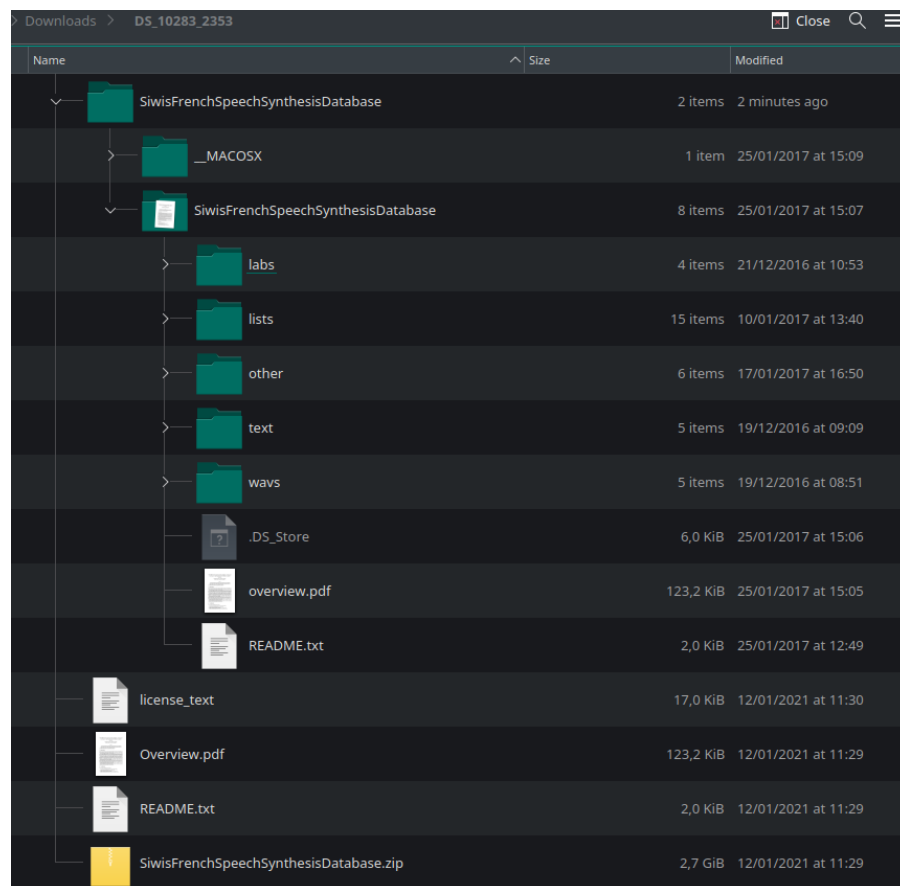


Use Case 1a: aligning French speech data

In this notebook we will do a simple use case by aligning French speech data. We will use the SIWIS French Speech Synthesis Database (Yamagishi et al., 2016) that contains utterances from political debates, book (chapters) and more spoken by one speaker.

1) Download resources

First, download the data from the link: <https://datashare.ed.ac.uk/handle/10283/2353>. This zip contains another zip (SiwisFrenchSpeechSynthesisDatabase.zip) with the data. You should see the following structure:



We need the wav files from dataset. Download also the UseCase1_transcriptions.zip.

