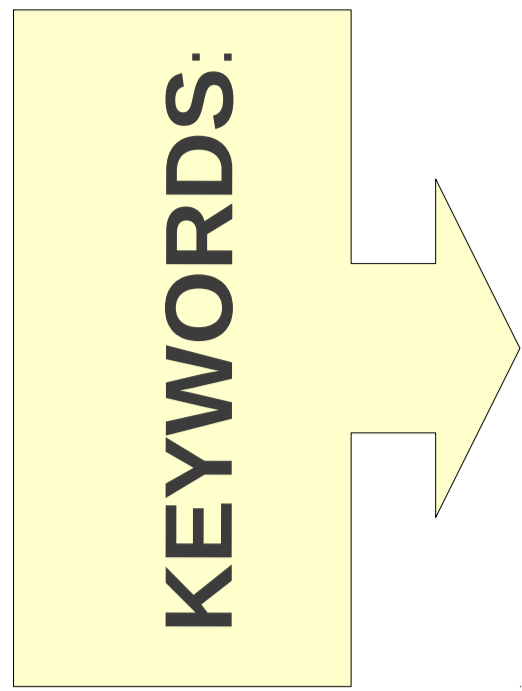


SPPAS: a tool for the automatic analysis of speech prosody.

Brigitte Bigi, Daniel Hirst

Laboratoire Parole et Langage, CNRS & Aix-Marseille Université
5 avenue Pasteur, BP 80975, 13604 Aix en Provence, France
{brigitte.bigi, daniel.hirst}@lpl-aix.fr



