Use case 3: Xhosa data

In all the preceding tutorials we covered languages that do have pre-trained models and dictionaries in the MFA repository. In this notebook we will work on Xhosa data, a language that has neither a MFA dictionary nor an acoustic model. Fortunately, it is sometimes possible to use resources from other projects. An example is the LWAZI project that aimed to develop speech technology resources for several South African languages. We can use the dictionary that was created in this project for our needs (for the original files, see here.

Download the Xhosa_dataset.zip and the Xhosa.txt and Xhosa_wordlist.txt files, that can be found in the Google Drive. The Xhosa dataset contains audio recordings and transcriptions, the Xhosa.txt is a dictionary and Xhosa_wordlist.txt is a list of all words in the transcription.

Unzip the dataset first. We will run at first the validate command of the MFA:

```
In [10]:  ! mfa validate ~/Workspace/corpora/Xhosa_dataset/ ~/Workspace/corpora/X
```

```
WARNING - The previous run had a different configuration than the curre
nt, which
                  may cause issues. Please see the log for details or
use
                  --clean flag if issues are encountered.
INFO - Setting up corpus information...
INFO - Loading corpus from source files...
2420it [00:01, 1520.57it/s]
INFO - Found 12 speakers across 2420 files, average number of utterance
s per
                  speaker: 201.66666666666666
INFO - Initializing multiprocessing jobs...
INFO - Normalizing text...
100%|████████████████████████████████| 2420/2420 [00:01<00:00, 129
3.30it/s]
INFO - Creating corpus split for feature generation...
100%|████████████████████████████████| 4840/4840 [00:01<00:00, 426
3.60it/s]
INFO - Generating MFCCs...
2439it [00:18, 130.99it/s]
INFO - Calculating CMVN...
INFO - Generating final features...
100%|████████████████████████████████| 2420/2420 [00:01<00:00, 145
2.04it/s]
INFO - Creating corpus split with features...
100%|████████████████████████████████| 2420/2420 [00:01<00:00, 193
9.66it/s]

INFO - ******
INFO - Corpus
INFO - ******

INFO -    2420 sound files
INFO -    2420 text files
INFO -    12 speakers
INFO -    2420 utterances
INFO -    11179.284 seconds total duration

INFO -    Sound file read errors
INFO -    ======================

INFO -       There were no issues reading sound files.

INFO -    Feature generation
INFO -    ==================

INFO -       There were no utterances missing features.

INFO -    Files without transcriptions
INFO -    ============================

INFO -       There were no sound files missing transcriptions.

INFO -    Transcriptions without sound files
INFO -    ================================

INFO -       There were no transcription files missing sound files.


INFO - **********
```

```
INFO - Dictionary
INFO - **********

INFO -    Out of vocabulary words
INFO -    =======================

INFO -      44 OOV word types
INFO -      94043 total OOV tokens

INFO -     For a full list of the word types, please see:

INFO -          /home/philipp/Documents/MFA/Xhosa_dataset/oovs_found.txt

INFO -     For a by-utterance breakdown of missing words, see:

INFO -          /home/philipp/Documents/MFA/Xhosa_dataset/utterance_oov
s.txt


INFO - ********
INFO - Training
INFO - ********

INFO - Creating subset directory with 2000 utterances...
100%|████████████████████████████████| 2000/2000 [00:01<00:00, 175
1.71it/s]
INFO - Initializing training for monophone...
INFO - Compiling training graphs...
100%|████████████████████████████████| 2000/2000 [00:01<00:00, 130
1.77it/s]
INFO - Generating initial alignments...
100%|████████████████████████████████| 2000/2000 [00:02<00:00, 91
5.42it/s]
INFO - Initialization complete!
INFO - monophone - Iteration 1 of 40
INFO - Generating alignments...
 95%|███████████████████████████████  | 1892/2000 [00:20<00:01, 9
0.39it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2000/2000 [00:01<00:00, 115
3.60it/s]
INFO - monophone - Iteration 2 of 40
INFO - Generating alignments...
100%|████████████████████████████████| 2000/2000 [00:09<00:00, 21
3.70it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2000/2000 [00:01<00:00, 109
2.71it/s]
INFO - monophone - Iteration 3 of 40
INFO - Generating alignments...
100%|████████████████████████████████| 2000/2000 [00:07<00:00, 27
0.56it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2000/2000 [00:01<00:00, 107
1.10it/s]
INFO - monophone - Iteration 4 of 40
INFO - Generating alignments...
100%|████████████████████████████████| 2000/2000 [00:06<00:00, 28
8.96it/s]
```

