

OS: MacOS

Disclaimer: I used Google and ChatGPT when I got stuck.

I used the following commands. Many times, it failed, but below all worked.

# From any shell

```
conda deactivate 2>/dev/null || true
```

# Create an env that *forces* arm builds for this one command

```
CONDA_SUBDIR=osx-arm64 conda create -n mfa-arm -c conda-forge -y \
  python=3.10 montreal-forced-aligner=3.3.8 llvm-openmp
```

```
conda activate mfa-arm
```

```
which python
```

```
python -V
```

# make sure you install into THIS python

```
python -m pip install --upgrade pip
```

```
python -m pip install textgrid
```

# Persist the ARM64 setting inside this env so future installs stay arm64

```
conda config --env --set subdir osx-arm64
```

```
alias mfa="python -m montreal_forced_aligner"
```

```
mfa version
```

```
mfa model download acoustic english_us_arpa
```

```
mfa model download dictionary english_us_arpa
```

```
mfa validate corpus
```

```
~/Documents/MFA/pretrained_models/dictionary/english_mfa.
dict
```

**Usage:** `python -m montreal_forced_aligner align [OPTIONS]  
CORPUS_DIRECTORY DICTIONARY_PATH ACOUSTIC_MODEL_PATH  
OUTPUT_DIRECTORY`

```
mfa align ~/Downloads/Assignment/corpus
~/Documents/MFA/pretrained_models/dictionary/english_mfa.
dict
~/Documents/MFA/pretrained_models/acoustic/english_mfa
~/Downloads/Assignment/aligned
```

This created following files as below:

```
cd aligned/speaker1 [from inside Assignment folder do cd]
```

```
ls
```

```
F2BJRLP1.TextGrid
```

```
ISLE_SESS0131_BLOCKD02_01_sprt1.TextGrid
```

```
F2BJRLP2.TextGrid
```

```
ISLE_SESS0131_BLOCKD02_02_sprt1.TextGrid
```

```
F2BJRLP3.TextGrid
```

```
ISLE_SESS0131_BLOCKD02_03_sprt1.TextGrid
```

Installed praat to see the textgrid file but later realised that i can see whats inside these files using a text editor

Consider F2BJRLP2.TextGrid (GENERATED SNAPSHOT CSV FILE FOR THIS AS BELOW)

## F2BJRLP2.words

start_s	end_s	duration_s	label
<b>0.000000</b>	0.170000	0.170000	in
<b>0.170000</b>	0.540000	0.370000	nineteen
<b>0.540000</b>	0.910000	0.370000	seventy
<b>0.910000</b>	1.380000	0.470000	six
<b>1.380000</b>	2.050000	0.670000	democratic
<b>2.050000</b>	2.410000	0.360000	governor
<b>2.410000</b>	2.750000	0.340000	michael
<b>2.750000</b>	3.350000	0.600000	<unk>
<b>3.630000</b>	4.060000	0.430000	fulfilled
<b>4.060000</b>	4.120000	0.060000	a
<b>4.120000</b>	4.670000	0.550000	campaign
<b>4.670000</b>	5.220000	0.550000	promise
<b>5.290000</b>	5.410000	0.120000	to
<b>5.410000</b>	5.570000	0.160000	de
<b>5.570000</b>	6.520000	0.950000	politicize
<b>6.520000</b>	7.050000	0.530000	judicial
<b>7.050000</b>	7.760000	0.710000	appointments
<b>7.790000</b>	8.160000	0.370000	he
<b>8.160000</b>	8.500000	0.340000	named
<b>8.500000</b>	9.180000	0.680000	republican

.....

<All words as intervals[word] for every 'word' in the transcript  
Each word (interval[i]) has start time, end time and text(word).  
Interestingly some words are missing. For example in the above file,  
the word 'dukakis' is present in transcript but not present in the above  
corresponding TextGrid file (SEE THE CSV FILE WITH START TIME 2.75  
TO 3.35 -> 9TH LINE WITH <UNK>).

## THE CORRESPONDING TRANSCRIPT FILE HAS THE FOLLOWING

*In nineteen seventy- six, Democratic Governor Michael Dukakis*

*fulfilled a campaign promise to de-politicize judicial appointments.*

*He named Republican Edward Hennessy to head the State Supreme*

*Judicial Court. For Hennessy, it was another step along a*

*distinguished career that began as a trial lawyer*

*and led to an appointment as associate Supreme Court*

*Justice in nineteen seventy- one. That year Thomas Maffy,*

*now president of the Massachusetts Bar Association, was Hennessy's*

*law clerk.*

See the below:

**NOTE: [export\\_textgrid\\_csv3.py](#) is in /aligned/speaker1**

**python export\_textgrid\_csv3.py ~/Downloads/Assignment/aligned**  
**OUTPUT of the above:**

-> wrote speaker1/F2BJRLP2.words.csv

=> wrote speaker1/F2BJRLP2.phones.csv

**F2BJRLP2: words=75, null=11, transcript\_tokens=75,**  
**missing\_in\_alignment=1 -> ['dukakis'] → In this transcript file,**

**there are 75 words, but in textgrid file, 1 word is missing(i dont know why the acoustic model is missing few wordS. I did not explore much. NULL includes the gaps between the end time and the next start time(may be the speaker pauses). For ex., see the line numbers 11, 14,19,24,28,31,41,47,49,61,72 in F2BJRLPU2.words.csv file and check with previous line numbers' end times. It does not continue with that end time of previous word. I think these are due to speaker pauses in between).**

-> wrote speaker1/F2BJRLP3.words.csv

=> wrote speaker1/F2BJRLP3.phones.csv

F2BJRLP3: words=84, null=15, transcript\_tokens=84,  
missing\_in\_alignment=1 -> ['judgeships']

-> wrote speaker1/ISLE\_SESS0131\_BLOCKD02\_01\_sprt1.words.csv

=> wrote speaker1/ISLE\_SESS0131\_BLOCKD02\_01\_sprt1.phones.csv

ISLE\_SESS0131\_BLOCKD02\_01\_sprt1: words=5, null=2,  
transcript\_tokens=5, missing\_in\_alignment=0

-> wrote speaker1/ISLE\_SESS0131\_BLOCKD02\_03\_sprt1.words.csv

=> wrote speaker1/ISLE\_SESS0131\_BLOCKD02\_03\_sprt1.phones.csv

ISLE\_SESS0131\_BLOCKD02\_03\_sprt1: words=5, null=2,  
transcript\_tokens=5, missing\_in\_alignment=0

-> wrote speaker1/ISLE\_SESS0131\_BLOCKD02\_02\_sprt1.words.csv

=> wrote speaker1/ISLE\_SESS0131\_BLOCKD02\_02\_sprt1.phones.csv

ISLE\_SESS0131\_BLOCKD02\_02\_sprt1: words=5, null=2,  
transcript\_tokens=5, missing\_in\_alignment=0

-> wrote speaker1/F2BJRLP1.words.csv

=> wrote speaker1/F2BJRLP1.phones.csv

F2BJRLP1: words=72, null=12, transcript\_tokens=72,  
missing\_in\_alignment=0