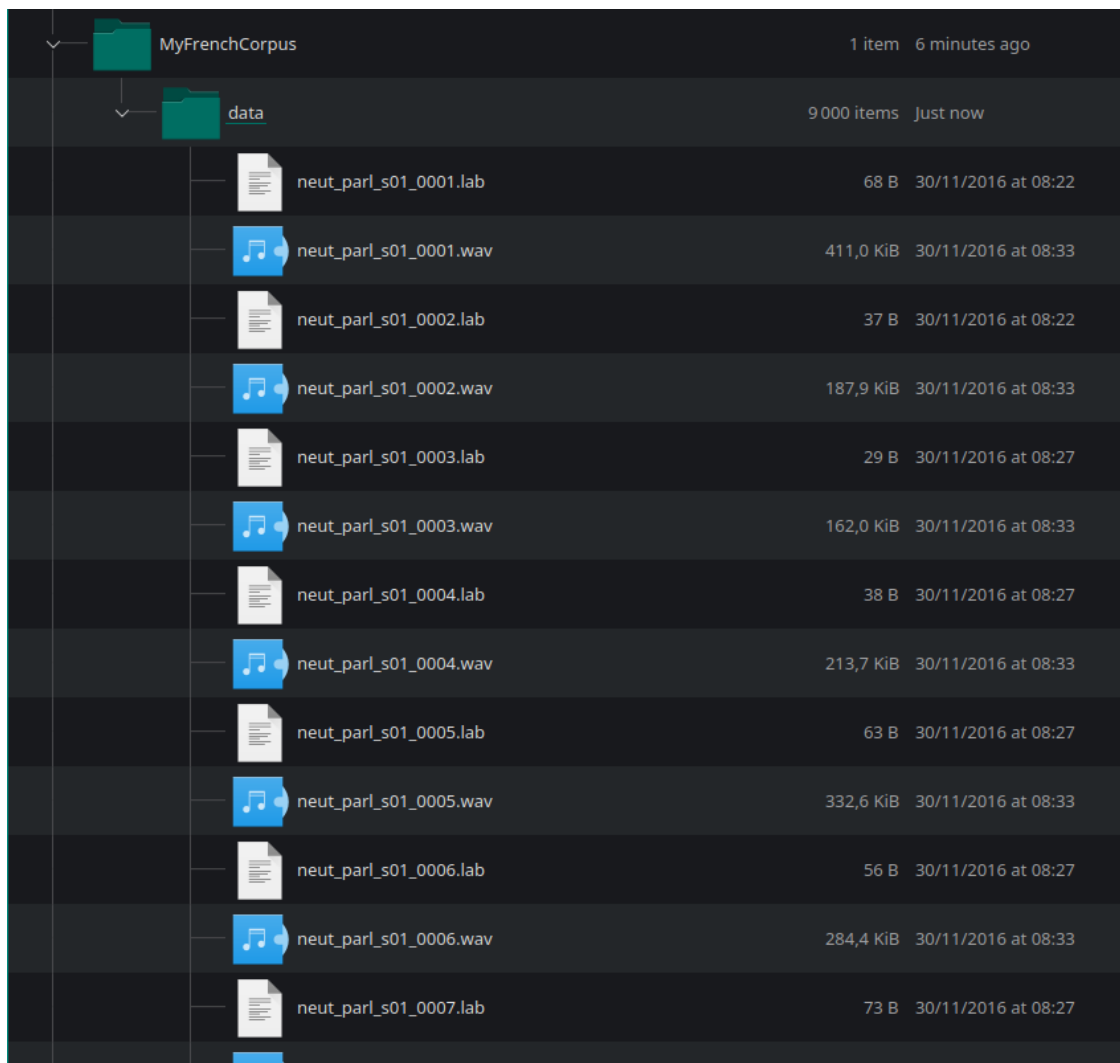
Note: These are not the .lab files from the SIWIS dataset!

2) Create a corpus

The next step is the basis for the alignments. Just create a folder that will contain the data of your corpus. In this case, I created a folder in the documents directory called “MyFrenchCorpus” and another folder that will contain the data.



You can create your corpus folder anywhere you want on your computer, but be sure to note path. Now we can copy & paste the wav files we want to align. In this case, I used the subset “part1” of the SIWIS corpus, but you can use any or all of the subsets. Copy & paste also the respective lab files from the UseCase1_transcriptions.zip. The folder structure should look like this:



In general, the MFA needs both wav files and annotation files for the automatic alignment. The annotation files contain the utterances of the wav files either in the TextGrid format or as .lab files. The .lab files are simple text files with .lab suffix instead of .txt. Regardless of the format, the annotation files contain only the orthographic transcriptions of the utterances. Furthermore, wav files and corresponding annotation files need to have the same file names.