INFO - Accumulating statistics...
100%|████████████████████████████████| 2000/2000 [00:01<00:00, 108
5.19it/s]
INFO - monophone - Iteration 5 of 40
INFO - Generating alignments...
100%|████████████████████████████████| 2000/2000 [00:06<00:00, 29
0.85it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2000/2000 [00:01<00:00, 108
3.86it/s]
INFO - monophone - Iteration 6 of 40
INFO - Generating alignments...
100%|████████████████████████████████| 2000/2000 [00:06<00:00, 29
7.35it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2000/2000 [00:01<00:00, 106
0.83it/s]
INFO - monophone - Iteration 7 of 40
INFO - Generating alignments...
100%|████████████████████████████████| 2000/2000 [00:06<00:00, 29
1.11it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2000/2000 [00:01<00:00, 106
0.48it/s]
INFO - monophone - Iteration 8 of 40
INFO - Generating alignments...
100%|████████████████████████████████| 2000/2000 [00:06<00:00, 29
2.19it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2000/2000 [00:01<00:00, 107
6.13it/s]
INFO - monophone - Iteration 9 of 40
INFO - Generating alignments...
100%|████████████████████████████████| 2000/2000 [00:06<00:00, 29
9.62it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2000/2000 [00:01<00:00, 105
1.04it/s]
INFO - monophone - Iteration 10 of 40
INFO - Generating alignments...
100%|████████████████████████████████| 2000/2000 [00:06<00:00, 29
8.90it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2000/2000 [00:01<00:00, 102
0.19it/s]
INFO - monophone - Iteration 11 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████| 2000/2000 [00:01<00:00, 103
8.84it/s]
INFO - monophone - Iteration 12 of 40
INFO - Generating alignments...
100%|████████████████████████████████| 2000/2000 [00:06<00:00, 30
8.90it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2000/2000 [00:01<00:00, 102
0.06it/s]
INFO - monophone - Iteration 13 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████| 2000/2000 [00:01<00:00, 102
1.72it/s]

```
INFO - monophone - Iteration 14 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 2000/2000 [00:06<00:00, 30
5.77it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2000/2000 [00:01<00:00, 101
9.34it/s]
INFO - monophone - Iteration 15 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2000/2000 [00:01<00:00, 101
7.80it/s]
INFO - monophone - Iteration 16 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 2000/2000 [00:06<00:00, 30
2.56it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2000/2000 [00:02<00:00, 98
1.34it/s]
INFO - monophone - Iteration 17 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2000/2000 [00:02<00:00, 98
8.30it/s]
INFO - monophone - Iteration 18 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 2000/2000 [00:06<00:00, 29
6.88it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2000/2000 [00:02<00:00, 99
7.45it/s]
INFO - monophone - Iteration 19 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2000/2000 [00:02<00:00, 98
5.38it/s]
INFO - monophone - Iteration 20 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 2000/2000 [00:06<00:00, 29
9.49it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2000/2000 [00:02<00:00, 99
0.41it/s]
INFO - monophone - Iteration 21 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2000/2000 [00:01<00:00, 100
2.82it/s]
INFO - monophone - Iteration 22 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2000/2000 [00:01<00:00, 100
5.73it/s]
INFO - monophone - Iteration 23 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 2000/2000 [00:06<00:00, 29
5.34it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2000/2000 [00:02<00:00, 95
6.39it/s]
INFO - monophone - Iteration 24 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2000/2000 [00:02<00:00, 95
8.15it/s]
INFO - monophone - Iteration 25 of 40
```

```
INFO - Accumulating statistics...
100%|████████████████████████████| 2000/2000 [00:02<00:00, 97
1.81it/s]
INFO - monophone - Iteration 26 of 40
INFO - Generating alignments...
100%|████████████████████████████| 2000/2000 [00:06<00:00, 29
6.84it/s]
INFO - Accumulating statistics...
100%|████████████████████████████| 2000/2000 [00:02<00:00, 93
2.14it/s]
INFO - monophone - Iteration 27 of 40
INFO - Accumulating statistics...
100%|████████████████████████████| 2000/2000 [00:02<00:00, 91
9.58it/s]
INFO - monophone - Iteration 28 of 40
INFO - Accumulating statistics...
100%|████████████████████████████| 2000/2000 [00:02<00:00, 95
6.49it/s]
INFO - monophone - Iteration 29 of 40
INFO - Generating alignments...
100%|████████████████████████████| 2000/2000 [00:06<00:00, 29
0.48it/s]
INFO - Accumulating statistics...
100%|████████████████████████████| 2000/2000 [00:02<00:00, 93
9.38it/s]
INFO - monophone - Iteration 30 of 40
INFO - Accumulating statistics...
100%|████████████████████████████| 2000/2000 [00:02<00:00, 95
6.76it/s]
INFO - monophone - Iteration 31 of 40
INFO - Accumulating statistics...
100%|████████████████████████████| 2000/2000 [00:02<00:00, 92
5.77it/s]
INFO - monophone - Iteration 32 of 40
INFO - Generating alignments...
100%|████████████████████████████| 2000/2000 [00:07<00:00, 27
7.58it/s]
INFO - Accumulating statistics...
100%|████████████████████████████| 2000/2000 [00:02<00:00, 93
9.51it/s]
INFO - monophone - Iteration 33 of 40
INFO - Accumulating statistics...
100%|████████████████████████████| 2000/2000 [00:02<00:00, 92
9.48it/s]
INFO - monophone - Iteration 34 of 40
INFO - Accumulating statistics...
100%|████████████████████████████| 2000/2000 [00:02<00:00, 93
2.78it/s]
INFO - monophone - Iteration 35 of 40
INFO - Generating alignments...
100%|████████████████████████████| 2000/2000 [00:06<00:00, 28
6.70it/s]
INFO - Accumulating statistics...
100%|████████████████████████████| 2000/2000 [00:02<00:00, 93
0.04it/s]
INFO - monophone - Iteration 36 of 40
INFO - Accumulating statistics...
100%|████████████████████████████| 2000/2000 [00:02<00:00, 93
4.92it/s]
INFO - monophone - Iteration 37 of 40
```

```
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2000/2000 [00:02<00:00, 93
5.43it/s]
INFO - monophone - Iteration 38 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 2000/2000 [00:06<00:00, 28
9.34it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2000/2000 [00:02<00:00, 90
2.82it/s]
INFO - monophone - Iteration 39 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2000/2000 [00:02<00:00, 93
3.19it/s]
INFO - monophone - Iteration 40 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2000/2000 [00:02<00:00, 92
8.39it/s]
INFO - Training complete!
INFO - Compiling training graphs...
 83%|███████████████████████████         | 2000/2420 [00:01<00:00, 128
9.88it/s]
INFO - Generating alignments...
 83%|███████████████████████████         | 2000/2420 [00:07<00:01, 28
1.68it/s]
INFO - Accumulating transition stats...
 83%|███████████████████████████         | 2000/2420 [00:01<00:00, 169
7.55it/s]
INFO - Finished accumulating transition stats!
  0%|                                    | 0/2420 [00:00
<?, ?it/s]INFO - Collecting phone and word alignments from monophone_al
i lattices...
 83%|███████████████████████████         | 2000/2420 [00:02<00:00, 82
0.37it/s]
INFO - Beginning phone LM training...
INFO - Collecting training data...
 83%|███████████████████████████         | 2000/2420 [00:01<00:00, 182
9.47it/s]
INFO - Training model...
INFO - Completed training in 294.91183066368103 seconds!
INFO - Done! Everything took 324.102 seconds
```

As we can see under "Dictionary", we have a lot of OOV words, so the dictionary covers not the entire corpus. Therefore we have to add all the entries manually, which will be a lot of work, or we create a new dictionary using a g2p model. A g2p model is also missing, so we have to train one on the Xhosa.txt dictionary.

The training of the g2p model is achieved by the following command (adapt paths):

```
In [5]: ! mfa train_g2p ~/Workspace/corpora/Xhosa.txt ~/Workspace/corpora/Xhosa
```

```
WARNING - The previous run had a different configuration than the curre
nt, which
                    may cause issues. Please see the log for details or
use
                    --clean flag if issues are encountered.
INFO - Training aligner
INFO - Calculating alignments...
100%|████████████████████████████████████| 250/250 [01:45<00:00,
2.37it/s]
INFO - Best likelihood: -11.8614
INFO - Completed computing alignments!
INFO - Encoding the alignments as FSAs
INFO - Success! FAR path: /home/philipp/Documents/MFA/xhosa/train_g2p/x
hosa.far;
                    encoder path:
                    /home/philipp/Documents/MFA/xhosa/train_g2p/xhosa.enc
INFO - Saved model to /home/philipp/Workspace/corpora/xhosa_g2p.zip
INFO - Done! Everything took 109.028 seconds
```

Since we have now our g2p model, we can create the new dictionary based on the wordlist that contains all the words in our transcriptions. We use the following command:

In [7]: `! mfa g2p ~/Workspace/corpora/Xhosa_wordlist.txt ~/Workspace/corpora/Xh`

```
INFO - Generating pronunciations...
100%|████████████████████████████████████| 6754/6754 [00:02<00:00, 322
1.04it/s]
INFO - Done! Everything took 2.519 seconds
```

When we look at the dictionary, we will see that this is not perfect. There are a lot of pronunciation alternations generated by the g2p model and most of them may not be accurate. However, we do this whole procedure to illustrate the workflow. Normally, at this point we would correct the dictionary and remove all unnecessary entries.

