



Automatic Phonetic Labeling with SPPAS

Brigitte Bigi

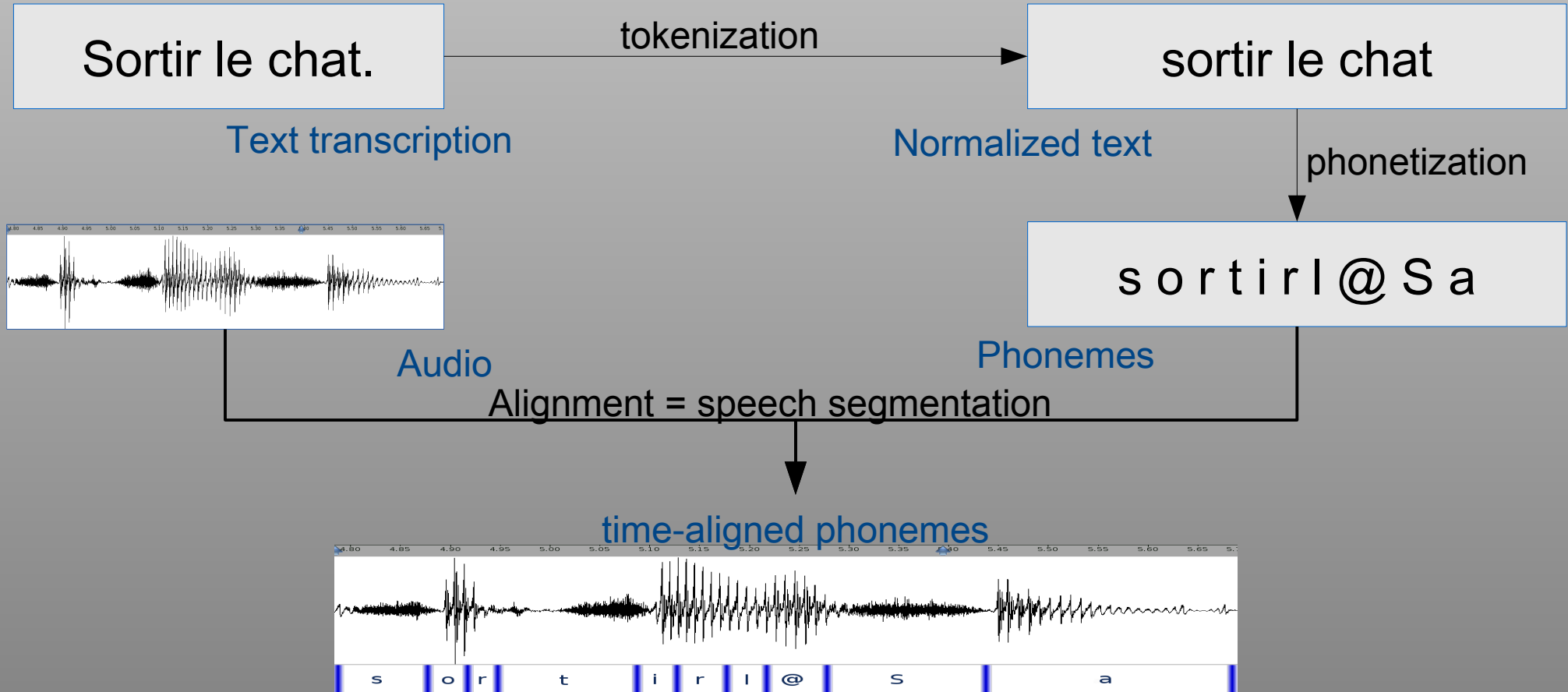
LPL - Aix-en-Provence - France



This work has been carried out thanks to the support of the A*MIDEX project (n° ANR-11-IDEX-0001-02) funded by the « Investissements d'Avenir » French Government program, managed by the French National Research Agency (ANR)