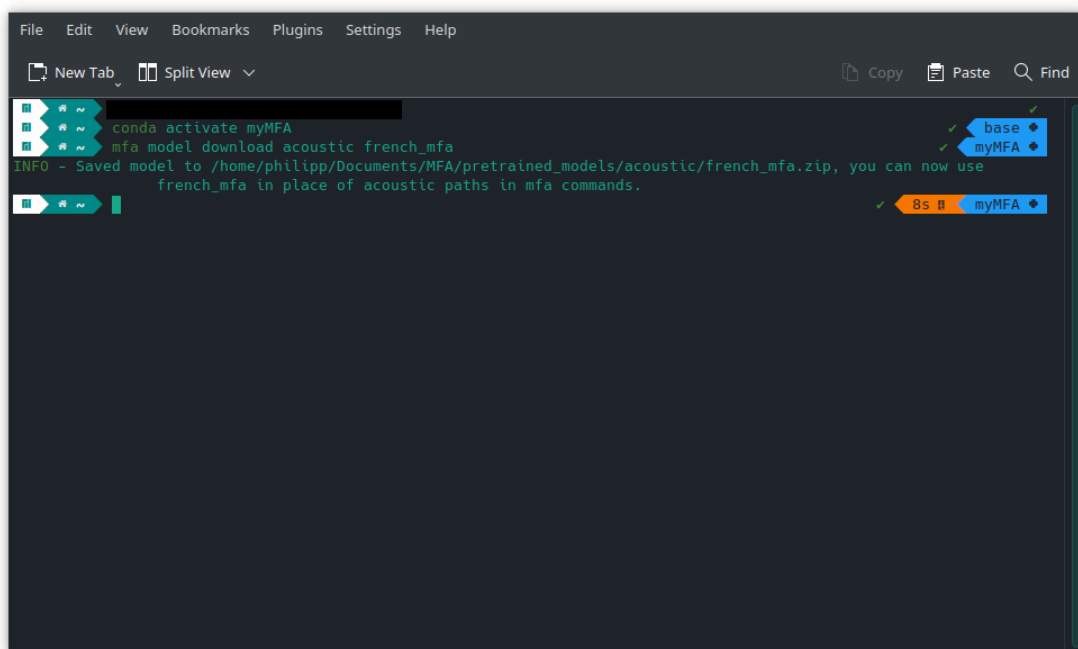
3) open Anaconda and activate your MFA environment

In case Anaconda is not opened, open your terminal/console/command line and activate the environment in which you installed the MFA by (replace “myMFA” by the name you use for the environment):

```
[ ]: conda activate myMFA
```

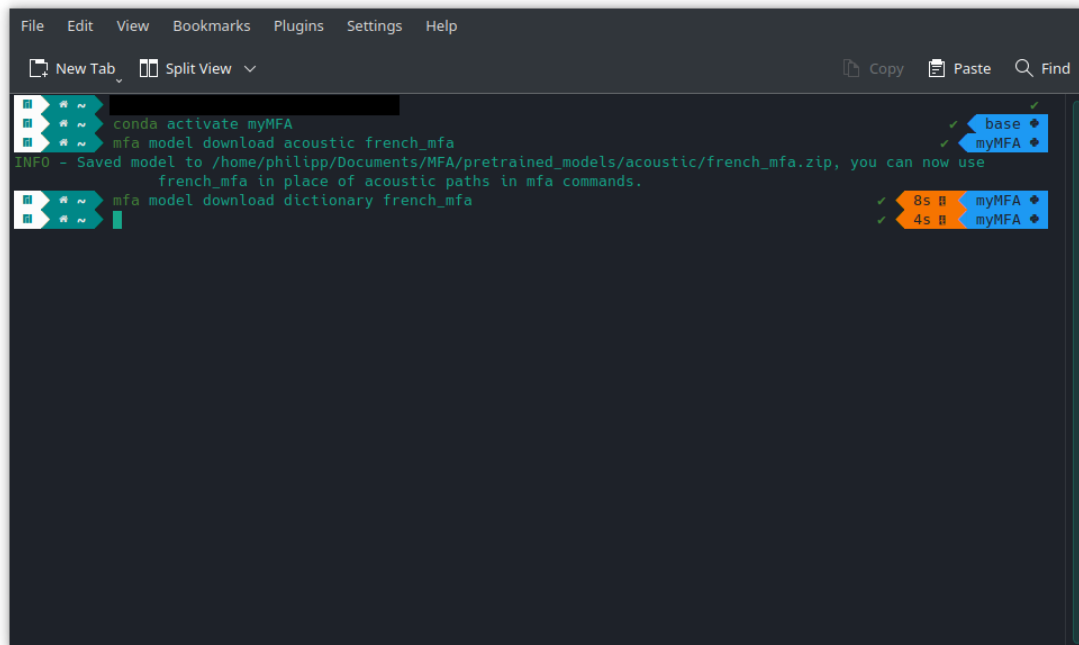
Before we can align the french data, we have to install an acoustic model and dictionary. Pre-trained acoustic models and dictionary can be found here: <https://mfa-models.readthedocs.io/en/latest/> . In this case, we install the “french_mfa” acoustic model:

```
[ ]: mfa model download acoustic french_mfa
```



