Accelerating the analysis of your audio recordings with Untrained Forced Speech Alignment

Beginning Alignment: A new recording

Go to http://icldc-align.appspot.com/ Before you start, you will need to download this free software: Praat http://www.fon.hum.uva.nl/praat/ If you are using a Mac computer, you also need to download Sublime Text from https://sublime-text.en.softonic.com/mac If you want to use textedit on the Mac, you can, but you need to save your files as plain text by selecting that option in the format menu (or shift+ command +T).

1) **Prepare your audio data**

   In Praat, select a very short audio segment (3-5 seconds):
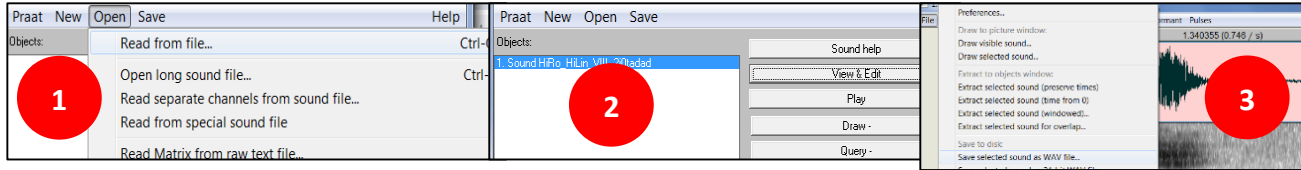
2) **Transcribe your data**

   Produce a transcription file in a plain text .txt file. Use Notepad on Windows, and Sublime on MacOS. Put each phrase on a separate line, with tabs separating each column. Use Praat to find out what the start and end time of each phrase is, in seconds. The format of the transcription file is the following (make sure you don't have blank lines at the end of the file, and do not use any punctuation).

   [speaker name]   [info]   [start time]   [end time]   [phrase]

   ```
   Jean    CIM    0.14    1.0     e oti reia
   Jean    CIM    1.6     2.6     ax
   ```
   Example:
   ```
   Jean    CIM    4       4.96    te openga i reia o texrax
   ```

3) **Create a dictionary**

   In another .txt file, specify the correspondence between word and phoneme string. Each unique word should be on a separate line, separated from the phonemes by a tab. Phonemes should be separated from each other by a space. You will use the ARPAbet transcription system.

   https://icldc-align.appspot.com/examples/arpabet.pdf

   (make sure you don't have blank lines at the end of the file, and do not use any punctuation).

   ```
   ax       AE1
   e        EH1
   i        IY1
   o        OW1
   openga   OW1 P EH1 NG AE1
   oti      OW1 T IY1
   reia     R EH1 IY1 AE1
   te       T EH1
   ```
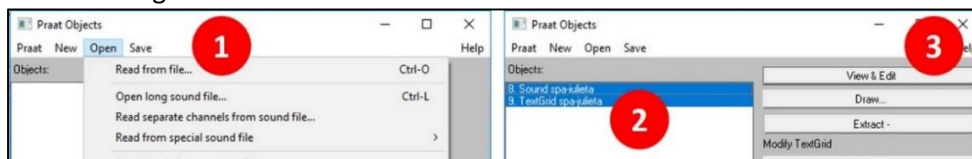   **Example:**
   ```
   texrax   T EH1 R AE1
   ```

4) **Align your recording!**

   Using the interface at https://icldc-align.appspot.com/aligner.jsp upload your audio file, your transcript and your dictionary file. Specify your email address and align!
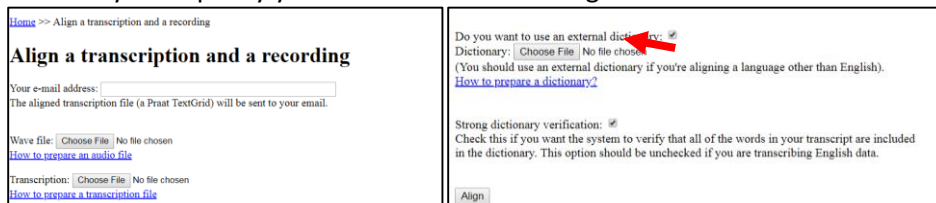
5) **Check your recording**

Extraction and visualization of phonetic information

6) Extraction of phonetic information

Next, you will automatically navigate pitch, formants, and other segmental information using the TextGrid you have produced. Navigate to https://icldc-align.appspot.com/extractphonetics.jsp in the interface. Write your email, upload your audio file, and the TextGrid produced by alignment. It is recommended to keep the remaining settings default unless you are comfortable with phonetic analysis.

Extract phonetic information

We **STRONGLY RECOMMEND** that you manually correct the alignment before you try to extract phonetic information from it. Once you have hand corrected it, you can use the TextGrid and its corresponding wave file to extract phonetic info from the recording (e.g. formants, duration, intensity).

How do I hand correct a TextGrid?

E-mail:

Wave file: Choose File No file chosen
How to prepare an audio file

TextGrid: Choose File No file chosen
How to automatically generate the TextGrid for a recording

[Extract phonetic values] [Reset to default values]

7) Visualize phonetic information

Using the phonetic information extracted in the previous step, you will produce a vowel chart. Navigate to https://icldc-align.appspot.com/uploadForTriangle.jsp in the interface. Write your email and upload the phonetic information produced in step 6. It is recommended to keep the default settings unless you are comfortable with phonetic analysis. After clicking "Generate vowel triangle," manually click the vowels you would like to plot.

Draw a vowel triangle

E-mail:

Phonetic information: Choose File No file chosen

[Generate vowel triangle] [Reset the form]

Advanced alignment: Automatic dictionary generation

8) Choose a new recording, and make a new transcription, following steps 1 and 2 above.

9) Make an ARPAbet equivalences file.

The equivalence file has two tab-separated columns. The first column has the letter in the orthography of your language (Warning: Unicode is not currently supported. Do not use accent marks in your orthography). The second column has the closest ARPAbet equivalent.

Example:

n	N
ng	NG
o	OW1
ox	OW1
p	P

10) Automatically generate a dictionary using your new ARPAbet equivalences file.

Navigate to https://icldc-align.appspot.com/generatedictionary.jsp in the interface. Upload your new transcription, and the ARPAbet equivalences file. You can optionally choose to augment the previous dictionary you created in step 3. Remember to put your email!

Generate a dictionary, or add new words to an existing dictionary

E-mail:

Transcription: Choose File No file chosen
How to prepare a transcription file

Arpabet Equivalences: Choose File No file chosen
How to prepare an ARPAbet equivalences file

Do you want to augment an already existing dictionary: ☐
Dictionary: Choose File No file chosen
How to prepare a dictionary? - When should I provide a dictionary?

Do you want to provide an already existing 3-column dictionary: ☐
3-Column dictionary: Choose File No file chosen
How to prepare a 3-column dictionary? - When should I provide a 3-column dictionary?

[Generate]

11) Align and correct, using your **new custom dictionary**, following the directions in step 4 and 5.