Segmentation

Automatic Syllabification

Annotation Phonetization

Speech

Alignment

Phonetic

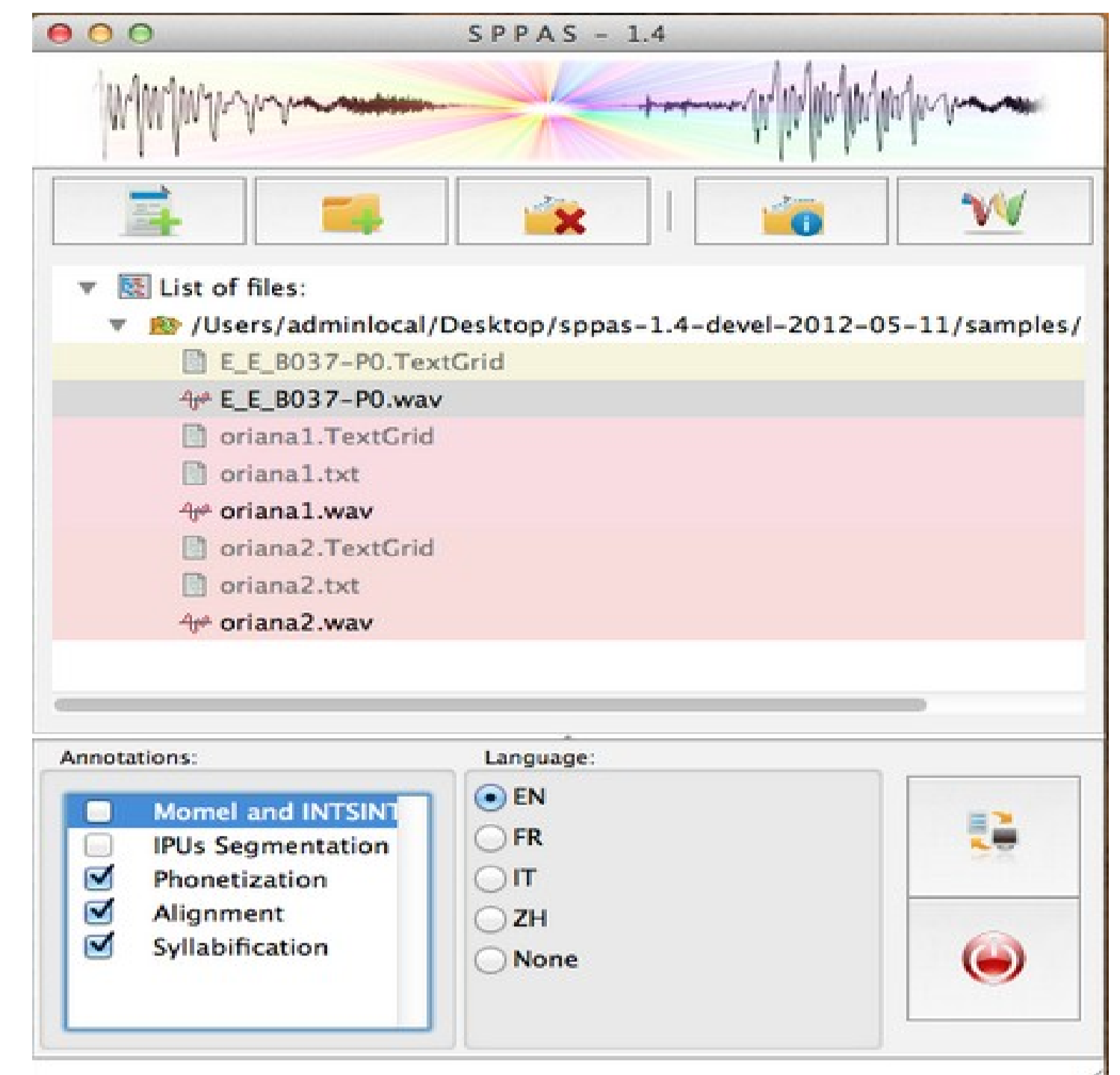
Prosody

Overview:

- A tool to produce automatically annotations which includes utterance, word, syllabic and phonemic segmentations from a recorded speech sound and its transcription
- Language-independent
- Currently designed for French, English, Italian and Chinese
 - and there is an easy way to add other Languages
- GNU Public License



