8b. MFA

The Montreal Forced Aligner (MFA) is a 3rd-party tool developed by Michael McAuliffe and others that can use words with phonemic transcriptions, and the corresponding audio, to force-align words and phones; i.e. determine the start and end time of each speech sound within each word, and thus the start/end times of the words.

The annotator can work in two modes:

- *Train and Align* - acoustic models are trained on the data you want to align, which can be in any language as long as you have a pronunciation dictionary for it.
- *Pre-trained Models/Dictionaries* - pre-trained models and pronunciation dictionaries are supplied by the Montreal Forced Aligner and used for forced alignment. Languages for which dictionaries are available include:
  - English
  - French
  - German
  - Brazilian Portuguese
  - Spanish
  - Catalan

As the data we have is in English, we will use the *Pre-trained Models/Dictionaries* approach in this exercise.

In this exercise you will

- install the MFA Layer Manager,
- force-align the speech of all of the participants in your database, and
- check and manually correct the alignments.

> **!** Important
>
> In this exercise, you will set up Praat Integration in your web browser. There is currently no Praat integration support for Microsoft's Edge' browser, so if you normally use 'Edge' on Windows, you may need to swap to another browser for this exercise - e.g. *Google Chrome*, or *Mozilla Firefox*.

Installation

MFA

MFA is a 3rd-party tool (https://montrealcorpustools.github.io/Montreal-Forced-Aligner/) that LaBB-CAT integrates with via a Layer Manager module. MFA is not included as part of LaBB-CAT, and so it

must be installed on the server you have installed LaBB-CAT on before you can integrate LaBB-CAT with it.

If MFA has not been installed already, please follow the following steps, depending on the operatings system of your LaBB-CAT server:

Linux

To install the Montreal Forced Aligner on Linux systems for all users, so that your web server can access it if required:

1. Download Miniconda:
   `wget https://repo.anaconda.com/miniconda/Miniconda3-py38\_4.10.3-Linux-x86\_64.sh`
2. Start the installer:
   `sudo bash Miniconda3-py38\_4.10.3-Linux-x86\_64.sh`
3. When asked the location to install Miniconda, use:
   `/opt/conda`
4. When asked whether the installer should initialize Miniconda, this is unnecessary so you can respond no
5. Change ownership of the conda files):
   `sudo chown -R $USERNAME:$USERNAME /opt/conda`
6. Make conda accessible to all users (so you web server can access MFA):
   `chmod -R go-w /opt/conda`
   `chmod -R go+rX /opt/conda`
7. Install the Montreal Forced Aligner
   `/opt/conda/bin/conda create -n aligner -c conda-forge montreal-forced-aligner`

Windows

To install the Montreal Forced Aligner on Windows systems for all users, so that your web server can access it if required:

1. Download the Miniconda installer:
   https://repo.anaconda.com/miniconda/Miniconda3-latest-Windows-x86_64.exe
2. Start the installer by double-clicking it.
3. When asked, select the "Install for all users" option. This will install conda somewhere like
   C:\ProgramData\Miniconda3
4. When asked, tick the *add to PATH* option.
5. Install the Montreal Forced Aligner by specifying a path to the environment
   `conda create -c conda-forge -p C:\ProgramData\Miniconda3\envs\aligner`

```
montreal-forced-aligner
```

Docker Container

If your LaBB-CAT server is installed in a Docker Container, it can download and install Miniconda and MFA itself, as part of the process of installing the MFA Manager (below)

MFA Manager

Once MFA has been installed, you have to install the MFA Manager, which is the LaBB-CAT module that provides MFA with all the data it needs, and then saves to alignments MFA produces back to your database.

(1) Select the *layer managers* menu option.
(2) Follow the *List of layer managers that are not yet installed* link at the bottom.
(3) Find MFA Manager in the list, and press its *Install* button and then press *Install* again.

    As long as MFA has been installed for all users, you should see a box that's already filled in with the location that MFA was installed to in the *Path to MFA* box.

    If the *Path to MFA* box is empty and there's an *Attempt to Install MFA* button, click the button, and LaBB-CAT will try to install MFA on the server for you. This process can take a few minutes, and the *Configure* button will be disabled until it's finished.
(4) Press *Configure* to continue the layer manager installation.
    You will see a window open with some information about integrating with MFA, including the information you've already read above.
(5) Now you need to add a *phrase layer* for the MFA configuration:

    - **Layer ID:** `mfa`
    - **Type:** *Text*
    - **Alignment:** *Intervals*
    - **Manager:** *MFA Manager*
    - **Generate:** *always*
    - **Description:** `MFA alignment time`

> **Note**
>
> After saving the layer, the MFA configuraion form will appear.
> Initially, there will be a 'spinner' and the form will be disabled while LaBB-CAT is making internet requests in order to retrieve lists of available dictionaries and acoustic models.
> This process may take a few seconds, depending on LaBB-CAT's connection to the internet.