But for now we will train now our acoustic model on the corpus data and the dictionary by the following command:

In [8]: `! mfa train ~/Workspace/corpora/Xhosa_dataset/ ~/Workspace/corpora/Xhos`

INFO - Setting up corpus information...
INFO - Loading corpus from source files...
2420it [00:01, 1538.83it/s]
INFO - Found 12 speakers across 2420 files, average number of utterances per
                speaker: 201.66666666666
INFO - Initializing multiprocessing jobs...
INFO - Normalizing text...
100%|████████████████████████████████████| 2420/2420 [00:01<00:00, 1377.79it/s]
INFO - Creating corpus split for feature generation...
100%|████████████████████████████████████| 4840/4840 [00:01<00:00, 4293.13it/s]
INFO - Generating MFCCs...
2435it [00:18, 133.23it/s]
INFO - Calculating CMVN...
INFO - Generating final features...
100%|████████████████████████████████████| 2420/2420 [00:01<00:00, 1439.55it/s]
INFO - Creating corpus split with features...
100%|████████████████████████████████████| 2420/2420 [00:01<00:00, 2145.68it/s]
INFO - Initializing training for monophone...
INFO - Compiling training graphs...
100%|████████████████████████████████████| 2420/2420 [00:04<00:00, 562.54it/s]
INFO - Generating initial alignments...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 1034.04it/s]
INFO - Initialization complete!
INFO - monophone - Iteration 1 of 40
INFO - Generating alignments...
 60%|██████████████████████ | 1445/2420 [00:26<00:18, 53.72it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:01<00:00, 1555.24it/s]
INFO - monophone - Iteration 2 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 2420/2420 [00:10<00:00, 232.81it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 1184.79it/s]
INFO - monophone - Iteration 3 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 2420/2420 [00:07<00:00, 339.08it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 1189.17it/s]
INFO - monophone - Iteration 4 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 2420/2420 [00:05<00:00, 433.94it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 1200.53it/s]
INFO - monophone - Iteration 5 of 40
INFO - Generating alignments...

```
100%|████████████████████████████████| 2420/2420 [00:04<00:00, 49
0.23it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 116
4.01it/s]
INFO - monophone - Iteration 6 of 40
INFO - Generating alignments...
100%|████████████████████████████████| 2420/2420 [00:04<00:00, 50
9.24it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 118
6.80it/s]
INFO - monophone - Iteration 7 of 40
INFO - Generating alignments...
100%|████████████████████████████████| 2420/2420 [00:04<00:00, 52
2.47it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 119
5.38it/s]
INFO - monophone - Iteration 8 of 40
INFO - Generating alignments...
100%|████████████████████████████████| 2420/2420 [00:04<00:00, 53
6.37it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 118
2.68it/s]
INFO - monophone - Iteration 9 of 40
INFO - Generating alignments...
100%|████████████████████████████████| 2420/2420 [00:04<00:00, 53
4.61it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 117
8.15it/s]
INFO - monophone - Iteration 10 of 40
INFO - Generating alignments...
100%|████████████████████████████████| 2420/2420 [00:04<00:00, 54
3.44it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 117
5.28it/s]
INFO - monophone - Iteration 11 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 114
7.47it/s]
INFO - monophone - Iteration 12 of 40
INFO - Generating alignments...
100%|████████████████████████████████| 2420/2420 [00:04<00:00, 52
7.28it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 113
0.32it/s]
INFO - monophone - Iteration 13 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 116
0.11it/s]
INFO - monophone - Iteration 14 of 40
INFO - Generating alignments...
100%|████████████████████████████████| 2420/2420 [00:04<00:00, 52
2.71it/s]
INFO - Accumulating statistics...
```

```
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 116
2.48it/s]
INFO - monophone - Iteration 15 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 111
3.29it/s]
INFO - monophone - Iteration 16 of 40
INFO - Generating alignments...
100%|█████████████████████████████████| 2420/2420 [00:04<00:00, 52
8.59it/s]
INFO - Accumulating statistics...
100%|█████████████████████████████████| 2420/2420 [00:02<00:00, 113
0.82it/s]
INFO - monophone - Iteration 17 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 112
6.34it/s]
INFO - monophone - Iteration 18 of 40
INFO - Generating alignments...
100%|█████████████████████████████████| 2420/2420 [00:04<00:00, 53
2.32it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 111
1.79it/s]
INFO - monophone - Iteration 19 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 103
7.55it/s]
INFO - monophone - Iteration 20 of 40
INFO - Generating alignments...
100%|█████████████████████████████████| 2420/2420 [00:04<00:00, 50
4.99it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 109
9.40it/s]
INFO - monophone - Iteration 21 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 110
1.92it/s]
INFO - monophone - Iteration 22 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 107
5.27it/s]
INFO - monophone - Iteration 23 of 40
INFO - Generating alignments...
100%|█████████████████████████████████| 2420/2420 [00:04<00:00, 50
4.66it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 110
6.30it/s]
INFO - monophone - Iteration 24 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 110
7.23it/s]
INFO - monophone - Iteration 25 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 110
7.05it/s]
INFO - monophone - Iteration 26 of 40
INFO - Generating alignments...
```

```
100%|███████████████████████████████████████| 2420/2420 [00:04<00:00, 50
7.01it/s]
INFO - Accumulating statistics...
100%|███████████████████████████████████████| 2420/2420 [00:02<00:00, 112
5.90it/s]
INFO - monophone - Iteration 27 of 40
INFO - Accumulating statistics...
100%|███████████████████████████████████████| 2420/2420 [00:02<00:00, 110
3.79it/s]
INFO - monophone - Iteration 28 of 40
INFO - Accumulating statistics...
100%|███████████████████████████████████████| 2420/2420 [00:02<00:00, 110
6.28it/s]
INFO - monophone - Iteration 29 of 40
INFO - Generating alignments...
100%|███████████████████████████████████████| 2420/2420 [00:04<00:00, 51
5.68it/s]
INFO - Accumulating statistics...
100%|███████████████████████████████████████| 2420/2420 [00:02<00:00, 112
0.94it/s]
INFO - monophone - Iteration 30 of 40
INFO - Accumulating statistics...
100%|███████████████████████████████████████| 2420/2420 [00:02<00:00, 110
1.15it/s]
INFO - monophone - Iteration 31 of 40
INFO - Accumulating statistics...
100%|███████████████████████████████████████| 2420/2420 [00:02<00:00, 107
9.94it/s]
INFO - monophone - Iteration 32 of 40
INFO - Generating alignments...
100%|███████████████████████████████████████| 2420/2420 [00:04<00:00, 50
6.47it/s]
INFO - Accumulating statistics...
100%|███████████████████████████████████████| 2420/2420 [00:02<00:00, 104
8.88it/s]
INFO - monophone - Iteration 33 of 40
INFO - Accumulating statistics...
100%|███████████████████████████████████████| 2420/2420 [00:02<00:00, 109
7.46it/s]
INFO - monophone - Iteration 34 of 40
INFO - Accumulating statistics...
100%|███████████████████████████████████████| 2420/2420 [00:02<00:00, 108
8.85it/s]
INFO - monophone - Iteration 35 of 40
INFO - Generating alignments...
100%|███████████████████████████████████████| 2420/2420 [00:04<00:00, 51
5.99it/s]
INFO - Accumulating statistics...
100%|███████████████████████████████████████| 2420/2420 [00:02<00:00, 109
8.18it/s]
INFO - monophone - Iteration 36 of 40
INFO - Accumulating statistics...
100%|███████████████████████████████████████| 2420/2420 [00:02<00:00, 110
3.83it/s]
INFO - monophone - Iteration 37 of 40
INFO - Accumulating statistics...
100%|███████████████████████████████████████| 2420/2420 [00:02<00:00, 108
6.81it/s]
INFO - monophone - Iteration 38 of 40
INFO - Generating alignments...
```

```
100%|████████████████████████████████| 2420/2420 [00:04<00:00, 50
9.18it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 102
5.98it/s]
INFO - monophone - Iteration 39 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 102
