotation and analysis of speech

ANNOTATE

SPPAS is able to produce **automatically** speech annotations from a recorded speech sound and its orthographic transcription.

ANALYZE

SPPAS is helpful for the analysis of any annotated data: estimate statistical distributions, make requests, manage files, visualize annotations.

CONVERT

SPPAS offers a file converter from/to a wide range of formats: xra, TextGrid, eaf, trs...

Multi-platform (Linux, MacOS and Windows)
Open source

http://www.sppas.org/

Reference:
Brigitte Bigi (2015). SPPAS - Multi-lingual Approaches to the Automatic Annotation of Speech. The Phonetician, 111-112, pp. 54-69.

Nigerian Pidgin English: Naija

Post-creole
L1: 5 million people
L2: 70 million people

NaijaSynCor project:
A Corpus-based Macro-Syntactic Study of Naija
http://naijasyncor.huma-num.fr

Reference:
Brigitte Bigi, Bernard Caron, Oyelere S. Abiola (2017). Developing Resources for Automated Speech Processing of the African Language Naija (Nigerian Pidgin) LTC, pages 441-445, Poznan (Poland).

Recording → IPUs → Orthographic Transcription