

*Vers le développement d'une chaîne de traitement  
pour l'alignement automatique d'une transcription  
orthographique à un signal sonore*

**Nicolas DAVID**

(Doctorant – Laboratoire LIDILEM – Université Grenoble Alpes)

[davidnic@univ-grenoble-alpes.fr](mailto:davidnic@univ-grenoble-alpes.fr)

**Séminaire DéLiCorTAL – Université Grenoble Alpes**

**22 janvier 2021**

# Sommaire

1. Introduction et présentation générale
2. Méthodologie de l'alignement automatique
3. Esquisse d'une chaîne de traitement

# Introduction et présentation générale

## Cadre du travail

- Discours scientifique oral (routines sémantico-rhétoriques)
- Corpus multimodal (enregistrements audio-visuels | transcriptions)
- Outils de traitement automatique (TAL)
- Champs de recherche et d'analyse pluridisciplinaires

# Introduction et présentation générale

## Objectifs du travail

- Réaliser la transcription orthographique
  - Conventions de transcription et de segmentation
- Procéder à la vérification des transcriptions
  - Accord inter-annotateur
- Parvenir à l'alignement automatique
  - Structuration du corpus

# Introduction et présentation générale

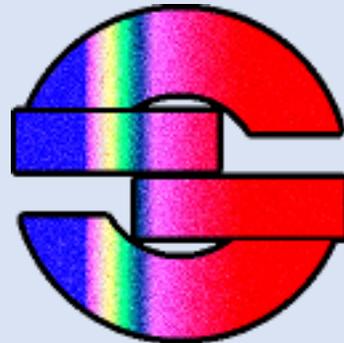
## Problématique

*Comment aboutir au développement d'une chaîne de traitement afin d'harmoniser et d'automatiser les diverses étapes du travail ?*

# Méthodologie de l'alignement automatique

**SPPAS**

**(SPeech Phonetization Alignment and Syllabification)**



# Méthodologie de l'alignement automatique

## SPPAS (Bigi, 2020)

- Créé par Brigitte Bigi en 2011 (LPL | Aix-Marseille Université)
- Outil d'analyse et d'annotation automatiques du signal de parole
- Développé sous Python (*GNU – General Public License*)
- Implémente un algorithme basé sur des HMM

# Méthodologie de l'alignement automatique

## SPPAS – Principe de fonctionnement

- **Données d'entrée (*input*)**
  - Enregistrement sonore
  - Transcription orthographique pré-segmentée
  - *PitchTier*
- **Résultats de sortie (*output*)**
  - Fichiers d'alignement : *tokens*, phonèmes, transcription orthographique et transcription phonétique

# Méthodologie de l'alignement automatique

## SPPAS – Formats de sortie

- |   |  |
|---|--|
| <ul style="list-style-type: none"><li>• SPPAS : xra</li><li>• Praat : TextGrid   PitchTier</li><li>• ELAN : eaf</li><li>• HTK : lab   mlf</li></ul> | <ul style="list-style-type: none"><li>• Sclite : ctm   stm</li><li>• Phonedit : mrk</li><li>• Tableurs : csv</li><li>• Sous-titres : sub   srt</li></ul> |
|---|--|

# Méthodologie de l'alignement automatique

## SPPAS – Contraintes, difficultés et solutions

- Nécessité d'avoir des données d'entrée « lissées »
- Diverses exigences au niveau des pré-traitements
- Échec tantôt partiel, tantôt total de l'alignement
- Recourir à un alignement forcé