6.38it/s]
INFO - monophone - Iteration 40 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 104
4.25it/s]
INFO - Training complete!
INFO - Compiling training graphs...
100%|████████████████████████████████| 2420/2420 [00:04<00:00, 51
6.96it/s]
INFO - Generating alignments...
100%|████████████████████████████████| 2420/2420 [00:05<00:00, 47
7.28it/s]
INFO - Initializing training for triphone...
INFO - Compiling training graphs...
100%|████████████████████████████████| 2420/2420 [00:05<00:00, 44
1.70it/s]
INFO - Converting alignments...
100%|████████████████████████████████| 2420/2420 [00:01<00:00, 212
1.67it/s]
INFO - Initialization complete!
INFO - triphone - Iteration 1 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 116
8.20it/s]
INFO - triphone - Iteration 2 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 118
8.15it/s]
INFO - triphone - Iteration 3 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 115
7.14it/s]
INFO - triphone - Iteration 4 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 118
1.11it/s]
INFO - triphone - Iteration 5 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 115
5.86it/s]
INFO - triphone - Iteration 6 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 117
9.67it/s]
INFO - triphone - Iteration 7 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 116
9.31it/s]
INFO - triphone - Iteration 8 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 113
9.89it/s]
```

```
INFO - triphone - Iteration 9 of 35
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 116
4.20it/s]
INFO - triphone - Iteration 10 of 35
INFO - Generating alignments...
100%|██████████████████████████████| 2420/2420 [00:04<00:00, 56
3.90it/s]
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 111
1.35it/s]
INFO - triphone - Iteration 11 of 35
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 114
4.49it/s]
INFO - triphone - Iteration 12 of 35
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 113
1.74it/s]
INFO - triphone - Iteration 13 of 35
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 115
0.31it/s]
INFO - triphone - Iteration 14 of 35
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 112
7.09it/s]
INFO - triphone - Iteration 15 of 35
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 113
8.75it/s]
INFO - triphone - Iteration 16 of 35
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 114
1.66it/s]
INFO - triphone - Iteration 17 of 35
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 112
1.27it/s]
INFO - triphone - Iteration 18 of 35
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 113
8.11it/s]
INFO - triphone - Iteration 19 of 35
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 112
2.31it/s]
INFO - triphone - Iteration 20 of 35
INFO - Generating alignments...
100%|██████████████████████████████| 2420/2420 [00:04<00:00, 54
0.34it/s]
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 111
8.41it/s]
INFO - triphone - Iteration 21 of 35
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 111
7.01it/s]
INFO - triphone - Iteration 22 of 35
INFO - Accumulating statistics...
```

```
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 111
6.35it/s]
INFO - triphone - Iteration 23 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 108
5.49it/s]
INFO - triphone - Iteration 24 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 108
2.82it/s]
INFO - triphone - Iteration 25 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 109
9.99it/s]
INFO - triphone - Iteration 26 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 107
0.25it/s]
INFO - triphone - Iteration 27 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 109
4.51it/s]
INFO - triphone - Iteration 28 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 108
0.99it/s]
INFO - triphone - Iteration 29 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 109
7.40it/s]
INFO - triphone - Iteration 30 of 35
INFO - Generating alignments...
100%|████████████████████████████████████| 2420/2420 [00:04<00:00, 52
4.88it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 108
9.30it/s]
INFO - triphone - Iteration 31 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 107
9.33it/s]
INFO - triphone - Iteration 32 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 108
6.01it/s]
INFO - triphone - Iteration 33 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 110
2.70it/s]
INFO - triphone - Iteration 34 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 104
9.29it/s]
INFO - triphone - Iteration 35 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 107
3.34it/s]
INFO - Training complete!
INFO - Compiling training graphs...
100%|████████████████████████████████████| 2420/2420 [00:05<00:00, 44
```

```
9.56it/s]
INFO - Generating alignments...
100%|████████████████████████████████| 2420/2420 [00:04<00:00, 49
3.03it/s]
INFO - Initializing training for lda...
100%|████████████████████████████████| 2420/2420 [00:01<00:00, 164
6.01it/s]
INFO - Compiling training graphs...
100%|████████████████████████████████| 2420/2420 [00:05<00:00, 44
9.16it/s]
INFO - Converting alignments...
100%|████████████████████████████████| 2420/2420 [00:01<00:00, 211
1.13it/s]
INFO - Initialization complete!
INFO - lda - Iteration 1 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:01<00:00, 123
0.50it/s]
INFO - lda - Iteration 2 of 35
INFO - Re-calculating LDA...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 102
2.38it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:01<00:00, 121
5.68it/s]
INFO - lda - Iteration 3 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:01<00:00, 121
1.65it/s]
INFO - lda - Iteration 4 of 35
INFO - Re-calculating LDA...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 101
0.20it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 119
8.97it/s]
INFO - lda - Iteration 5 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 120
6.35it/s]
INFO - lda - Iteration 6 of 35
INFO - Re-calculating LDA...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 97
7.39it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 118
9.01it/s]
INFO - lda - Iteration 7 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 119
0.19it/s]
INFO - lda - Iteration 8 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 118
4.51it/s]
INFO - lda - Iteration 9 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 117
9.62it/s]
INFO - lda - Iteration 10 of 35
```

```
INFO - Generating alignments...
100%|████████████████████████████████| 2420/2420 [00:04<00:00, 58
8.97it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 116
3.60it/s]
INFO - lda - Iteration 11 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 117
0.87it/s]
INFO - lda - Iteration 12 of 35
INFO - Re-calculating LDA...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 98
9.11it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 115
8.45it/s]
INFO - lda - Iteration 13 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 114
1.04it/s]
INFO - lda - Iteration 14 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 114
6.39it/s]
INFO - lda - Iteration 15 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 114
7.38it/s]
INFO - lda - Iteration 16 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 113
6.87it/s]
INFO - lda - Iteration 17 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 113
6.74it/s]
INFO - lda - Iteration 18 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 112
6.74it/s]
INFO - lda - Iteration 19 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 113
2.32it/s]
INFO - lda - Iteration 20 of 35
INFO - Generating alignments...
100%|████████████████████████████████| 2420/2420 [00:04<00:00, 55
1.01it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 111
6.91it/s]
INFO - lda - Iteration 21 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 111
0.68it/s]
INFO - lda - Iteration 22 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 109
8.70it/s]
```

```
INFO - lda - Iteration 23 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 110
1.86it/s]
INFO - lda - Iteration 24 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 110
7.66it/s]
INFO - lda - Iteration 25 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 108