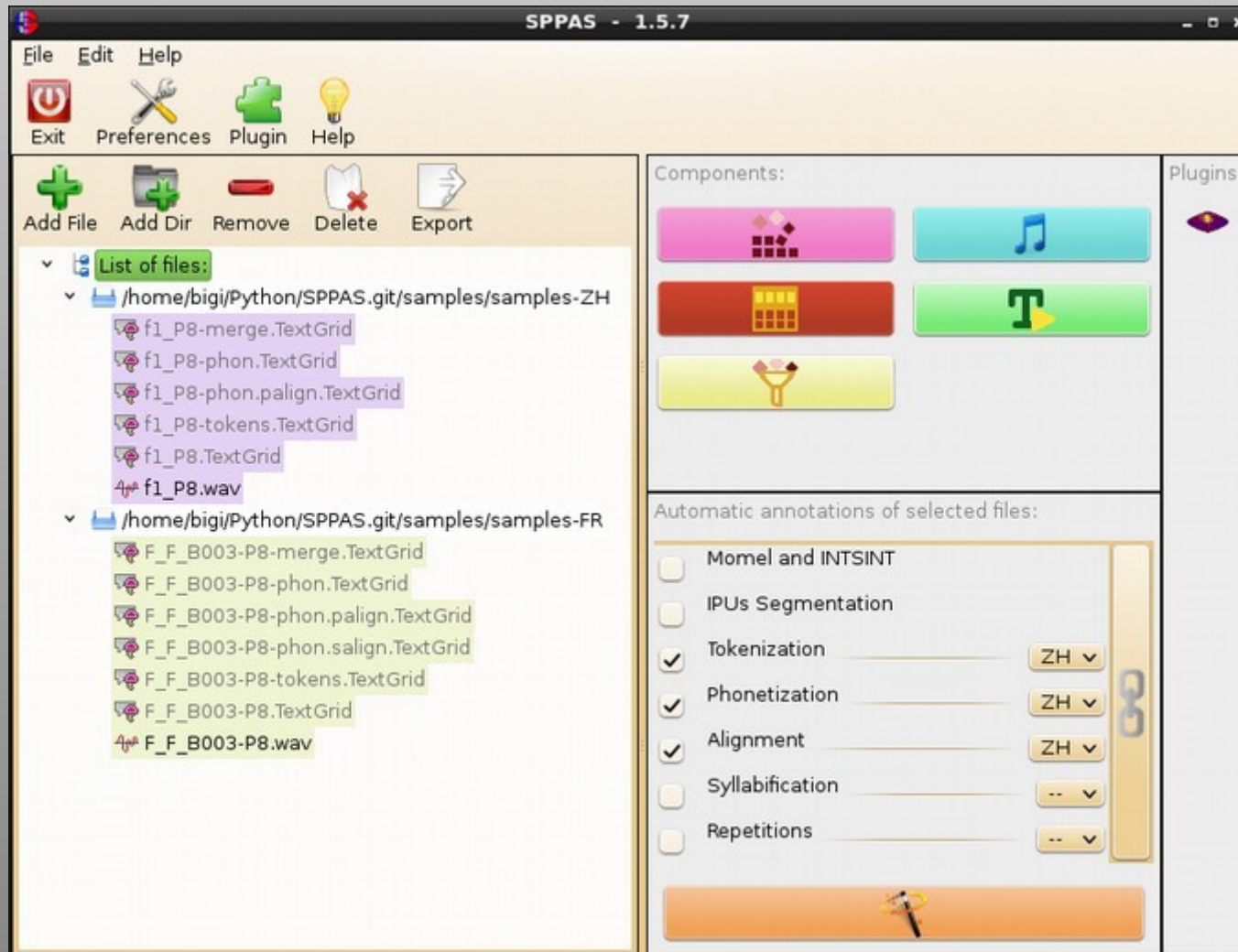
Automatic Phonetic Labeling

- the process of taking the *text transcription* of an audio speech segment and determining where in time particular phonemes occur in the speech segment



SPPAS

◆ Graphical User Interface: (Bigi, 2012)





Key-points



◆ Language-independent algorithms

- Tokenization: (Bigi, 2011),
- Phonetization: (Bigi, 2013)
- Alignment: (Bigi, 2012)