In the next step, we download the corresponding french dictionary:

```
[ ]: mfa model download dictionary french_mfa
```

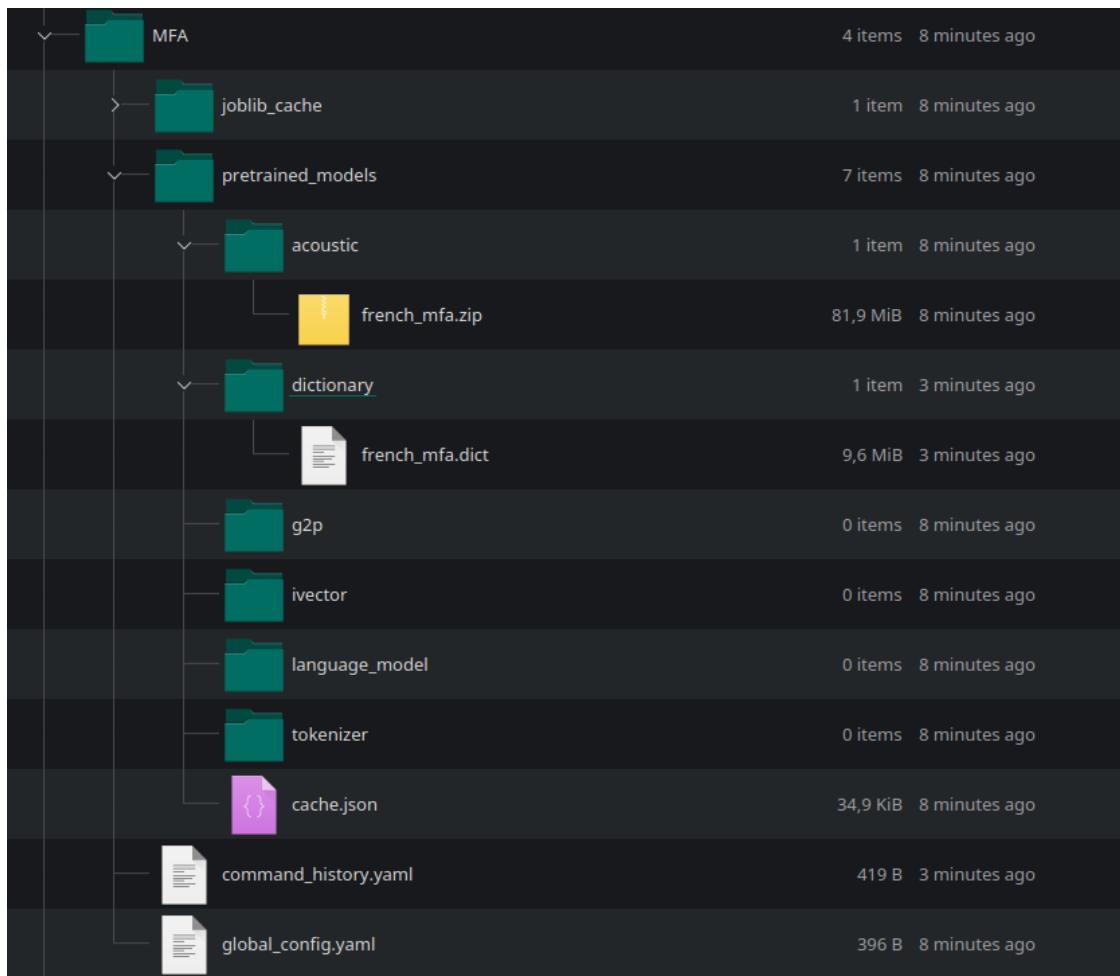


The screenshot shows a terminal window with a dark background. At the top, there is a menu bar with 'File', 'Edit', 'View', 'Bookmarks', 'Plugins', 'Settings', and 'Help'. Below the menu bar, there are tabs for 'New Tab' and 'Split View'. The terminal content shows a series of commands and their outputs:

```
conda activate myMFA
mfa model download acoustic french_mfa
INFO - Saved model to /home/philipp/Documents/MFA/pretrained_models/acoustic/french_mfa.zip, you can now use
      french_mfa in place of acoustic paths in mfa commands.
mfa model download dictionary french_mfa
```