6.64it/s]
INFO - lda - Iteration 26 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 101
5.98it/s]
INFO - lda - Iteration 27 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 101
8.48it/s]
INFO - lda - Iteration 28 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 102
8.99it/s]
INFO - lda - Iteration 29 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 101
5.38it/s]
INFO - lda - Iteration 30 of 35
INFO - Generating alignments...
100%|████████████████████████████████| 2420/2420 [00:04<00:00, 53
0.16it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 107
2.13it/s]
INFO - lda - Iteration 31 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 105
5.00it/s]
INFO - lda - Iteration 32 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 106
5.47it/s]
INFO - lda - Iteration 33 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 106
7.96it/s]
INFO - lda - Iteration 34 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 109
0.68it/s]
INFO - lda - Iteration 35 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 107
7.41it/s]
INFO - Training complete!
INFO - Compiling training graphs...
100%|████████████████████████████████| 2420/2420 [00:05<00:00, 43
7.83it/s]
INFO - Generating alignments...
```

```
100%|████████████████████████████████████| 2420/2420 [00:04<00:00, 49
9.26it/s]
INFO - Initializing training for sat...
INFO - Calculating fMLLR for speaker adaptation...
100%|████████████████████████████████████| 12/12 [00:05<00:00,
2.32it/s]
INFO - Converting alignments...
100%|████████████████████████████████████| 2420/2420 [00:01<00:00, 213
6.24it/s]
INFO - Compiling training graphs...
100%|████████████████████████████████████| 2420/2420 [00:06<00:00, 37
9.72it/s]
INFO - Initialization complete!
INFO - sat - Iteration 1 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 118
1.34it/s]
INFO - sat - Iteration 2 of 35
INFO - Calculating fMLLR for speaker adaptation...
100%|████████████████████████████████████| 12/12 [00:04<00:00,
2.50it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 117
4.34it/s]
INFO - sat - Iteration 3 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 114
1.11it/s]
INFO - sat - Iteration 4 of 35
INFO - Calculating fMLLR for speaker adaptation...
100%|████████████████████████████████████| 12/12 [00:04<00:00,
2.50it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 111
5.23it/s]
INFO - sat - Iteration 5 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 117
4.44it/s]
INFO - sat - Iteration 6 of 35
INFO - Calculating fMLLR for speaker adaptation...
100%|████████████████████████████████████| 12/12 [00:04<00:00,
2.54it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 112
1.02it/s]
INFO - sat - Iteration 7 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 114
9.14it/s]
INFO - sat - Iteration 8 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 114
8.34it/s]
INFO - sat - Iteration 9 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:02<00:00, 116
7.74it/s]
INFO - sat - Iteration 10 of 35
INFO - Generating alignments...
```

```
100%|████████████████████████████████| 2420/2420 [00:04<00:00, 59
5.67it/s]
WARNING - No files were aligned, this likely indicates serious problems
with the
                    aligner.
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 115
1.03it/s]
INFO - sat - Iteration 11 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 116
1.15it/s]
INFO - sat - Iteration 12 of 35
INFO - Calculating fMLLR for speaker adaptation...
100%|████████████████████████████████| 12/12 [00:04<00:00,
2.57it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 116
1.97it/s]
INFO - sat - Iteration 13 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 114
5.35it/s]
INFO - sat - Iteration 14 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 115
1.86it/s]
INFO - sat - Iteration 15 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 114
7.79it/s]
INFO - sat - Iteration 16 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 114
5.39it/s]
INFO - sat - Iteration 17 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 111
6.20it/s]
INFO - sat - Iteration 18 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 113
6.94it/s]
INFO - sat - Iteration 19 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 113
7.81it/s]
INFO - sat - Iteration 20 of 35
INFO - Generating alignments...
100%|████████████████████████████████| 2420/2420 [00:04<00:00, 57
2.05it/s]
WARNING - No files were aligned, this likely indicates serious problems
with the
                    aligner.
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 109
5.18it/s]
INFO - sat - Iteration 21 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 111
```

```
5.07it/s]
INFO - sat - Iteration 22 of 35
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 107
0.45it/s]
INFO - sat - Iteration 23 of 35
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 110
4.56it/s]
INFO - sat - Iteration 24 of 35
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 110
4.66it/s]
INFO - sat - Iteration 25 of 35
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 110
0.93it/s]
INFO - sat - Iteration 26 of 35
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 108
3.58it/s]
INFO - sat - Iteration 27 of 35
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 107
8.33it/s]
INFO - sat - Iteration 28 of 35
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 109
6.33it/s]
INFO - sat - Iteration 29 of 35
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 109
0.60it/s]
INFO - sat - Iteration 30 of 35
INFO - Generating alignments...
100%|████████████████████████████| 2420/2420 [00:04<00:00, 55
6.75it/s]
WARNING - No files were aligned, this likely indicates serious problems
with the
                      aligner.
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 108
8.32it/s]
INFO - sat - Iteration 31 of 35
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 109
9.22it/s]
INFO - sat - Iteration 32 of 35
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 108
7.98it/s]
INFO - sat - Iteration 33 of 35
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 107
9.44it/s]
INFO - sat - Iteration 34 of 35
INFO - Accumulating statistics...
100%|████████████████████████████| 2420/2420 [00:02<00:00, 108
6.67it/s]
INFO - sat - Iteration 35 of 35
```

```
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 109
1.99it/s]
INFO - Creating alignment model for speaker-independent features...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 92
7.51it/s]
INFO - Training complete!
INFO - Compiling training graphs...
100%|████████████████████████████████| 2420/2420 [00:05<00:00, 43
2.91it/s]
INFO - Generating alignments...
100%|████████████████████████████████| 2420/2420 [00:04<00:00, 53
4.12it/s]
INFO - Calculating fMLLR for speaker adaptation...
100%|██████████████████████████████████| 12/12 [00:04<00:00,
2.51it/s]
INFO - Generating alignments...
100%|████████████████████████████████| 2420/2420 [00:04<00:00, 53
1.98it/s]
INFO - Initializing training for sat_2...
INFO - Converting alignments...
100%|████████████████████████████████| 2420/2420 [00:01<00:00, 213
5.23it/s]
INFO - Compiling training graphs...
100%|████████████████████████████████| 2420/2420 [00:05<00:00, 43
5.39it/s]
INFO - Initialization complete!
INFO - sat_2 - Iteration 1 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 120
6.60it/s]
INFO - sat_2 - Iteration 2 of 35
INFO - Calculating fMLLR for speaker adaptation...
100%|██████████████████████████████████| 12/12 [00:04<00:00,
2.47it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 113
7.46it/s]
INFO - sat_2 - Iteration 3 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 115
3.10it/s]
INFO - sat_2 - Iteration 4 of 35
INFO - Calculating fMLLR for speaker adaptation...
100%|██████████████████████████████████| 12/12 [00:04<00:00,
2.48it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 114
7.54it/s]
INFO - sat_2 - Iteration 5 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 114
9.94it/s]
INFO - sat_2 - Iteration 6 of 35
INFO - Calculating fMLLR for speaker adaptation...
100%|██████████████████████████████████| 12/12 [00:04<00:00,
2.54it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 113
6.61it/s]
```