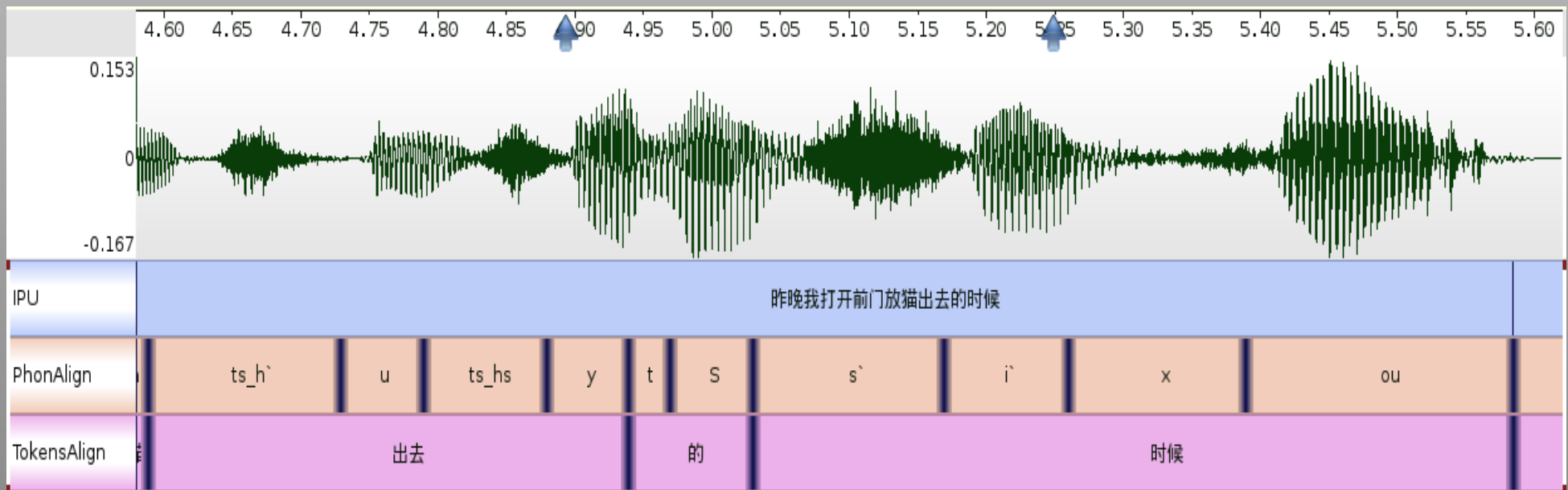
◆ A tool dedicated to computer scientists **and** linguists

◆ GNU Public License



On which languages?

- ◆ SPPAS can perform forced-alignment of:
 - French, English, Italian, Spanish, Chinese, Taiwanese, Japanese.





Orthographic transcription



- ◆ Manual annotation (for 1 min of speech → 20 to 75 min)
- ◆ Representation of what is “percieved” in the signal
- ◆ It must includes:
 - Filled pauses
 - Percieved short pauses
 - Repeats
 - Noises and Laugh (currently only for French)
- ◆ It can includes:
 - Specific pronunciations
 - Elisions



Requirements: tokenization



- ◆ Tokenization: *A list of words (lexicon)*
- ◆ Transcription:
 - Et euh donc donc du coup c'est toi c'est un peu toi q(ui) a les premiers contacts avec le avec le gosse quoi + et puis là ils te demandent le prénom donc faut ce soit prêt là @ parce que putain.
 - 昨晚我打开前门放猫出去的时候
- ◆ Tokenized:
 - et euh donc donc du coup c' est toi c'est un_peu toi qui a les premiers contacts avec le avec le gosse quoi + et puis là ils te demandent le prénom donc faut ce soit prêt là @ parce_que putain
 - 昨晚 我 打开 前门 放 猫 出去 的 时候