On the right side of the terminal, there are environment indicators: 'base' and 'myMFA' with a green checkmark, and '8s' and '4s' with a green checkmark and a blue arrow pointing to 'myMFA'.

After you first downloaded an acoustic model, dictionary, g2p model etc., you will see that a folder “MFA” was created in your documents directory. This folder contains all models you downloaded:



If you want to delete a specific acoustic model, dictionary, etc., simply delete the respective file. If you want to delete all models, just delete the entire “MFA” folder.

4) Generate alignments for the French data

Now that we have a corpus, an acoustic model and a dictionary, we can create the alignments for the data. We have to type in the following command to the command line:

```
[ ]: mfa align --clean /path/to/your/data /path/to/your/dictionary /path/to/your/  
↪acoustic_model /path/to/your/output/
```

Since this will be the first alignment, the `--clean` flag can be omitted. You can use the `--clean` flag if you run the aligner multiple times, since the files from previous runs will be removed. On Linux and Mac, you can abbreviate your home directory with `~/`. Note: use the full path on Windows: `C:/Users/...`

Since we downloaded the `french_mfa` acoustic model and dictionary, we do not have to enter the full paths, we can just enter `french_mfa`:

```

File Edit View Bookmarks Plugins Settings Help
New Tab Split View Copy Paste Find
B ~$ conda activate myMFA
B ~$ mfa model download acoustic french_mfa
INFO - Saved model to /home/philipp/Documents/MFA/pretrained_models/acoustic/french_mfa.zip, you can now use french_mfa in place of acoustic paths in mfa commands.
B ~$ mfa model download dictionary french_mfa
INFO - Saved dictionary to /home/philipp/Documents/MFA/pretrained_models/dictionary/french_mfa.zip, you can now use french_mfa in place of dictionary paths in mfa commands.
B ~$ mfa align --clean ~/Documents/MyFrenchCorpus/data/ french_mfa french_mfa ~/Documents/MyFrenchCorpus/alignments/

```

This may take some minutes, because we have a lot of data. You should see this after a successful run:

```

File Edit View Bookmarks Plugins Settings Help
New Tab Split View Copy Paste Find
B ~$ conda activate myMFA
B ~$ mfa model download acoustic french_mfa
INFO - Saved model to /home/philipp/Documents/MFA/pretrained_models/acoustic/french_mfa.zip, you can now use french_mfa in place of acoustic paths in mfa commands.
B ~$ mfa model download dictionary french_mfa
INFO - Saved dictionary to /home/philipp/Documents/MFA/pretrained_models/dictionary/french_mfa.zip, you can now use french_mfa in place of dictionary paths in mfa commands.
B ~$ mfa align --clean ~/Documents/MyFrenchCorpus/data/ french_mfa french_mfa ~/Documents/MyFrenchCorpus/alignments/
INFO - Setting up corpus information...
INFO - Loading corpus from source files...
4500it [00:02, 1762.28it/s]
INFO - Found 1 speaker across 4500 files, average number of utterances per speaker: 4500.0
INFO - Initializing multiprocessing jobs...
WARNING - Number of jobs was specified as 3, but due to only having 1 speakers, MFA will only use 1 jobs. Use the --single-speaker flag if you would like to split utterances across jobs regardless of their speaker.
INFO - Normalizing text...
100% | 4500/4500 [00:07:00:00, 560.00it/s]
INFO - Creating corpus split for feature generation...
100% | 9000/9000 [00:01:00:00, 7393.79it/s]
INFO - Generating MFCCs...
4500it [00:54, 82.66it/s]
INFO - Calculating CMVN...
INFO - Generating final features...
100% | 4500/4500 [00:03:00:00, 1402.28it/s]
INFO - Creating corpus split with features...
100% | 4500/4500 [00:02:00:00, 2123.05it/s]
INFO - Compiling training graphs...
100% | 4500/4500 [00:00:00:00, 552.46it/s]
INFO - Performing first-pass alignment...
INFO - Generating alignments...
100% | 4490/4500 [00:24:00:00, 101.89it/s]
INFO - Calculating FHLR for speaker adaptation...
100% | 1/1 [00:12:00:00, 12.78it/s]
INFO - Performing second-pass alignment...
INFO - Generating alignments...
100% | 4499/4500 [00:22:00:00, 190.21it/s]
INFO - Collecting phone and word alignments from alignment lattices...
100% | 0/4500 [00:00:00:00, 7it/s]
INFO - Exporting alignment TextGrids to /home/philipp/Documents/MyFrenchCorpus/alignments...
0% | 0/4500 [00:00:00:00, 7it/s]
Exception during reset or similar
Traceback (most recent call last):
  File "/home/philipp/anaconda3/envs/myMFA/lib/python3.10/site-packages/sqlalchemy/pool/base.py", line 991, in _finalize_fairy
    fairy._reset()
  File "/home/philipp/anaconda3/envs/myMFA/lib/python3.10/site-packages/sqlalchemy/pool/base.py", line 1440, in _reset
    pool._dialect.do_rollback(self)
  File "/home/philipp/anaconda3/envs/myMFA/lib/python3.10/site-packages/sqlalchemy/engine/default.py", line 657, in do_rollback
    @api.connection.rollback()
psycopg2.OperationalError: server closed the connection unexpectedly
This probably means the server terminated abnormally
before or while processing the request.
100% | 4500/4500 [00:26:00:00, 171.29it/s]
INFO - Finished exporting TextGrids to /home/philipp/Documents/MyFrenchCorpus/alignments/
INFO - Done! Everything took 230.328 seconds

```

And here is an example for Windows:

```
(base) C:\Users\p>conda activate myMFA

(myMFA) C:\Users\p>mfa model download acoustic french_mfa

(myMFA) C:\Users\p>mfa model download dictionary french_mfa

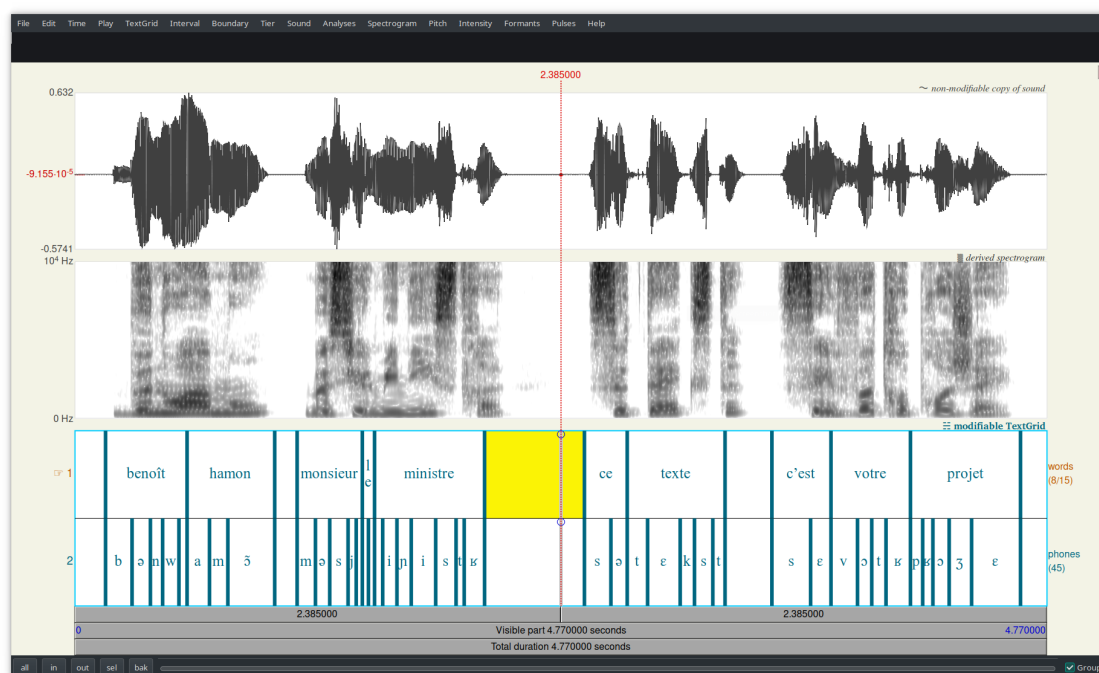
(myMFA) C:\Users\p>mfa align --clean C:/Users/p/Documents/MyFrenchCorpus/data/ french_mfa french_mfa C:/Users/p/Documents/MyFrenchCorpus/alignments/
INFO - Setting up corpus information...
INFO - Loading corpus from source files...
4500it [00:04, 1085.64it/s]
INFO - Found 1 speaker across 4500 files, average number of utterances per speaker: 4500.0
INFO - Initializing multiprocessing jobs...
WARNING - Number of jobs was specified as 3, but due to only having 1 speakers, MFA will only use 1 jobs. Use the
--single_speaker flag if you would like to split utterances across jobs regardless of their speaker.
INFO - Normalizing text...
100% | 4500/4500 [00:11<00:00, 408.98it/s]
INFO - Creating corpus split for feature generation...
100% | 9000/9000 [00:03<00:00, 2848.20it/s]
INFO - Generating MFCCs...
4504it [03:33, 21.13it/s]
INFO - Calculating CMVN...
INFO - Generating final features...
100% | 4500/4500 [00:05<00:00, 824.41it/s]
INFO - Creating corpus split with features...
100% | 4500/4500 [00:04<00:00, 1106.43it/s]
INFO - Compiling training graphs...
100% | 4500/4500 [00:12<00:00, 349.20it/s]
INFO - Performing first-pass alignment...
INFO - Generating alignments...
100% | 4498/4500 [00:33<00:00, 134.65it/s]
INFO - Calculating fMLLR for speaker adaptation...
100% | 1/1 [00:14<00:00, 14.72s/it]
INFO - Performing second-pass alignment...
INFO - Generating alignments...
100% | 4499/4500 [00:28<00:00, 156.82it/s]
0% | 0/4500 [00:00<, ?it/s]
INFO - Collecting phone and word alignments from alignment lattices...
100% | 4499/4500 [00:11<00:00, 408.68it/s]
INFO - Exporting alignment TextGrids to C:\Users\p\Documents\MyFrenchCorpus\alignments...
100% | 4500/4500 [00:30<00:00, 146.77it/s]
INFO - Finished exporting TextGrids to C:\Users\p\Documents\MyFrenchCorpus\alignments!
INFO - Done! Everything took 433.137 seconds

(myMFA) C:\Users\p>
```

You should find a folder called “alignments” in your corpus directory containing the TextGrids:

MyFrenchCorpus	2 items	6 minutes ago
alignments	4 499 items	5 minutes ago
neut_parl_s01_0001.TextGrid	6,1 KiB	5 minutes ago
neut_parl_s01_0002.TextGrid	3,2 KiB	5 minutes ago
neut_parl_s01_0003.TextGrid	2,5 KiB	5 minutes ago
neut_parl_s01_0004.TextGrid	3,8 KiB	5 minutes ago
neut_parl_s01_0005.TextGrid	5,8 KiB	5 minutes ago
neut_parl_s01_0006.TextGrid	4,9 KiB	5 minutes ago
neut_parl_s01_0007.TextGrid	6,2 KiB	5 minutes ago
neut_parl_s01_0008.TextGrid	6,6 KiB	5 minutes ago
neut_parl_s01_0009.TextGrid	3,4 KiB	5 minutes ago
neut_parl_s01_0010.TextGrid	2,5 KiB	5 minutes ago
neut_parl_s01_0011.TextGrid	7,7 KiB	5 minutes ago

When you open them in Praat, you will see that word and phone alignments were generated:



References

Yamagishi, Junichi; Honnet, Pierre-Edouard; Garner, Philip; Lazaridis, Alexandros. (2017). The SIWIS French Speech Synthesis Database, 2016 [dataset]. University of Edinburgh. School of Informatics. The Centre for Speech Technology Research. <https://doi.org/10.7488/ds/1705>.