Contribution ID: **274**   Type: **Poster**

Speech Analysis Module: An Open-Source Audio Processing Library for Onset Detection and Stimuli Preparation in Psycholinguistics

Wednesday 23 July 2025 11:40 (20 minutes)

Poster

Speech Analysis Module: An Open-Source Audio Processing Library for Onset Detection and Stimuli Preparation in Psycholinguistics

Author

Emma Rico Martín

Affiliation

Instituto Universitario de Neurociencia. Universidad de La Laguna, Tenerife, Spain; Departamento de Psicología Cognitiva, Social y Organizacional. Universidad de La Laguna, Tenerife, Spain.

Abstract

Introduction:

Language research has gained increasing importance in recent years, especially with the rise of AI. Precise measurement of naming reaction times is critical in psycholinguistic research, particularly in experimental psychology paradigms involving spoken responses. Although some approaches accurately detect voice onset, most fail or provide less accurate predictions in the presence of noise or poor-quality signals, and some are not free. Another common requirement in psycholinguistic research is the preparation of auditory stimuli, along with other essential tasks such as subject response transcription, preprocessing, and standardization. Here we present a novel Python-based audio processing library that integrates several open-source libraries—such as Librosa, SciPy, NumPy, Matplotlib and Whisper—to facilitate robust onset detection and stimuli preparation. To evaluate its performance, we conducted a naming experiment, demonstrating the library's effectiveness in accurately estimating reaction times from spoken responses.

Methods:

The library encompasses a comprehensive suite of functions tailored to experimental settings. Key functionalities include:

• Stimuli Synchronization: Audio files are systematically retrieved and prepared using functions that trim silence (via Librosa's trimming algorithms), adjusting the duration with precise zero-padding. This is useful for stimuli preparation, particularly in neuroscience research where precise synchronization with brain data is essential.

• Signal Filtering: A Butterworth bandpass filter can be applied to refine the audio signal by reducing noise and isolating specific frequency bands typically associated with the human voice.

• Noise Gating: A dynamic noise gate further enhances signal quality by attenuating background noise below a predetermined threshold.

• Onset Detection: After applying signal filtering and noise gating, the library leverages Librosa's onset strength envelope to identify transient events. By computing the energy envelope of the audio signal and applying backtracking and sensitivity adjustments, this approach ensures accurate detection of speech onsets, crucial for measuring reaction times.

• Visualization and Spectral Analysis: The library offers visualization of waveforms and spectrograms, allowing researchers to verify signal quality and processing outcomes.

To evaluate the effectiveness of this approach, we compared the accuracy of the library's automatic onset

detection in a naming experiment with manual selection using Praat visualization. Additionally, we provide an example of audio preprocessing in a natural speech inhibition task.

Results:

Analysis revealed that the automatic naming reaction time estimator performed robustly, with an absolute difference of 38.29 ms and an R^2 of 0.87 when compared against manual measurements across 322 audio files. These metrics indicate high accuracy and reliability in capturing the temporal dynamics of spoken responses.

Conclusions:

This open-source audio processing library offers a versatile, transparent, and effective tool for psycholinguistic research. By integrating well-established Python libraries, it simplifies the complex process of audio signal processing—from onset detection to stimuli preparation—thus reducing manual effort and increasing measurement precision in language research. The encouraging results obtained underscore its potential as a valuable resource for experimental psychology, enabling researchers to derive meaningful insights from auditory data with enhanced accuracy.

Keywords

Psycholinguistics, Onset-Detection, Speech-Analysis, Audio-Processing, Naming

Primary author: Ms RICO MARTÍN, Emma (Instituto Universitario de Neurociencia, Universidad de La Laguna, Tenerife, Spain; Departamento de Psicología Cognitiva, Social y Organizacional. Universidad de La Laguna, Tenerife, Spain.)

Co-authors: Dr BIRBA, Agustina (Cognitive Neuroscience Center. University of San Andrés, Vito Dumas 284, B1644BID, Victoria, Buenos Aires, Argentina); Dr DOMÍNGUEZ MARTÍNEZ, Alberto (Instituto Universitario de Neurociencia, Universidad de La Laguna, Tenerife, Spain; Departamento de Psicología Cognitiva, Social y Organizacional. Universidad de La Laguna, Tenerife, Spain.); Dr BERMÚDEZ MARGARETTO, Beatriz (Departamento de Psicología Básica, Psicobiología y Metodología de las Ciencias del Comportamiento, Facultad de Psicología, Universidad de Salamanca, Salamanca, Spain; Instituto de Integración en la Comunidad - INICO, Universidad de Salamanca, Salamanca, Spain.)

Presenter: Ms RICO MARTÍN, Emma (Instituto Universitario de Neurociencia, Universidad de La Laguna, Tenerife, Spain; Departamento de Psicología Cognitiva, Social y Organizacional. Universidad de La Laguna, Tenerife, Spain.)

Session Classification: Poster Session 1

Track Classification: Applications/Substantive areas: Applications/Substantive areas