Requirements: phonetization



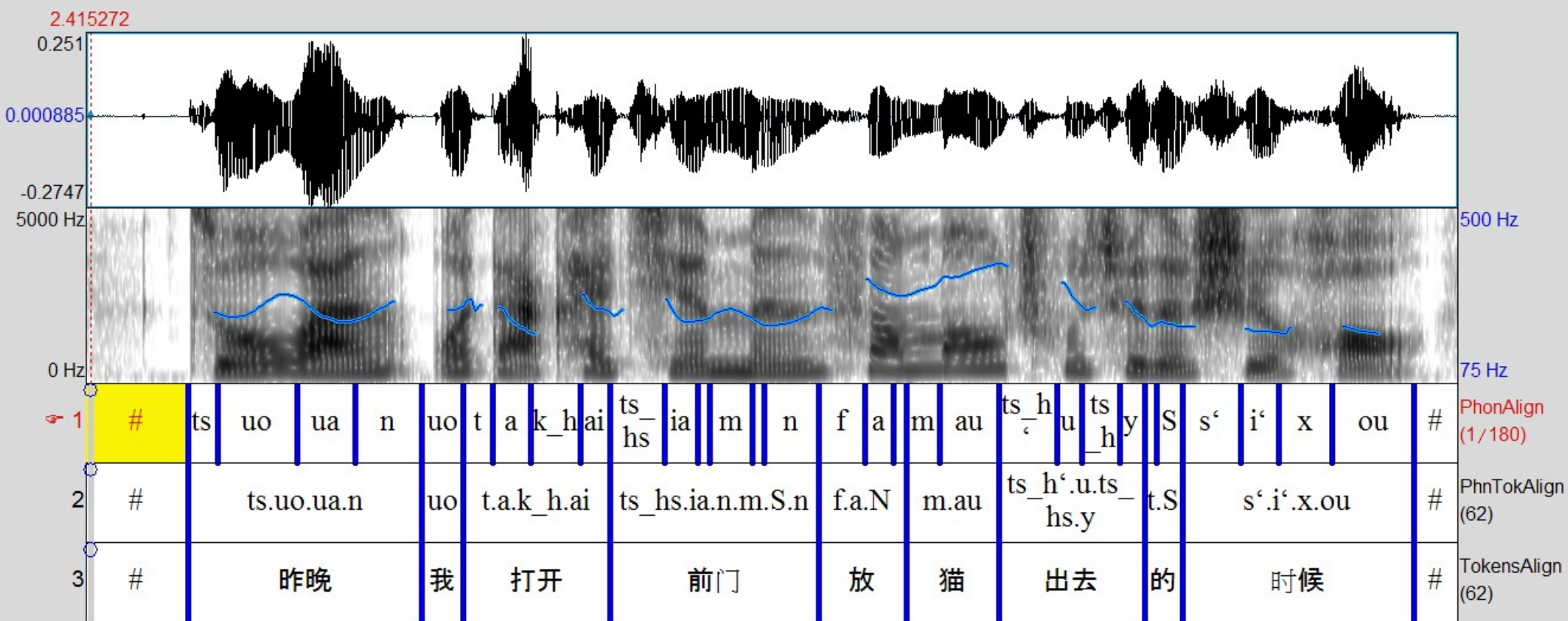
- ◆ Phonetization: **A pronunciation dictionary**
- ◆ e|E|E.t fp d.o~.k|d.o~ d.o~.k|d.o~ d.y.k.u s.e t.w.a|t.w.a s.e U~.p.EU|
U~.p.EU t.w.a|t.w.a k.i|k.j|k a l.E|l.E.z|l.e|l.e.z|l.EU p.r.EU.m.j.e|
p.r.EU.m.j.e.z|p.r.EU.m.j.e|p.r.EU.m.j.e.z k.o~.t.a.k.t.z|k.o~.t.a.k.t
a.v.E.k|a.v.E.k |l.EU a.v.E.k|a.v.E.k |l.EU g.o.s.EU|g.o.s k.w.a|k.w.a sil
e.p.H.i|e.p.H.i|e.p.H.i.z|E.p.H.i|E.p.H.i.z l.a|l.a i.|i.l.z|i|i.z t.EU|t
d.EU.m.a~.d.EU.t|d.EU.m.a~.d |l.EU p.r.e.n.o~|p.r.e.n.o~|p.r.e.n.o~
d.o~.k|d.o~ f.o.t|f.o|f s.EU|s s.w.a|s.w.a.t|s.w.a|s.w.a.t p.r.E|p.r.E.t|
p.r.e.t|p.r.e l.a|l.a @@ p.a.r.s.k|p.a.r.s.EU.k.EU|p.s.k|p.a.s.k.EU|
p.a.s.EU p.y.t.U~
- ◆ ts.uo.ua.n uo t.a.k_h.ai ts_hs.ia.n.m.S.n|ts_hs.ia.n.m.S.n f.a.N m.au
ts_h` .u.ts_hs.y t.S|t.i|t.i s`.i`.x.ou



Requirements: alignment



- ◆ Alignment: An acoustic model





Advices to include a new language



- ◆ A lexicon
- ◆ A pronunciation dictionary
- ◆ Acoustic model:
 - Requirements for a Monophone:
 - at least 3 minutes of time-aligned data
 - 30-60 minutes manually phonetized data
 - Requirements for a Triphone:
 - at least 8 hours of “well”-transcribed speech
 - From these data, I can train an acoustic model and add the new language in SPPAS!

Perspectives: Variamu Project

- ◆ Forced Alignment on Children Speech (FACS)
 - FA = Phonetization + Speech Segmentation (Bigi, 2011)
 - EVALITA 2014.



- ◆ Creating a multilingual model:

- Speech segmentation of an un-trained language
- Speech segmentation of a known-language even if code-switching (with another known-language)





References



- ◆ Bigi, B. (2012). The SPPAS participation to Evalita 2011. In EVALITA 2011: Workshop on Evaluation of NLP and Speech Tools for Italian.
- ◆ Bigi, B. (2012). SPPAS: a tool for the phonetic segmentation of speech. In LREC (Vol. 8, pp. 1748-1754).
- ◆ Bigi, B., & Hirst, D. (2012). SPeech Phonetization Alignment and Syllabification (SPPAS): a tool for the automatic analysis of speech prosody. In Proc. of Speech Prosody (Vol. 2012).
- ◆ Bigi, B., Péri, P., & Bertrand, R. (2012). Orthographic Transcription: which Enrichment is required for phonetization?. In LREC (Vol. 8, pp. 1756-1763).
- ◆ Bigi, B. (2011). A multilingual tokenization approach. LTC'2011. Poznan.
- ◆ Bigi, B. (2013). A phonetization approach for the forced-alignment task. LTC'2013. Poznan.