INFO - sat_2 - Iteration 7 of 35
INFO - Accumulating statistics...
100%|██████████████████████████████| 2420/2420 [00:02<00:00, 114
0.43it/s]
INFO - sat_2 - Iteration 8 of 35
INFO - Accumulating statistics...
100%|██████████████████████████████| 2420/2420 [00:02<00:00, 111
6.03it/s]
INFO - sat_2 - Iteration 9 of 35
INFO - Accumulating statistics...
100%|██████████████████████████████| 2420/2420 [00:02<00:00, 110
9.13it/s]
INFO - sat_2 - Iteration 10 of 35
INFO - Generating alignments...
100%|██████████████████████████████| 2420/2420 [00:04<00:00, 55
1.73it/s]
WARNING - No files were aligned, this likely indicates serious problems
with the
                        aligner.
INFO - Accumulating statistics...
100%|██████████████████████████████| 2420/2420 [00:02<00:00, 109
6.57it/s]
INFO - sat_2 - Iteration 11 of 35
INFO - Accumulating statistics...
100%|██████████████████████████████| 2420/2420 [00:02<00:00, 107
4.44it/s]
INFO - sat_2 - Iteration 12 of 35
INFO - Calculating fMLLR for speaker adaptation...
100%|██████████████████████████████| 12/12 [00:04<00:00,
2.49it/s]
INFO - Accumulating statistics...
100%|██████████████████████████████| 2420/2420 [00:02<00:00, 106
3.13it/s]
INFO - sat_2 - Iteration 13 of 35
INFO - Accumulating statistics...
100%|██████████████████████████████| 2420/2420 [00:02<00:00, 104
8.97it/s]
INFO - sat_2 - Iteration 14 of 35
INFO - Accumulating statistics...
100%|██████████████████████████████| 2420/2420 [00:02<00:00, 103
4.18it/s]
INFO - sat_2 - Iteration 15 of 35
INFO - Accumulating statistics...
100%|██████████████████████████████| 2420/2420 [00:02<00:00, 103
6.92it/s]
INFO - sat_2 - Iteration 16 of 35
INFO - Accumulating statistics...
100%|██████████████████████████████| 2420/2420 [00:02<00:00, 100
8.15it/s]
INFO - sat_2 - Iteration 17 of 35
INFO - Accumulating statistics...
100%|██████████████████████████████| 2420/2420 [00:02<00:00, 100
4.07it/s]
INFO - sat_2 - Iteration 18 of 35
INFO - Accumulating statistics...
100%|██████████████████████████████| 2420/2420 [00:02<00:00, 98
2.25it/s]
INFO - sat_2 - Iteration 19 of 35
INFO - Accumulating statistics...
100%|██████████████████████████████| 2420/2420 [00:02<00:00, 97

```
0.50it/s]
INFO - sat_2 - Iteration 20 of 35
INFO - Generating alignments...
100%|████████████████████████████████| 2420/2420 [00:04<00:00, 48
9.01it/s]
WARNING - No files were aligned, this likely indicates serious problems
with the
                        aligner.
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 95
6.59it/s]
INFO - sat_2 - Iteration 21 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 96
0.03it/s]
INFO - sat_2 - Iteration 22 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 93
2.68it/s]
INFO - sat_2 - Iteration 23 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 92
5.60it/s]
INFO - sat_2 - Iteration 24 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 90
2.80it/s]
INFO - sat_2 - Iteration 25 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 88
7.65it/s]
INFO - sat_2 - Iteration 26 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 86
8.24it/s]
INFO - sat_2 - Iteration 27 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 83
5.84it/s]
INFO - sat_2 - Iteration 28 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 85
2.47it/s]
INFO - sat_2 - Iteration 29 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 84
6.72it/s]
INFO - sat_2 - Iteration 30 of 35
INFO - Generating alignments...
100%|████████████████████████████████| 2420/2420 [00:05<00:00, 44
2.21it/s]
WARNING - No files were aligned, this likely indicates serious problems
with the
                        aligner.
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 83
8.56it/s]
INFO - sat_2 - Iteration 31 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 84
```

```
8.11it/s]
INFO - sat_2 - Iteration 32 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 83
3.67it/s]
INFO - sat_2 - Iteration 33 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 83
6.40it/s]
INFO - sat_2 - Iteration 34 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 84
1.18it/s]
INFO - sat_2 - Iteration 35 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 82
9.22it/s]
INFO - Creating alignment model for speaker-independent features...
100%|████████████████████████████████| 2420/2420 [00:03<00:00, 68
5.38it/s]
INFO - Training complete!
INFO - Compiling training graphs...
100%|████████████████████████████████| 2420/2420 [00:05<00:00, 42
0.12it/s]
INFO - Generating alignments...
100%|████████████████████████████████| 2420/2420 [00:05<00:00, 41
3.93it/s]
INFO - Calculating fMLLR for speaker adaptation...
100%|██████████████████████████████████| 12/12 [00:05<00:00,
2.37it/s]
INFO - Generating alignments...
100%|████████████████████████████████| 2420/2420 [00:05<00:00, 42
4.13it/s]
INFO - Generating pronunciations...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 103
6.06it/s]
INFO - Compiling training graphs...
100%|████████████████████████████████| 2420/2420 [00:05<00:00, 43
0.62it/s]
INFO - Generating alignments...
100%|████████████████████████████████| 2420/2420 [00:05<00:00, 42
6.17it/s]
INFO - Calculating fMLLR for speaker adaptation...
100%|██████████████████████████████████| 12/12 [00:05<00:00,
2.39it/s]
INFO - Generating alignments...
100%|████████████████████████████████| 2420/2420 [00:05<00:00, 42
6.31it/s]
INFO - Initializing training for sat_3...
INFO - Converting alignments...
100%|████████████████████████████████| 2420/2420 [00:01<00:00, 213
7.61it/s]
INFO - Compiling training graphs...
100%|████████████████████████████████| 2420/2420 [00:05<00:00, 43
5.80it/s]
INFO - Initialization complete!
INFO - sat_3 - Iteration 1 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████| 2420/2420 [00:02<00:00, 120
8.47it/s]
```