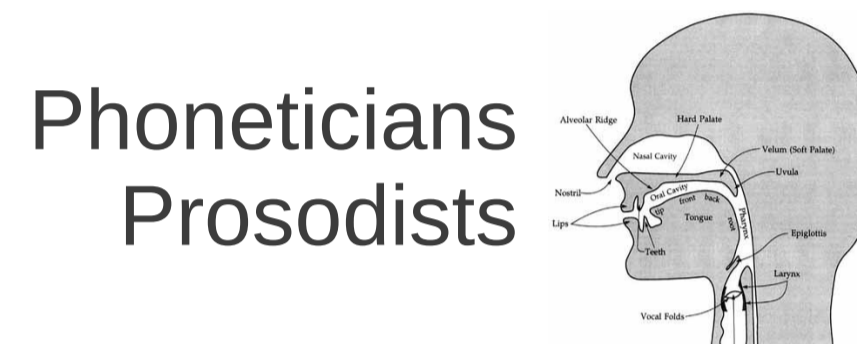
<http://www.lpl-aix.fr/~bigi/sppas/>



SPPAS in the Speech Tools communities



Computer scientists



Phoneticians Prosodists

Doing speech technologies with computers

Public tools

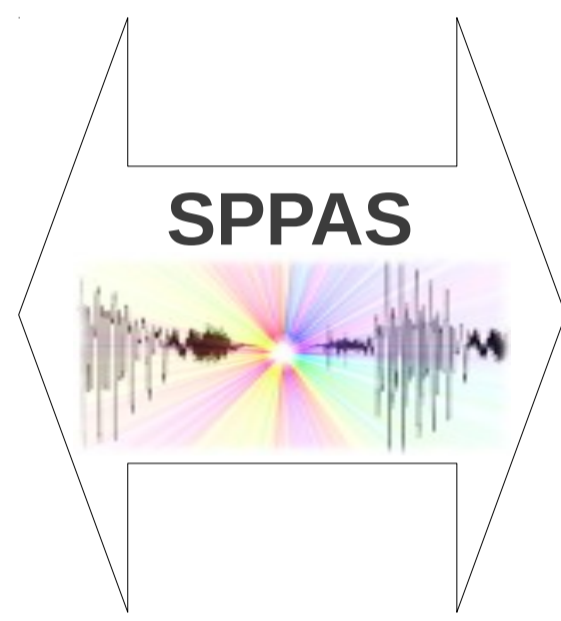
→ Julius CSR



Local tools

Mixed Public/Local tools

Always automatic annotations



Doing phonetic/prosody with computers

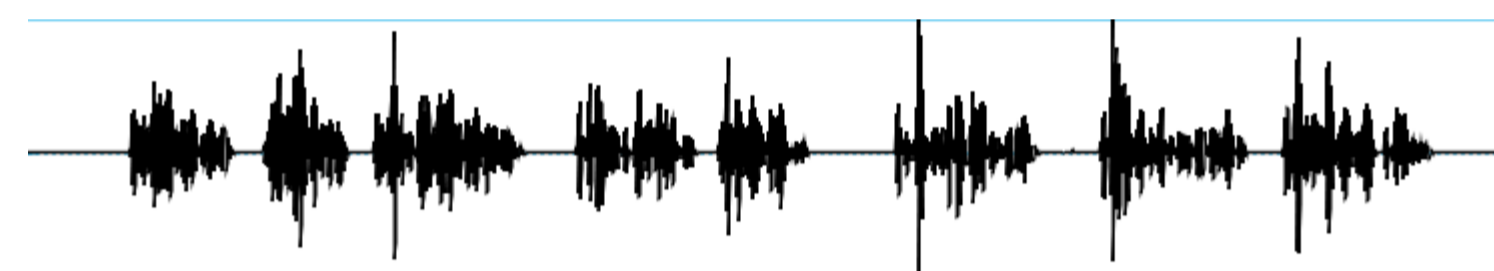
Public tools

→ Praat

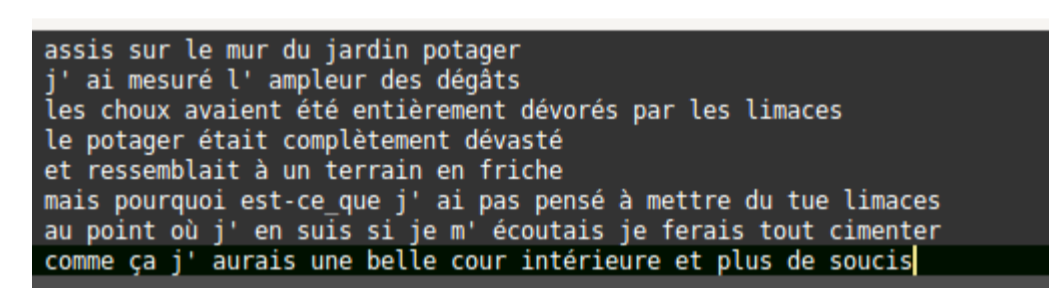
Other tools

Most of time manual annotations

SPPAS inputs:

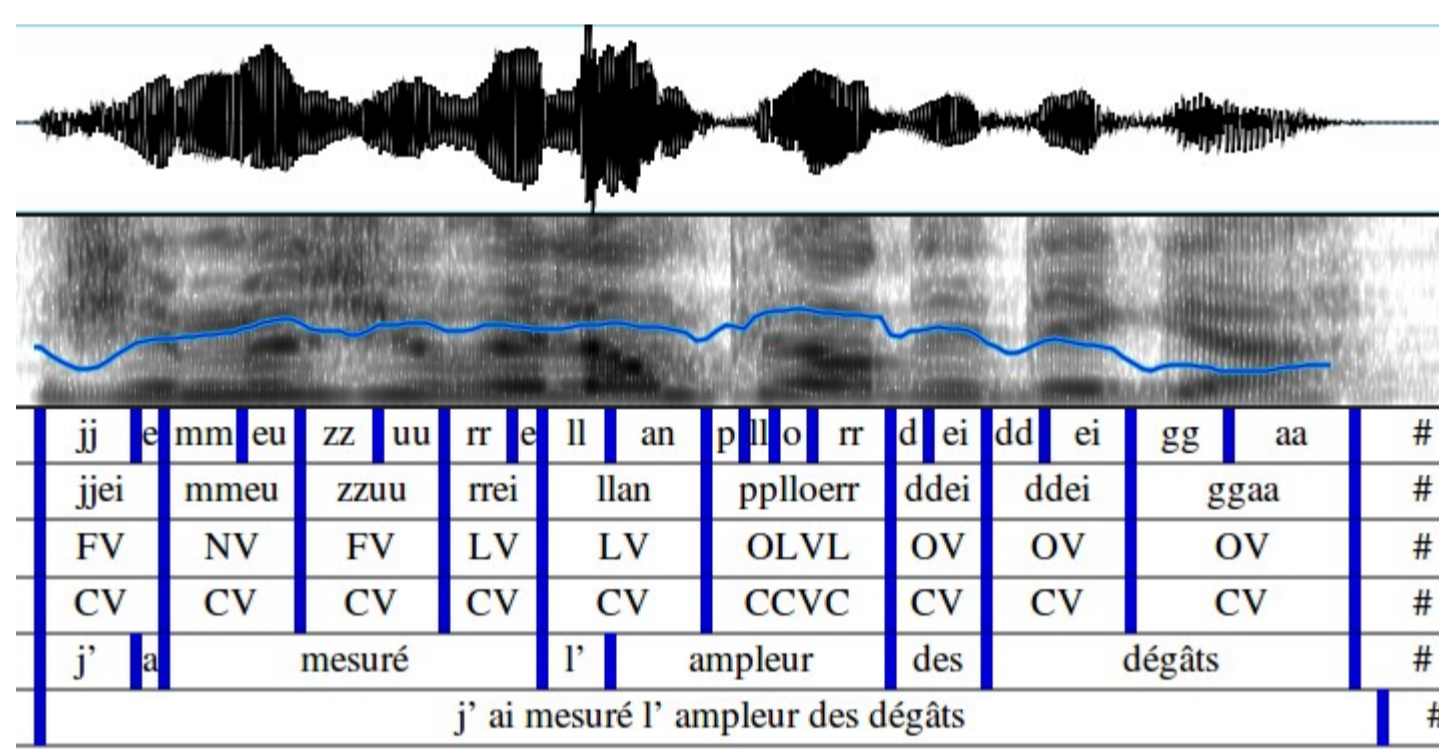


A speech wav file



A transcription (time-aligned or not)

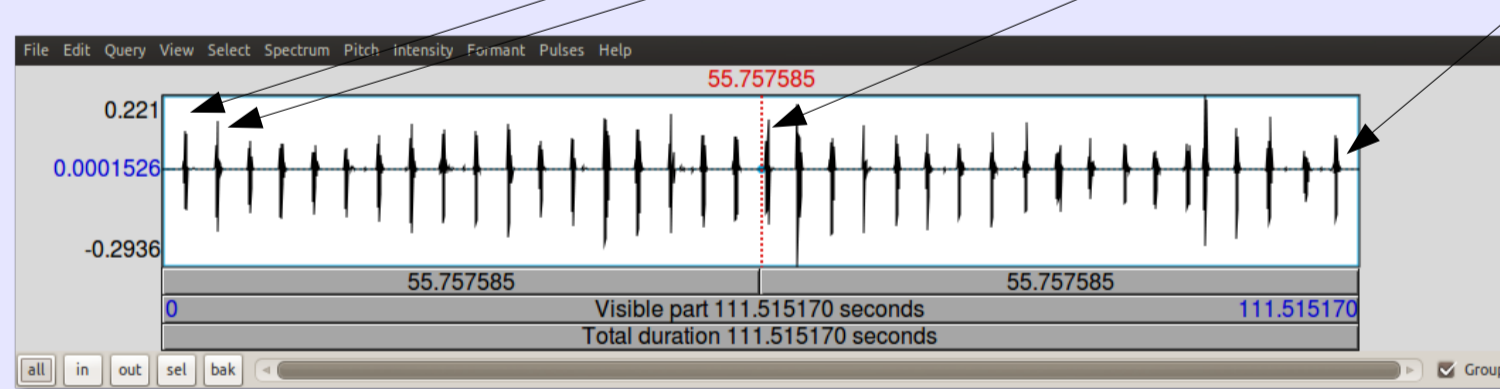
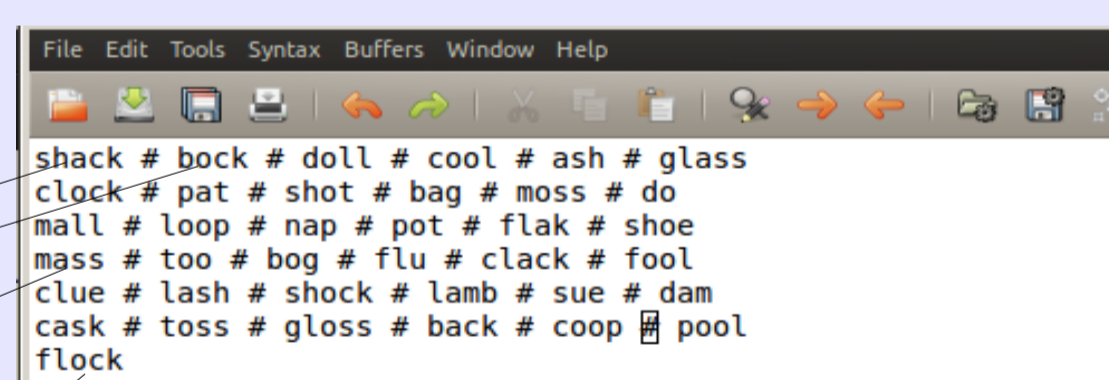
SPPAS outputs:



A set of TextGrid files

Inter-Pausal Units segmentation

The algorithm computes a heuristics based on the detection of silences



Phonetization

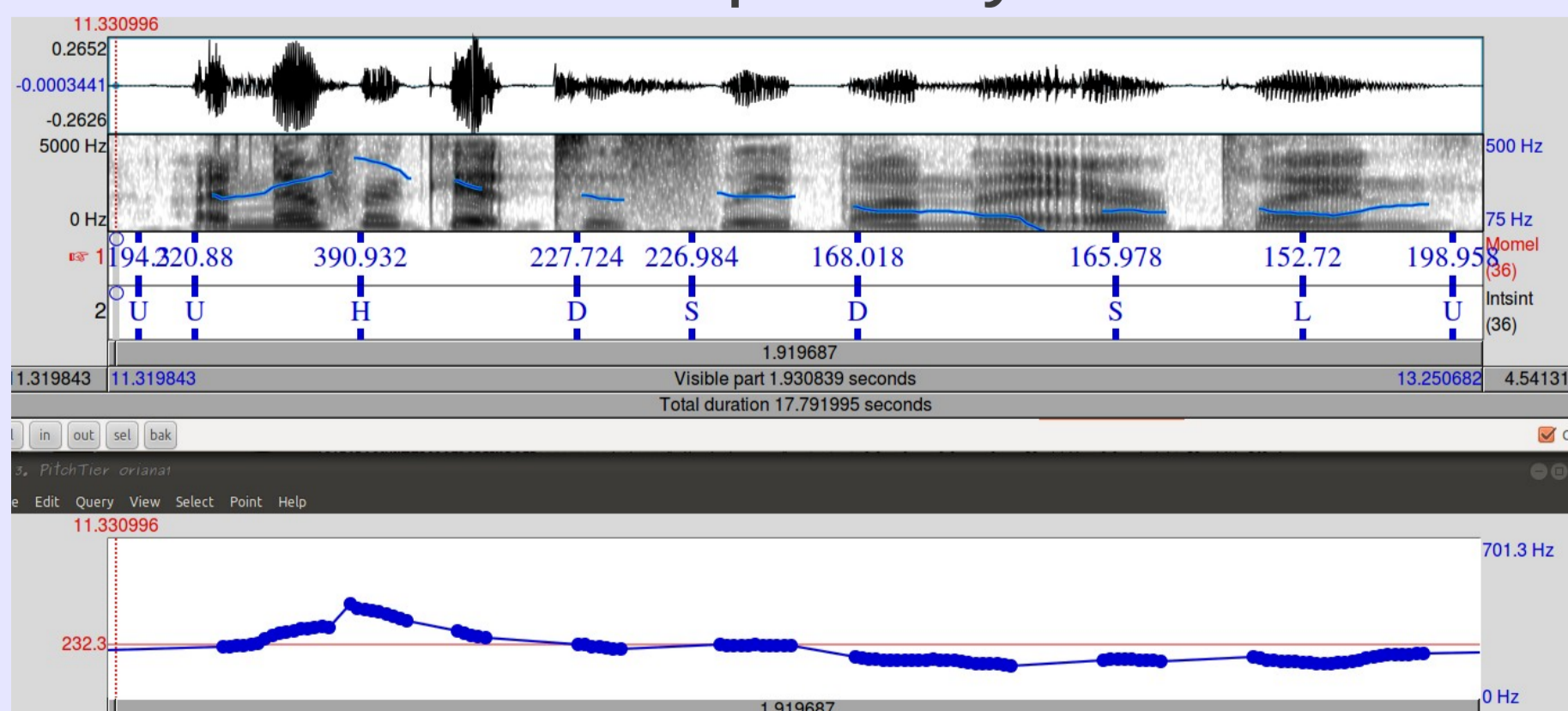
- Process of representing sounds with phonetic signs.
- SPPAS uses the dictionary-based approach:
 - the phonetization is the equivalent of a sequence of dictionary-look-ups.
 - Phonetic variants are proposed to the aligner.

	FR	IT	ZH	EN
Dictionary:	350k words and 300k variants	390k words and 5k variants	350 syllables	121k words and 10k variants
Acoustic model:	Triphones	Triphones	Monophones	Triphones
Data to train	7h30 Conversational +30min read	3h30 map-task	50min read	See voxforge.org

Available resources in SPPAS

Momel and INTSINT

Momel: (Modelling melody) is an algorithm for the analysis and synthesis of intonation patterns.
INTSINT: International Transcription System for INTonation.



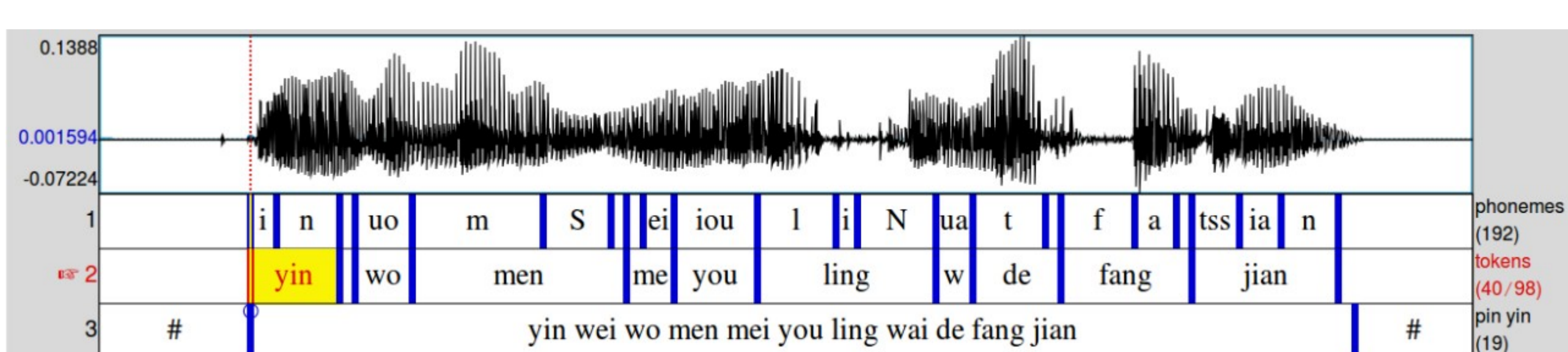
Alignment

- Forced-alignment in SPPAS is based on Julius. Need:
 - a finite state grammar that describes sentence patterns to be recognized;
 - an acoustic model.
- The alignment task is a 2-step process:
 - the first one: choose the phonetization;
 - the second one: perform the segmentation.

Syllabification

- French or Italian
- A Rule-Based System for automatic syllabification of phonemes' strings. The syllabification is based on 2 principles:
 - a syllable contains a vowel, and only one;
 - a pause is a syllable boundary.

V C C V



SPPAS output example for the utterance 因为我们没有另外的房间 (Because we do not have another room)
Pinyin: Yinwei wómen méiyǒu língwài de fāngjiān