# Méthodologie de l'alignement automatique

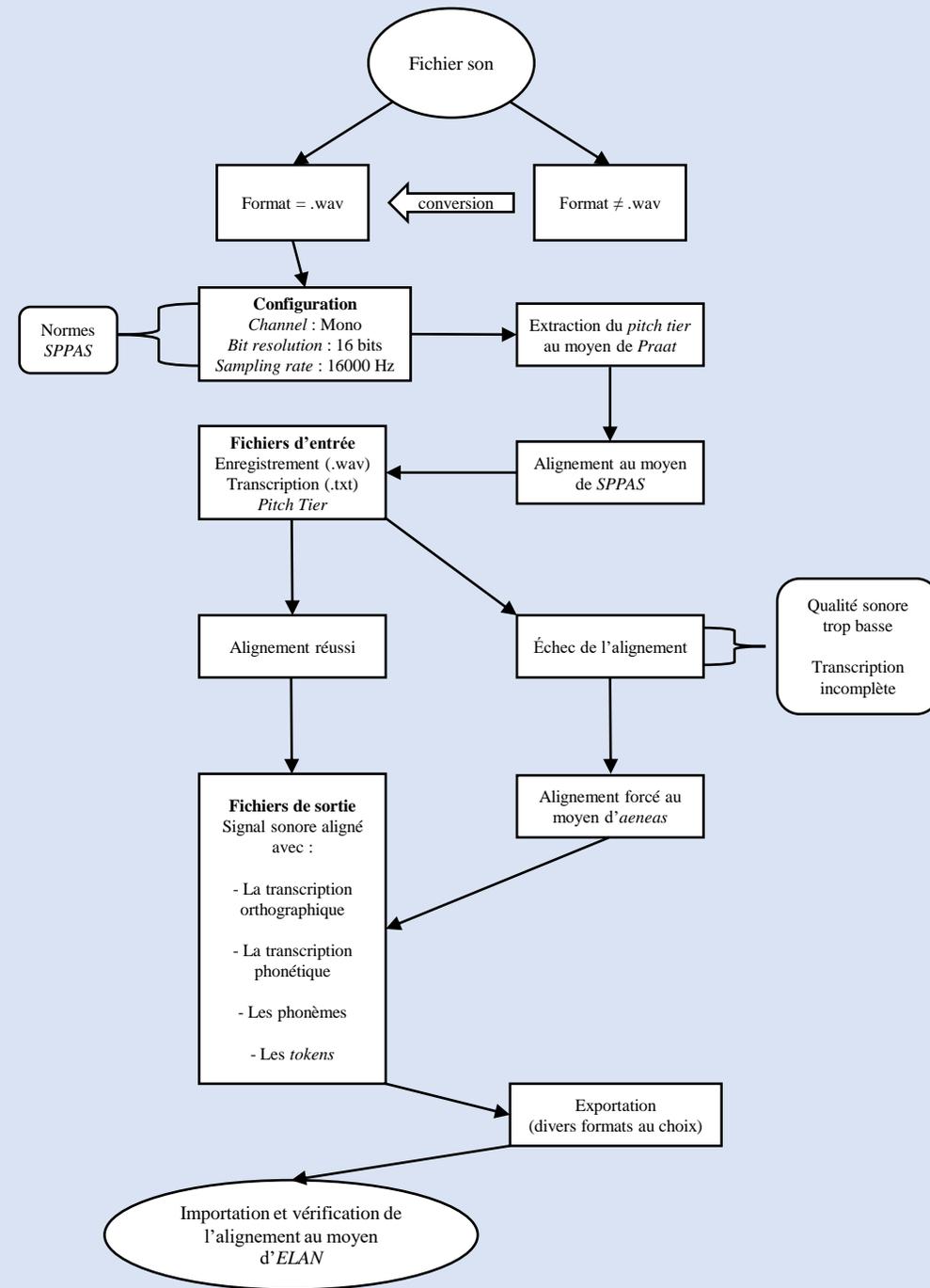
## **aeneas** **(ReadBeyond, 2017)**

- Bibliothèque Python et C
- Boîte à outils « d'alignement forcé »
- Développé sous Python et C (*GNU – Affero General Public License*)
- Implémente un algorithme DTW

# Esquisse d'une chaîne de traitement

*Le pipeline ALTS*  
**(David, Lee et Tutin, 2020)**

**A**Lignement automatique d'une **T**ranscription orthographique à un **S**ignal sonore



**Merci de votre attention !**

# Références bibliographiques

- Bigi, B. (2015). SPPAS – Multi-lingual Approaches to the Automatic Annotation of Speech. In “the Phonetician” – International Society of Phonetic Sciences, ISSN 0741-6164, Number 111-112 / 2015-I-II, pages 54-69.
- Bigi, B. & Hirst, D. (2013). What’s new in SPPAS 1.5?, Tools and Resources for the Analysis of Speech Prosody, Aix-en-Provence, France, pp. 62-65.
- Bigi, B. (2012). “SPPAS: a tool for the phonetic segmentations of Speech”, Language Resources and Evaluation Conference, ISBN 978-2-9517408-7-7, pages 1748-1755, Istanbul (Turkey).
- Bigi, B. & Hirst, D. (2012). “SPeech Phonetization Alignment and Syllabification (SPPAS): a tool for the automatic analysis of speech prosody”, Speech Prosody, Tongji University Press, ISBN 978-7-5608-4869-3, pages 19-22, Shanghai (China).

# Références bibliographiques

- Leung, H. & Zue, V. W. (1984). A procedure for automatic alignment of phonetic transcription with continuous speech. *Proceedings of ICASSP 1984*: 73-76.
- Pellegrini, T. (2008). *Transcription automatique de langues peu dotées*. Paris : Université Paris-Sud – Paris XI.
- Wightman, C. & Talkin, D. (1997). The Aligner: Text to speech alignment using Markov Models. In J. van Santen, R. Sproat, J. Olive, and J. Hirschberg (ed.), *Progress in Speech Synthesis*, 313-323, Springer Verlag, New York.
- Yuan, J., Lai, W., Cieri, C., & Liberman, M. (2018). *Using Forced Alignment for Phonetics Research*. Philadelphia: University of Pennsylvania.