INFO - sat_3 - Iteration 2 of 35
INFO - Calculating fMLLR for speaker adaptation...
100%|███████████████████████████████| 12/12 [00:04<00:00,
2.51it/s]
INFO - Accumulating statistics...
100%|███████████████████████████████| 2420/2420 [00:02<00:00, 114
3.63it/s]
INFO - sat_3 - Iteration 3 of 35
INFO - Accumulating statistics...
100%|███████████████████████████████| 2420/2420 [00:02<00:00, 114
5.77it/s]
INFO - sat_3 - Iteration 4 of 35
INFO - Calculating fMLLR for speaker adaptation...
100%|███████████████████████████████| 12/12 [00:05<00:00,
2.39it/s]
INFO - Accumulating statistics...
100%|███████████████████████████████| 2420/2420 [00:02<00:00, 107
6.62it/s]
INFO - sat_3 - Iteration 5 of 35
INFO - Accumulating statistics...
100%|███████████████████████████████| 2420/2420 [00:02<00:00, 105
9.10it/s]
INFO - sat_3 - Iteration 6 of 35
INFO - Calculating fMLLR for speaker adaptation...
100%|███████████████████████████████| 12/12 [00:05<00:00,
2.36it/s]
INFO - Accumulating statistics...
100%|███████████████████████████████| 2420/2420 [00:02<00:00, 99
9.60it/s]
INFO - sat_3 - Iteration 7 of 35
INFO - Accumulating statistics...
100%|███████████████████████████████| 2420/2420 [00:02<00:00, 100
6.45it/s]
INFO - sat_3 - Iteration 8 of 35
INFO - Accumulating statistics...
100%|███████████████████████████████| 2420/2420 [00:02<00:00, 98
6.72it/s]
INFO - sat_3 - Iteration 9 of 35
INFO - Accumulating statistics...
100%|███████████████████████████████| 2420/2420 [00:02<00:00, 92
6.43it/s]
INFO - sat_3 - Iteration 10 of 35
INFO - Generating alignments...
100%|███████████████████████████████| 2420/2420 [00:05<00:00, 45
4.66it/s]
WARNING - No files were aligned, this likely indicates serious problems
with the
                    aligner.
INFO - Accumulating statistics...
100%|███████████████████████████████| 2420/2420 [00:02<00:00, 90
5.02it/s]
INFO - sat_3 - Iteration 11 of 35
INFO - Accumulating statistics...
100%|███████████████████████████████| 2420/2420 [00:02<00:00, 85
6.08it/s]
INFO - sat_3 - Iteration 12 of 35
INFO - Calculating fMLLR for speaker adaptation...
100%|███████████████████████████████| 12/12 [00:05<00:00,
2.31it/s]
INFO - Accumulating statistics...

```
100%|████████████████████████████████████| 2420/2420 [00:03<00:00, 78
0.73it/s]
INFO - sat_3 - Iteration 13 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:05<00:00, 48
2.69it/s]
INFO - sat_3 - Iteration 14 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:11<00:00, 21
8.60it/s]
INFO - sat_3 - Iteration 15 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:18<00:00, 13
1.92it/s]
INFO - sat_3 - Iteration 16 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:18<00:00, 13
0.66it/s]
INFO - sat_3 - Iteration 17 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:18<00:00, 13
1.48it/s]
INFO - sat_3 - Iteration 18 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:18<00:00, 13
0.46it/s]
INFO - sat_3 - Iteration 19 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:18<00:00, 13
1.85it/s]
INFO - sat_3 - Iteration 20 of 35
INFO - Generating alignments...
100%|████████████████████████████████████| 2420/2420 [00:26<00:00, 8
9.71it/s]
WARNING - No files were aligned, this likely indicates serious problems
with the
                       aligner.
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:20<00:00, 11
9.26it/s]
INFO - sat_3 - Iteration 21 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:19<00:00, 12
6.08it/s]
INFO - sat_3 - Iteration 22 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:19<00:00, 12
6.22it/s]
INFO - sat_3 - Iteration 23 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:19<00:00, 12
4.48it/s]
INFO - sat_3 - Iteration 24 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:18<00:00, 12
8.36it/s]
INFO - sat_3 - Iteration 25 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:19<00:00, 12
6.20it/s]
```

```
INFO - sat_3 - Iteration 26 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:18<00:00, 12
7.45it/s]
INFO - sat_3 - Iteration 27 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:21<00:00, 11
2.92it/s]
INFO - sat_3 - Iteration 28 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:21<00:00, 11
2.60it/s]
INFO - sat_3 - Iteration 29 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:20<00:00, 11
7.14it/s]
INFO - sat_3 - Iteration 30 of 35
INFO - Generating alignments...
100%|████████████████████████████████████| 2420/2420 [00:26<00:00, 9
2.91it/s]
WARNING - No files were aligned, this likely indicates serious problems
with the
                          aligner.
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:26<00:00, 9
2.67it/s]
INFO - sat_3 - Iteration 31 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:21<00:00, 11
3.89it/s]
INFO - sat_3 - Iteration 32 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:25<00:00, 9
4.43it/s]
INFO - sat_3 - Iteration 33 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:24<00:00, 9
9.52it/s]
INFO - sat_3 - Iteration 34 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:24<00:00, 9
7.74it/s]
INFO - sat_3 - Iteration 35 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 2420/2420 [00:20<00:00, 11
9.15it/s]
INFO - Creating alignment model for speaker-independent features...
100%|████████████████████████████████████| 2420/2420 [00:20<00:00, 11
6.94it/s]
INFO - Training complete!
INFO - Exiting training early to save time as the corpus is below the s
ubset
                  size for later training stages
INFO - Compiling training graphs...
100%|████████████████████████████████████| 2420/2420 [00:05<00:00, 43
5.60it/s]
INFO - Generating alignments...
100%|████████████████████████████████████| 2420/2420 [00:24<00:00, 10
0.21it/s]
INFO - Calculating fMLLR for speaker adaptation...
```

```
100%|████████████████████████████████████████| 12/12 [00:05<00:00,
2.26it/s]
INFO - Generating alignments...
100%|████████████████████████████████████████| 2420/2420 [00:23<00:00, 10
2.09it/s]
INFO - Accumulating transition stats...
100%|████████████████████████████████████████| 2420/2420 [00:01<00:00, 203
2.67it/s]
INFO - Finished accumulating transition stats!
  0%|                                         | 0/2420 [00:00
<?, ?it/s]INFO - Collecting phone and word alignments from sat_3_ali la
ttices...
100%|████████████████████████████████████████| 2420/2420 [00:04<00:00, 56
9.52it/s]
INFO - Beginning phone LM training...
INFO - Collecting training data...
  0%|                                         | 0/2420 [00:00
<?, ?it/s]Exception during reset or similar
Traceback (most recent call last):
  File "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-pack
ages/sqlalchemy/pool/base.py", line 991, in _finalize_fairy
    fairy._reset(
  File "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-pack
ages/sqlalchemy/pool/base.py", line 1440, in _reset
    pool._dialect.do_rollback(self)
  File "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-pack
ages/sqlalchemy/engine/default.py", line 657, in do_rollback
    dbapi_connection.rollback()
psycopg2.OperationalError: server closed the connection unexpectedly
        This probably means the server terminated abnormally
        before or while processing the request.

Exception during reset or similar
Traceback (most recent call last):
  File "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-pack
ages/sqlalchemy/pool/base.py", line 991, in _finalize_fairy
    fairy._reset(
  File "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-pack
ages/sqlalchemy/pool/base.py", line 1440, in _reset
    pool._dialect.do_rollback(self)
  File "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-pack
ages/sqlalchemy/engine/default.py", line 657, in do_rollback
    dbapi_connection.rollback()
psycopg2.OperationalError: server closed the connection unexpectedly
        This probably means the server terminated abnormally
        before or while processing the request.

100%|████████████████████████████████████████| 2420/2420 [00:01<00:00, 211
3.04it/s]
INFO - Training model...
Exception during reset or similar
Traceback (most recent call last):
  File "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-pack
ages/sqlalchemy/pool/base.py", line 991, in _finalize_fairy
    fairy._reset(
  File "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-pack
ages/sqlalchemy/pool/base.py", line 1440, in _reset
    pool._dialect.do_rollback(self)
  File "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-pack
ages/sqlalchemy/engine/default.py", line 657, in do_rollback
```

```
    dbapi_connection.rollback()
psycopg2.OperationalError: server closed the connection unexpectedly
        This probably means the server terminated abnormally
        before or while processing the request.

INFO - Completed training in 1591.5464088916779 seconds!
INFO - Saved model to /home/philipp/Workspace/corpora/Xhosa_am.zip
INFO - Done! Everything took 1632.307 seconds
```

