

Myprosody

Prosody is the study of the tune and rhythm of speech and how these features contribute to meaning. Prosody is the study of those aspects of speech that typically apply to a level above that of the individual phoneme and very often to sequences of words (in prosodic phrases). Features above the level of the phoneme (or "segment") are referred to as suprasegmentals.

A phonetic study of prosody is a study of the suprasegmental features of speech. At the phonetic level, prosody is characterised by:

1. vocal pitch (fundamental frequency)
2. acoustic intensity
3. rhythm (phoneme and syllable duration)

MyProsody is a Python library for measuring these acoustic features of speech (simultaneous speech, high entropy) compared to ones of native speech. The acoustic features of native speech patterns have been observed and established by employing Machine Learning algorithms. An acoustic model (algorithm) breaks recorded utterances (48 kHz & 32 bit sampling rate and bit depth respectively) and detects syllable boundaries, fundamental frequency contours, and formants. Its built-in functions recognize/measures:

- Average_syll_pause_duration
- No._long_pause
- Speaking-time
- No._of_words_in_minutes
- Articulation_rate
- No._words_in_minutes
- formants_index
- f0_index ((f0 is for fundamental frequency)
- f0_quantile_25_index
- f0_quantile_50_index
- f0_quantile_75_index
- f0_std
- f0_max
- f0_min
- No._detected_vowel
- perc%._correct_vowel
- (f2/f1)_mean (1st and 2nd formant frequencies)
- (f2/f1)_std
- no._of_words
- no._of_pauses
- intonation_index
- (voiced_syll_count)/(no_of_pause)
- TOEFL_Scale_Score
- Score_Shannon_index
- speaking_rate
- gender_recognition
- speech_mood (semantic analysis)
- pronunciation_posterior_score

- articulation-rate
- speech rate
- filler words
- f0 statistics

The library was developed based upon the idea introduced by Klaus Zechner et al “Automatic scoring of non-native spontaneous speech in tests of spoken English” Speech Communication vol 51-2009, Nivja DeJong and Ton Wempe [1], Paul Boersma and David Weenink [2], Carlo Gussenhoven [3], S.M Witt and S.J. Young [4] , and Yannick Jadoul [5].

Peaks in intensity (dB) that are preceded and followed by dips in intensity are considered as potential syllable cores.

MyProsody is unique in its aim to provide a complete quantitative and analytical way to study acoustic features of a speech. Moreover, those features could be analysed further by employing Python's functionality to provide more fascinating insights into speech patterns.

This library is for Linguists, scientists, developers, speech and language therapy clinics and researchers.

Please note that MyProsody Analysis is currently in initial state though in active development. While the amount of functionality that is currently present is not huge, more will be added over the next few months.

Installation

Myprosody can be installed like any other Python library, using (a recent version of) the Python package manager pip, on Linux, macOS, and Windows:

pip install myprosody

or, to update your installed version to the latest release:

pip install -u myprosody

NOTE:

After installing Myprosody, download the files

1. myspsolution.praat
2. MLTRNL.praat
3. stats.csv

from

<https://github.com/Shahabks/myprosody>

and save in the directory where you will save audio files for analysis.

Audio files must be in *.wav format, recorded at 48 kHz sample frame and 24-32 bits of resolution.

To check how the myprosody functions behave, please check

EXAMPLES.docx on

<https://github.com/Shahabks/myprosody>

Development

Myprosody was developed by MYOLUTIONS Lab in Japan. It is part of New Generation of Voice Recognition and Acoustic & Language modelling Project in MYSOLUTIONS Lab. That is planned to rich the functionality of My-Voice Analysis by adding more advanced functions.

References and Acknowledgements

1. DeJong N.H, and Ton Wempe [2009]; “Praat script to detect syllable nuclei and measure speech rate automatically”; Behavior Research Methods, 41(2).385-390.
 2. Paul Boersma and David Weenink; <http://www.fon.hum.uva.nl/praat/>
 3. Gussenhoven C. [2002]; “ Intonation and Interpretation: Phonetics and Phonology”; Centre for Language Studies, Univerity of Nijmegen, The Netherlands.
 4. Witt S.M and Young S.J [2000]; “Phone-level pronunciation scoring and assessment or interactive language learning”; Speech Communication, 30 (2000) 95-108.
 5. Jadoul Y. <https://parselmouth.readthedocs.io/en/latest/installation.html>
-

MIT License

- Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

- The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.
- THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.