(6) When you configure the layer, set the following options:

- **Dictionary Name:** *english_mfa*
- **Pretrained Acoustic Models:** *english_mfa*
- The rest of the options can be left as their default values.

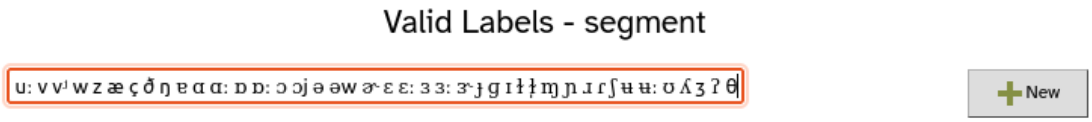
> **Tip**
>
> If you're curious about what the configuration options do, hover your mouse over each option to see a 'tool tip' that describes what the option is for.

(7) Press *Set Parameters* We will not click *Regenerate* to force-align the whole corpus just yet. We need to tweak a few other settings, and then align a single participant's speech.

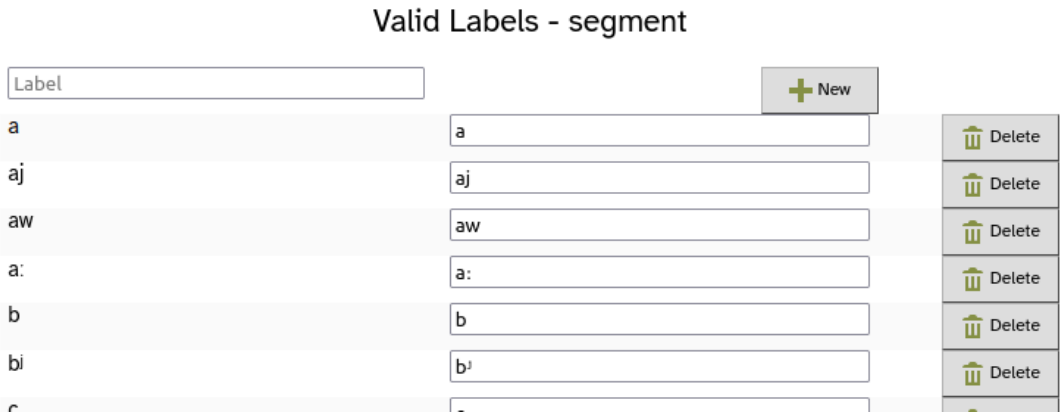
When forced-alignment is done, the resulting aligned phones will be saved on the *segment* layer. By default, this has it's type set to *Phonological*, which assumes that the phoneme symbols will be those defined by the CELEX DISC encoding. However, MFA uses its own phoneme symbols, which are different from the CELEX DISC ones.

To make later processing easier, we're going to change the type of the *segment* layer to *Text*, and let LaBB-CAT know what the possible phoneme labels are.

(8) Select the *segment layers* menu option.
You will see a single layer listed, called *segment*.
(9) Change the *Type* of the *segment* layer to *Text*.
(10) Press *Save*.
A new icon appears with a tag icon - if you hover the mouse over this button, you'll see it's for setting *Valid labels*.
(11) Click the *Valid labels* icon
You will see a page that allows you to add a list of possible labels.
But we don't know what the valid labels are yet. The labels used by MFA depend on the dictionary/models selected when defining the layer. The details for all dictionaries are in MFA Models documentation.
(12) Open the MFA Models documentation for *english_mfa*.
You will see a page that includes information about the dictionary including the *Phones* used by it.
(13) Select and copy the list of phones used by the *english_mfa* dictionary - i.e. all of these:
a aj aw a□ b b□ c c□ d d□ d□ e ej f f□ h i i□ j k k□ l m m□ m̩ n n̩ o ow p p□ p□ s t t□ t□ t□ u u□ v v□ w z æ ç ð ŋ □ □ □□ □ □□ □ □j □ □w □ □ □□ □ □□ □ □ □



(14) Press *New*.
You will see that all of the labels are separately added to the list of valid labels.



Defining this list for the *segment* layer means that, when you search the *segment* layer for specific phones, LaBB-CAT can display a clickable list of possibilities.

(15) At the bottom of the list, there's a *Save* button. Click it to save your changes.

Alignment

(16) Select the *participants* option on the menu.
(17) Find the participant *UC207YW* and tick their checkbox.

> **Note**
>
> Although we're only going to force-align the utterances of a single speaker in this exercise, you can align the utterances of multiple speakers at once, by ticking all their checkboxes on the *participants* page before continuing with the next step…

(18) Press the *All Utterances* button above the list of participants
(19) Press *List*.
You will see a progress bar while all their utterances are identified. Then a results page will be displayed, listing the first 20 utterances.
(20) Click the *Mfa* button at the bottom.
You will see a progress bar appear, while LaBB-CAT gathers the files that MFA needs, runs MFA, and parses the resulting alignments. This will take a few minutes.

Inspection/Correction

Once forced alignment is complete, you can inspect/correct alignments using LaBB-CAT's integration with Praat.

(21) Go to the *transcripts* page and open the *UC207YW.eaf* transcript.
(22) Tick both the *mfa* layer and the *segment* layer.
You will see which lines have been force-aligned, as they have an MFA timestamp, and have the *segment* layer filled in.

The interactive transcript page doesn't show you the alignments of the words or phones, but you can see those using Praat. You can open individual utterances in Praat directly from the transcript page, but first, the LaBB-CAT/Praat integration has to be set up; this only has to be done once:

i. On the top-right of the page, above the playback controls, there's a Praat icon - click it.
ii. Follow the instructions that appear (these vary depending on what web browser you use).
You may be asked whether to allow the "LaBB-CAT Integration Applet" to run. If you tick the "Do not show this again" option, then this message will not appear every time you open a transcript.
You may need to grant a browser extension permission to install, and it's possible you will need a connection to the internet in order to download this extension.
You also may be asked where Praat is installed; Navigate to the location where Praat is installed, and double-click the "Praat.exe" file (on some systems the file may simply be called "Praat"). The Praat program may open, and then immediately close, as LaBB-CAT tests it can communicate with Praat.

Now Praat integration has been set up, and you should be able to access Praat options in the transcript page from now on…

(23) Click on a line that has been aligned, and select the *Open Text Grid in Praat* option on the menu.
You may be asked you if want to allow access to the "LaBB-CAT Integration Applet" - if so, tick "Do not show this again", and click *Allow*.
Praat should open, and show you a spectrogram of the line's audio, with a TextGrid below that includes the words and the segments.
(24) If you click on a word, and hit the `tab` key, the word's interval is played. Try out various words, and see what you think about how accurate HTK has been with its alignment.
Try this out with different lines in the transcript.
You will see that in some cases the alignment is pretty good, and in other cases, it's not so good. In the not-so-good cases, see if you can figure out why HTK got it wrong.

You may have noticed that, each time you open an utterance in Praat, a button appears in the transcript to the left of the line, labelled *Import Changes*. This button allows you to save any adjustments you might want to make to the alignments back into the LaBB-CAT database.

27. If you feel confident using Praat, open an utterance TextGrid, adjust the alignments of the words an phones so that they're more accurate, and then click the *Import Changes* button in the transcript.

> ⚠ Warning
>
> These changes are flagged as manual edits, so if forced-alignment is run again, they will not be over-written with new bad alignments. Therefore it's important that the changes you make are actually improvements, because HTK will never change them again.

> ❗ Important
>
> There are some rules about what you can change: 1. You're not allowed to add or delete words (if this is necessary, it should be done by correcting the transcript instead). 2. All the phones must be within the bounds of their own word. 3. The start of the first phone should line up with the start of the word, and the end of the last phone should line up with the end of the word. 4. You should not change the alignment of the utterance itself (which would only be possible if you select the *Open Text Grid incl. ± 1 utterance in Praat* option).

In this exercise, you have seen how MFA can be used to compute word and phone alignments automatically from your data, and when using a pronunciation dictionary and pre-trained acoustic models, the process is very straightforward.

Such dictionaries/models are only available for a limited number of languages, but if you have a pronunciation dictionary for the language your corpus uses, MFA can also be used to train its own acoustic models from your corpus, and then use them for forced alignment. This process involves a fair amount of careful transcription, tagging, and dictionary filling.

Perfect automatic alignments are not guaranteed, but LaBB-CAT has a mechanism for manually correcting poor alignments.