We can now align our files using the dataset, our generated dictionary and the Xhosa acoustic model. The alignments will be saved into a folder called "Xhosa_alignments".
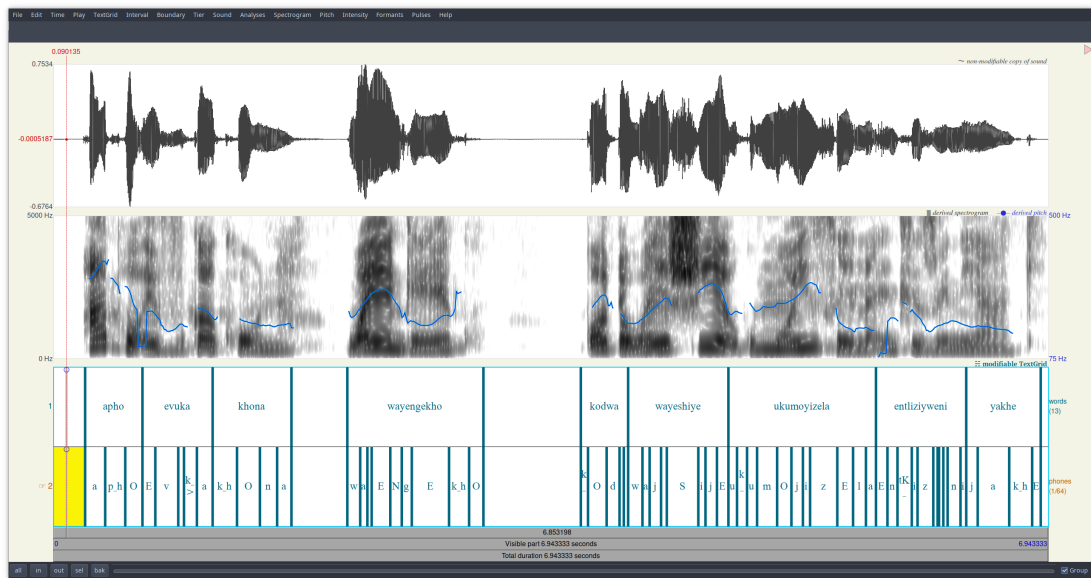
In [9]: `! mfa align ~/Workspace/corpora/Xhosa_dataset/ ~/Workspace/corpora/Xhos`

```
INFO - Setting up corpus information...
INFO - Loading corpus from source files...
100%|████████████████████████████████| 2420/2420 [00:01<00:00, 152
6.03it/s]
INFO - Found 12 speakers across 2420 files, average number of utterance
s per
                 speaker: 201.66666666666666
INFO - Initializing multiprocessing jobs...
INFO - Normalizing text...
100%|████████████████████████████████| 2420/2420 [00:01<00:00, 134
2.67it/s]
INFO - Creating corpus split for feature generation...
100%|████████████████████████████████| 4840/4840 [00:01<00:00, 426
6.85it/s]
INFO - Generating MFCCs...
2433it [00:19, 127.67it/s]
INFO - Calculating CMVN...
INFO - Generating final features...
100%|████████████████████████████████| 2420/2420 [00:01<00:00, 140
1.24it/s]
INFO - Creating corpus split with features...
100%|████████████████████████████████| 2420/2420 [00:01<00:00, 206
4.20it/s]
INFO - Compiling training graphs...
100%|████████████████████████████████| 2420/2420 [00:06<00:00, 38
8.34it/s]
INFO - Performing first-pass alignment...
INFO - Generating alignments...
100%|████████████████████████████████| 2420/2420 [00:30<00:00, 7
8.59it/s]
INFO - Calculating fMLLR for speaker adaptation...
100%|██████████████████████████████████| 12/12 [00:05<00:00,
2.20it/s]
INFO - Performing second-pass alignment...
INFO - Generating alignments...
100%|████████████████████████████████| 2420/2420 [00:29<00:00, 8
1.40it/s]
  0%|                                | 0/2420 [00:00
<?, ?it/s]INFO - Collecting phone and word alignments from alignment la
ttices...
100%|████████████████████████████████| 2420/2420 [00:04<00:00, 55
3.78it/s]
INFO - Exporting alignment TextGrids to
                 /home/philipp/Workspace/corpora/Xhosa_alignments...
100%|████████████████████████████████| 2420/2420 [00:08<00:00, 29
6.35it/s]
INFO - Finished exporting TextGrids to
                 /home/philipp/Workspace/corpora/Xhosa_alignments!
INFO - Done! Everything took 127.359 seconds
```

We can now look at our TextGrids. Words are well-aligned in general and phone alignments are usable. Recall again that we had a lot of variation in our dictionary and that we not corrected it, so the resulting phone alignments may not be accurate. It is thus important to check the pronunciations listed in the dictionary and to correct the dictionary if necessary.

In [ ]: