Training an acoustic model on Mandarin data

We will train an acoustic model of Mandarin using a large dataset (~10h) in this notebook. The data can be found in the Google Drive as THCHS30_part.zip. The zip files consists of folders for multiple speakers and that include wav and lab files. The dataset is a part of the THCHS30 corpus, the full corpus can be found here.

Download the data and unzip the file. In this example, the data is located in ~/Workspace/corpora/data_thchs30/. Instead of using a custom dictionary, we will use a dictionary from the MFA. Download the dictionary via the following command:

In [1]:
```
! mfa model download dictionary mandarin_china_mfa
```

The dataset from the Google Drive folder contains only the utterance where all words can be found in the downlaoded dictionary. However, we check first if there is an error in the corpus by using the validate command:

In [2]:
```
! mfa validate --clean ~/Workspace/corpora/thchs30_part/ mandarin_china
```

```
INFO - Setting up corpus information...
INFO - Loading corpus from source files...
4069it [00:02, 1940.99it/s]
INFO - Found 60 speakers across 4069 files, average number of utterance
s per
                speaker: 67.81666666666666
INFO - Initializing multiprocessing jobs...
INFO - Normalizing text...
100%|████████████████████████████████████| 4069/4069 [00:07<00:00, 51
4.22it/s]
INFO - Creating corpus split for feature generation...
100%|████████████████████████████████████| 8138/8138 [00:01<00:00, 675
4.94it/s]
INFO - Generating MFCCs...
4080it [00:48, 84.68it/s]
INFO - Calculating CMVN...
INFO - Generating final features...
100%|████████████████████████████████████| 4069/4069 [00:02<00:00, 139
5.21it/s]
INFO - Creating corpus split with features...
100%|████████████████████████████████████| 4069/4069 [00:01<00:00, 220
6.98it/s]

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

INFO -   4069 sound files
INFO -   4069 text files
INFO -   60 speakers
INFO -   4069 utterances
INFO -   37395.346 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 -     There were no missing words from the dictionary. If you plan
on using
                the a        model trained on this dataset to align oth
er
                datasets in the future, it is        recommended that t
here be
                at least some missing words.


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

INFO - Creating subset directory with 2000 utterances...
100%|████████████████████████████████████| 2000/2000 [00:01<00:00, 111
2.41it/s]
INFO - Initializing training for monophone...
INFO - Compiling training graphs...
100%|████████████████████████████████████| 2000/2000 [00:02<00:00, 77
3.91it/s]
INFO - Generating initial alignments...
100%|████████████████████████████████████| 2000/2000 [00:03<00:00, 61
3.46it/s]
INFO - Initialization complete!
INFO - monophone - Iteration 1 of 40
INFO - Generating alignments...
  0%|                                          | 9/2000 [00:50<3:05:31,
5.59s/it]
ERROR - There was an error in the run, please see the log.
Exception ignored in atexit callback: <bound method ExitHooks.history_s
ave_handler of <montreal_forced_aligner.command_line.mfa.ExitHooks obje
ct at 0x7f2b5d9e8df0>>
Traceback (most recent call last):
  File "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-pack
ages/montreal_forced_aligner/command_line/mfa.py", line 95, in history_
save_handler
    raise self.exception
  File "/home/philipp/anaconda3/envs/mfaligner/bin/mfa", line 11, in <m
odule>
    sys.exit(mfa_cli())
  File "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-pack
ages/click/core.py", line 1130, in __call__
    return self.main(*args, **kwargs)
  File "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-pack
ages/click/core.py", line 1055, in main
    rv = self.invoke(ctx)
  File "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-pack
ages/click/core.py", line 1657, in invoke
    return _process_result(sub_ctx.command.invoke(sub_ctx))
  File "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-pack
ages/click/core.py", line 1404, in invoke
    return ctx.invoke(self.callback, **ctx.params)
  File "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-pack
ages/click/core.py", line 760, in invoke
    return __callback(*args, **kwargs)
  File "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-pack
ages/click/decorators.py", line 26, in new_func
    return f(get_current_context(), *args, **kwargs)
  File "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-pack
```

```
ages/montreal_forced_aligner/command_line/validate.py", line 116, in va
lidate_corpus_cli
    validator.validate()
  File "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-pack
ages/montreal_forced_aligner/validation/corpus_validator.py", line 664,
in validate
    self.train()
  File "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-pack
ages/montreal_forced_aligner/acoustic_modeling/trainer.py", line 529, i
n train
    trainer.train()
  File "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-pack
ages/montreal_forced_aligner/acoustic_modeling/base.py", line 494, in t
rain
    self.train_iteration()
  File "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-pack
ages/montreal_forced_aligner/acoustic_modeling/base.py", line 466, in t
rain_iteration
    self.align_iteration()
  File "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-pack
ages/montreal_forced_aligner/acoustic_modeling/base.py", line 439, in a
lign_iteration
    self.align_utterances(training=True)
  File "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-pack
ages/montreal_forced_aligner/alignment/mixins.py", line 473, in align_u
tterances
    raise v
montreal_forced_aligner.exceptions.MultiprocessingError: Multiprocessin
gError:

  Job 2 encountered an error:
  Traceback (most recent call last):

    File
    "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-package
s/montreal_forced_aligner/abc.py",
     line 85, in run
      yield from self._run()

    File
    "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-package
s/montreal_forced_aligner/alignment/multiprocessing.py",
     line 926, in _run
      self.check_call(align_proc)

    File
    "/home/philipp/anaconda3/envs/mfaligner/lib/python3.10/site-package
s/montreal_forced_aligner/abc.py",
     line 112, in check_call
      raise KaldiProcessingError([self.log_path])

  montreal_forced_aligner.exceptions.KaldiProcessingError: KaldiProcess
ingError:

    There were 1 job(s) with errors when running Kaldi binaries.
    See the log files below for more information.
    /home/philipp/Documents/MFA/thchs30_part/monophone/log/align.1.2.lo
g
```

It seems that the corpus has no errors and no OOVs were found. In the next step, we train the acoustic model. I named the acoustic model thchs30_model.zip. This may take a while.

In [3]:
```
! mfa train --clean ~/Workspace/corpora/thchs30_part/ mandarin_china_mf
```

```
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...
4069it [00:02, 1975.60it/s]
INFO - Found 60 speakers across 4069 files, average number of utterance
s per
               speaker: 67.81666666666666
INFO - Initializing multiprocessing jobs...
INFO - Normalizing text...
100%|████████████████████████████████████| 4069/4069 [00:07<00:00, 53
4.90it/s]
INFO - Creating corpus split for feature generation...
100%|████████████████████████████████████| 8138/8138 [00:01<00:00, 669
2.57it/s]
INFO - Generating MFCCs...
4080it [00:46, 87.00it/s]
INFO - Calculating CMVN...
INFO - Generating final features...
100%|████████████████████████████████████| 4069/4069 [00:02<00:00, 136
2.49it/s]
INFO - Creating corpus split with features...
100%|████████████████████████████████████| 4069/4069 [00:01<00:00, 219
4.61it/s]
INFO - Initializing training for monophone...
INFO - Compiling training graphs...
100%|████████████████████████████████████| 4069/4069 [00:03<00:00, 104
6.46it/s]
INFO - Generating initial alignments...
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 74
6.40it/s]
INFO - Initialization complete!
INFO - monophone - Iteration 1 of 40
INFO - Generating alignments...
  1%|▌                                    | 39/4069 [01:29<2:34:03,
2.29s/it]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:01<00:00, 268
9.01it/s]
INFO - monophone - Iteration 2 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 4054/4069 [01:59<00:00, 3
3.83it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 96
4.88it/s]
INFO - monophone - Iteration 3 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 4069/4069 [00:32<00:00, 12
5.83it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 95
1.23it/s]
INFO - monophone - Iteration 4 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 4069/4069 [00:21<00:00, 18
9.25it/s]
```

```
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 93
8.59it/s]
INFO - monophone - Iteration 5 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 4069/4069 [00:15<00:00, 25
6.34it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 94
7.54it/s]
INFO - monophone - Iteration 6 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 4069/4069 [00:14<00:00, 28
4.78it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 95
0.05it/s]
INFO - monophone - Iteration 7 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 4069/4069 [00:12<00:00, 31
6.53it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 95
5.51it/s]
INFO - monophone - Iteration 8 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 4069/4069 [00:11<00:00, 34
6.47it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 94
8.03it/s]
INFO - monophone - Iteration 9 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 4068/4069 [00:11<00:00, 35
9.90it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 94
9.20it/s]
INFO - monophone - Iteration 10 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 4068/4069 [00:10<00:00, 37
7.76it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 93
2.12it/s]
INFO - monophone - Iteration 11 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 94
2.80it/s]
INFO - monophone - Iteration 12 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 4068/4069 [00:10<00:00, 38
1.93it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 94
3.47it/s]
INFO - monophone - Iteration 13 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 91
6.81it/s]
```

```
INFO - monophone - Iteration 14 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 4068/4069 [00:10<00:00, 37
9.85it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 95
5.31it/s]
INFO - monophone - Iteration 15 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 90
6.71it/s]
INFO - monophone - Iteration 16 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 4068/4069 [00:10<00:00, 37
9.58it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 92
1.98it/s]
INFO - monophone - Iteration 17 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 91
4.27it/s]
INFO - monophone - Iteration 18 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 4068/4069 [00:10<00:00, 37
5.73it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 94
0.79it/s]
INFO - monophone - Iteration 19 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 89
6.91it/s]
INFO - monophone - Iteration 20 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 4068/4069 [00:10<00:00, 37
6.37it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 91
3.02it/s]
INFO - monophone - Iteration 21 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 90
1.66it/s]
INFO - monophone - Iteration 22 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 90
3.76it/s]
INFO - monophone - Iteration 23 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 4068/4069 [00:10<00:00, 37
5.44it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 92
1.25it/s]
INFO - monophone - Iteration 24 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 92
5.09it/s]
INFO - monophone - Iteration 25 of 40
```

```
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 90
7.42it/s]
INFO - monophone - Iteration 26 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 4068/4069 [00:11<00:00, 36
7.68it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 91
4.17it/s]
INFO - monophone - Iteration 27 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 91
9.06it/s]
INFO - monophone - Iteration 28 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 91
1.58it/s]
INFO - monophone - Iteration 29 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 4068/4069 [00:10<00:00, 37
4.17it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 87
6.92it/s]
INFO - monophone - Iteration 30 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 88
3.91it/s]
INFO - monophone - Iteration 31 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 90
4.39it/s]
INFO - monophone - Iteration 32 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 4068/4069 [00:11<00:00, 35
5.68it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 89
5.01it/s]
INFO - monophone - Iteration 33 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 88
9.52it/s]
INFO - monophone - Iteration 34 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 89
4.69it/s]
INFO - monophone - Iteration 35 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 4068/4069 [00:10<00:00, 37
1.50it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 88
7.33it/s]
INFO - monophone - Iteration 36 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 89
6.55it/s]
INFO - monophone - Iteration 37 of 40
```

```
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 90
4.28it/s]
INFO - monophone - Iteration 38 of 40
INFO - Generating alignments...
100%|████████████████████████████████████| 4068/4069 [00:11<00:00, 36
6.25it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 89
4.27it/s]
INFO - monophone - Iteration 39 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 88
7.05it/s]
INFO - monophone - Iteration 40 of 40
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 86
3.36it/s]
INFO - Training complete!
INFO - Compiling training graphs...
100%|████████████████████████████████████| 4069/4069 [00:03<00:00, 103
1.74it/s]
INFO - Generating alignments...
100%|████████████████████████████████████| 4068/4069 [00:11<00:00, 34
5.97it/s]
INFO - Initializing training for triphone...
INFO - Compiling training graphs...
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 73
9.73it/s]
INFO - Converting alignments...
100%|████████████████████████████████████| 4068/4069 [00:01<00:00, 312
0.73it/s]
INFO - Initialization complete!
INFO - triphone - Iteration 1 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 92
3.35it/s]
INFO - triphone - Iteration 2 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 93
2.76it/s]
INFO - triphone - Iteration 3 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 89
9.10it/s]
INFO - triphone - Iteration 4 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 92
0.78it/s]
INFO - triphone - Iteration 5 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 93
7.65it/s]
INFO - triphone - Iteration 6 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 91
6.88it/s]
INFO - triphone - Iteration 7 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 89
```

```
6.36it/s]
INFO - triphone - Iteration 8 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 89
7.95it/s]
INFO - triphone - Iteration 9 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 89
4.92it/s]
INFO - triphone - Iteration 10 of 35
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:10<00:00, 37
4.80it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 88
7.12it/s]
INFO - triphone - Iteration 11 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 89
1.25it/s]
INFO - triphone - Iteration 12 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 90
0.82it/s]
INFO - triphone - Iteration 13 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 88
4.93it/s]
INFO - triphone - Iteration 14 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 89
5.02it/s]
INFO - triphone - Iteration 15 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 88
3.84it/s]
INFO - triphone - Iteration 16 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 87
8.09it/s]
INFO - triphone - Iteration 17 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 87
1.41it/s]
INFO - triphone - Iteration 18 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 85
5.73it/s]
INFO - triphone - Iteration 19 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 85
6.85it/s]
INFO - triphone - Iteration 20 of 35
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:11<00:00, 35
5.33it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 86
7.13it/s]
INFO - triphone - Iteration 21 of 35
```

```
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 85
0.40it/s]
INFO - triphone - Iteration 22 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 84
0.36it/s]
INFO - triphone - Iteration 23 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 84
0.52it/s]
INFO - triphone - Iteration 24 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 83
6.65it/s]
INFO - triphone - Iteration 25 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 84
4.73it/s]
INFO - triphone - Iteration 26 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 83
3.34it/s]
INFO - triphone - Iteration 27 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 84
4.12it/s]
INFO - triphone - Iteration 28 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 84
8.46it/s]
INFO - triphone - Iteration 29 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 82
3.29it/s]
INFO - triphone - Iteration 30 of 35
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:11<00:00, 34
6.66it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 82
3.84it/s]
INFO - triphone - Iteration 31 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 82
7.99it/s]
INFO - triphone - Iteration 32 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 83
7.22it/s]
INFO - triphone - Iteration 33 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 83
5.71it/s]
INFO - triphone - Iteration 34 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 82
8.45it/s]
INFO - triphone - Iteration 35 of 35
INFO - Accumulating statistics...
```

```
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 81
9.39it/s]
INFO - Training complete!
INFO - Compiling training graphs...
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 73
7.04it/s]
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:12<00:00, 33
4.29it/s]
INFO - Initializing training for lda...
100%|████████████████████████████████████| 4069/4069 [00:02<00:00, 173
1.64it/s]
INFO - Compiling training graphs...
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 75
4.23it/s]
INFO - Converting alignments...
100%|████████████████████████████████████| 4065/4069 [00:01<00:00, 308
1.88it/s]
INFO - Initialization complete!
INFO - lda - Iteration 1 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:03<00:00, 104
3.55it/s]
INFO - lda - Iteration 2 of 35
INFO - Re-calculating LDA...
100%|████████████████████████████████████| 4065/4069 [00:05<00:00, 81
0.69it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 101
3.67it/s]
INFO - lda - Iteration 3 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 101
1.01it/s]
INFO - lda - Iteration 4 of 35
INFO - Re-calculating LDA...
100%|████████████████████████████████████| 4065/4069 [00:04<00:00, 82
4.58it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 101
1.33it/s]
INFO - lda - Iteration 5 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 100
9.30it/s]
INFO - lda - Iteration 6 of 35
INFO - Re-calculating LDA...
100%|████████████████████████████████████| 4065/4069 [00:05<00:00, 81
2.54it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 98
0.54it/s]
INFO - lda - Iteration 7 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 99
1.95it/s]
INFO - lda - Iteration 8 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 97
5.88it/s]
```

```
INFO - lda - Iteration 9 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 965.64it/s]
INFO - lda - Iteration 10 of 35
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:10<00:00, 397.67it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 961.12it/s]
INFO - lda - Iteration 11 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 964.00it/s]
INFO - lda - Iteration 12 of 35
INFO - Re-calculating LDA...
100%|████████████████████████████████████| 4065/4069 [00:04<00:00, 824.56it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 957.59it/s]
INFO - lda - Iteration 13 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 945.77it/s]
INFO - lda - Iteration 14 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 922.44it/s]
INFO - lda - Iteration 15 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 934.04it/s]
INFO - lda - Iteration 16 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 923.65it/s]
INFO - lda - Iteration 17 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 922.94it/s]
INFO - lda - Iteration 18 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 916.84it/s]
INFO - lda - Iteration 19 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 910.82it/s]
INFO - lda - Iteration 20 of 35
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:10<00:00, 379.17it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 904.76it/s]
INFO - lda - Iteration 21 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 89
```

```
6.60it/s]
INFO - lda - Iteration 22 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 89
1.33it/s]
INFO - lda - Iteration 23 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 88
4.00it/s]
INFO - lda - Iteration 24 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 88
2.49it/s]
INFO - lda - Iteration 25 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 87
7.12it/s]
INFO - lda - Iteration 26 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 87
5.16it/s]
INFO - lda - Iteration 27 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 86
2.48it/s]
INFO - lda - Iteration 28 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 86
0.62it/s]
INFO - lda - Iteration 29 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 85
6.12it/s]
INFO - lda - Iteration 30 of 35
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:11<00:00, 36
1.19it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 87
0.79it/s]
INFO - lda - Iteration 31 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 86
1.15it/s]
INFO - lda - Iteration 32 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 86
9.65it/s]
INFO - lda - Iteration 33 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 87
1.01it/s]
INFO - lda - Iteration 34 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 86
8.26it/s]
INFO - lda - Iteration 35 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 86
4.64it/s]
```

```
INFO - Training complete!
INFO - Compiling training graphs...
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 74
9.20it/s]
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:11<00:00, 34
8.73it/s]
INFO - Initializing training for sat...
INFO - Calculating fMLLR for speaker adaptation...
100%|████████████████████████████████████████| 60/60 [00:11<00:00,
5.04it/s]
INFO - Converting alignments...
100%|████████████████████████████████████| 4065/4069 [00:01<00:00, 308
5.35it/s]
INFO - Compiling training graphs...
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 73
4.42it/s]
INFO - Initialization complete!
INFO - sat - Iteration 1 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:03<00:00, 102
1.59it/s]
INFO - sat - Iteration 2 of 35
INFO - Calculating fMLLR for speaker adaptation...
100%|████████████████████████████████████████| 60/60 [00:11<00:00,
5.10it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 99
3.32it/s]
INFO - sat - Iteration 3 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 100
1.33it/s]
INFO - sat - Iteration 4 of 35
INFO - Calculating fMLLR for speaker adaptation...
100%|████████████████████████████████████████| 60/60 [00:11<00:00,
5.08it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 99
5.51it/s]
INFO - sat - Iteration 5 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 100
0.22it/s]
INFO - sat - Iteration 6 of 35
INFO - Calculating fMLLR for speaker adaptation...
100%|████████████████████████████████████████| 60/60 [00:11<00:00,
5.02it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 99
4.70it/s]
INFO - sat - Iteration 7 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 99
1.90it/s]
INFO - sat - Iteration 8 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 97
9.44it/s]
INFO - sat - Iteration 9 of 35
```

```
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 97
9.28it/s]
INFO - sat - Iteration 10 of 35
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:09<00:00, 41
0.83it/s]
WARNING - No files were aligned, this likely indicates serious problems
with the
                    aligner.
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 97
2.20it/s]
INFO - sat - Iteration 11 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 94
8.52it/s]
INFO - sat - Iteration 12 of 35
INFO - Calculating fMLLR for speaker adaptation...
100%|██████████████████████████████████████████| 60/60 [00:11<00:00,
5.06it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 96
1.41it/s]
INFO - sat - Iteration 13 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 96
5.13it/s]
INFO - sat - Iteration 14 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 96
1.33it/s]
INFO - sat - Iteration 15 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 94
6.65it/s]
INFO - sat - Iteration 16 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 94
5.36it/s]
INFO - sat - Iteration 17 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 95
7.58it/s]
INFO - sat - Iteration 18 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 93
8.82it/s]
INFO - sat - Iteration 19 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 93
8.21it/s]
INFO - sat - Iteration 20 of 35
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:10<00:00, 39
8.54it/s]
WARNING - No files were aligned, this likely indicates serious problems
with the
                    aligner.
INFO - Accumulating statistics...
```

```
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 93
9.52it/s]
INFO - sat - Iteration 21 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 92
4.04it/s]
INFO - sat - Iteration 22 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 92
5.03it/s]
INFO - sat - Iteration 23 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 92
7.54it/s]
INFO - sat - Iteration 24 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 92
3.92it/s]
INFO - sat - Iteration 25 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 91
3.29it/s]
INFO - sat - Iteration 26 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 91
8.80it/s]
INFO - sat - Iteration 27 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 90
5.71it/s]
INFO - sat - Iteration 28 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 91
5.60it/s]
INFO - sat - Iteration 29 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 91
1.00it/s]
INFO - sat - Iteration 30 of 35
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:10<00:00, 37
1.51it/s]
WARNING - No files were aligned, this likely indicates serious problems
with the
                    aligner.
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 91
0.39it/s]
INFO - sat - Iteration 31 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 88
2.94it/s]
INFO - sat - Iteration 32 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 91
3.96it/s]
INFO - sat - Iteration 33 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 90
4.49it/s]
```

```
INFO - sat - Iteration 34 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 906.00it/s]
INFO - sat - Iteration 35 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 912.66it/s]
INFO - Creating alignment model for speaker-independent features...
100%|████████████████████████████████████| 4065/4069 [00:05<00:00, 722.75it/s]
INFO - Training complete!
INFO - Compiling training graphs...
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 735.29it/s]
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:10<00:00, 375.21it/s]
INFO - Calculating fMLLR for speaker adaptation...
100%|████████████████████████████████████████| 60/60 [00:11<00:00,  5.13it/s]
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:11<00:00, 365.77it/s]
INFO - Initializing training for sat_2...
INFO - Converting alignments...
100%|████████████████████████████████████| 4065/4069 [00:01<00:00, 3083.62it/s]
INFO - Compiling training graphs...
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 741.54it/s]
INFO - Initialization complete!
INFO - sat_2 - Iteration 1 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:03<00:00, 1025.09it/s]
INFO - sat_2 - Iteration 2 of 35
INFO - Calculating fMLLR for speaker adaptation...
100%|████████████████████████████████████████| 60/60 [00:11<00:00,  5.09it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 998.25it/s]
INFO - sat_2 - Iteration 3 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 1005.11it/s]
INFO - sat_2 - Iteration 4 of 35
INFO - Calculating fMLLR for speaker adaptation...
100%|████████████████████████████████████████| 60/60 [00:11<00:00,  5.05it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 991.27it/s]
INFO - sat_2 - Iteration 5 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 968.06it/s]
INFO - sat_2 - Iteration 6 of 35
INFO - Calculating fMLLR for speaker adaptation...
```

```
100%|████████████████████████████████████████| 60/60 [00:11<00:00,
5.08it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:04<00:00, 95
9.94it/s]
INFO - sat_2 - Iteration 7 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:04<00:00, 95
2.41it/s]
INFO - sat_2 - Iteration 8 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:04<00:00, 93
6.20it/s]
INFO - sat_2 - Iteration 9 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:04<00:00, 93
4.01it/s]
INFO - sat_2 - Iteration 10 of 35
INFO - Generating alignments...
100%|████████████████████████████████████████| 4065/4069 [00:10<00:00, 39
3.99it/s]
WARNING - No files were aligned, this likely indicates serious problems
with the
                    aligner.
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:04<00:00, 92
4.88it/s]
INFO - sat_2 - Iteration 11 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:04<00:00, 91
7.10it/s]
INFO - sat_2 - Iteration 12 of 35
INFO - Calculating fMLLR for speaker adaptation...
100%|████████████████████████████████████████| 60/60 [00:12<00:00,
5.00it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:04<00:00, 89
5.50it/s]
INFO - sat_2 - Iteration 13 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:04<00:00, 89
2.94it/s]
INFO - sat_2 - Iteration 14 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:04<00:00, 87
9.13it/s]
INFO - sat_2 - Iteration 15 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:04<00:00, 88
4.61it/s]
INFO - sat_2 - Iteration 16 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:04<00:00, 86
9.61it/s]
INFO - sat_2 - Iteration 17 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:04<00:00, 86
8.94it/s]
INFO - sat_2 - Iteration 18 of 35
INFO - Accumulating statistics...
```

```
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 85
1.24it/s]
INFO - sat_2 - Iteration 19 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 85
0.00it/s]
INFO - sat_2 - Iteration 20 of 35
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:11<00:00, 34
4.61it/s]
WARNING - No files were aligned, this likely indicates serious problems
with the
                    aligner.
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 84
1.35it/s]
INFO - sat_2 - Iteration 21 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 83
1.34it/s]
INFO - sat_2 - Iteration 22 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 82
7.05it/s]
INFO - sat_2 - Iteration 23 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 81
7.16it/s]
INFO - sat_2 - Iteration 24 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 81
1.52it/s]
INFO - sat_2 - Iteration 25 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 79
5.53it/s]
INFO - sat_2 - Iteration 26 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 79
1.17it/s]
INFO - sat_2 - Iteration 27 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 77
8.79it/s]
INFO - sat_2 - Iteration 28 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 78
5.28it/s]
INFO - sat_2 - Iteration 29 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 78
5.16it/s]
INFO - sat_2 - Iteration 30 of 35
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:12<00:00, 33
0.05it/s]
WARNING - No files were aligned, this likely indicates serious problems
with the
                    aligner.
INFO - Accumulating statistics...
```

```
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 78
8.90it/s]
INFO - sat_2 - Iteration 31 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 78
3.67it/s]
INFO - sat_2 - Iteration 32 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 78
5.41it/s]
INFO - sat_2 - Iteration 33 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 78
8.30it/s]
INFO - sat_2 - Iteration 34 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 79
1.07it/s]
INFO - sat_2 - Iteration 35 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 78
6.94it/s]
INFO - Creating alignment model for speaker-independent features...
100%|████████████████████████████████████| 4065/4069 [00:06<00:00, 62
5.85it/s]
INFO - Training complete!
INFO - Compiling training graphs...
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 73
8.80it/s]
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:12<00:00, 32
1.15it/s]
INFO - Calculating fMLLR for speaker adaptation...
100%|██████████████████████████████████████| 60/60 [00:11<00:00,
5.02it/s]
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:12<00:00, 32
1.10it/s]
INFO - Generating pronunciations...
100%|████████████████████████████████████| 4065/4069 [00:03<00:00, 127
7.41it/s]
INFO - Compiling training graphs...
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 74
3.57it/s]
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:12<00:00, 32
5.00it/s]
INFO - Calculating fMLLR for speaker adaptation...
100%|██████████████████████████████████████| 60/60 [00:11<00:00,
5.15it/s]
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:12<00:00, 32
3.91it/s]
INFO - Initializing training for sat_3...
INFO - Converting alignments...
100%|████████████████████████████████████| 4065/4069 [00:01<00:00, 305
1.53it/s]
INFO - Compiling training graphs...
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 73
2.67it/s]
```

```
INFO - Initialization complete!
INFO - sat_3 - Iteration 1 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 101
5.58it/s]
INFO - sat_3 - Iteration 2 of 35
INFO - Calculating fMLLR for speaker adaptation...
100%|████████████████████████████████████████████| 60/60 [00:11<00:00,
5.12it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 100
3.60it/s]
INFO - sat_3 - Iteration 3 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 97
7.04it/s]
INFO - sat_3 - Iteration 4 of 35
INFO - Calculating fMLLR for speaker adaptation...
100%|████████████████████████████████████████████| 60/60 [00:12<00:00,
4.96it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 94
8.69it/s]
INFO - sat_3 - Iteration 5 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 92
8.18it/s]
INFO - sat_3 - Iteration 6 of 35
INFO - Calculating fMLLR for speaker adaptation...
100%|████████████████████████████████████████████| 60/60 [00:11<00:00,
5.01it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 90
4.50it/s]
INFO - sat_3 - Iteration 7 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 89
3.06it/s]
INFO - sat_3 - Iteration 8 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 87
1.19it/s]
INFO - sat_3 - Iteration 9 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 83
2.05it/s]
INFO - sat_3 - Iteration 10 of 35
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:11<00:00, 35
3.86it/s]
WARNING - No files were aligned, this likely indicates serious problems
with the
                    aligner.
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 82
9.93it/s]
INFO - sat_3 - Iteration 11 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:04<00:00, 81
5.16it/s]
```

```
INFO - sat_3 - Iteration 12 of 35
INFO - Calculating fMLLR for speaker adaptation...
100%|████████████████████████████████████████| 60/60 [00:12<00:00,
4.90it/s]
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:05<00:00, 79
6.32it/s]
INFO - sat_3 - Iteration 13 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:05<00:00, 77
9.22it/s]
INFO - sat_3 - Iteration 14 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:05<00:00, 76
6.37it/s]
INFO - sat_3 - Iteration 15 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:05<00:00, 74
8.83it/s]
INFO - sat_3 - Iteration 16 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:05<00:00, 73
4.47it/s]
INFO - sat_3 - Iteration 17 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:05<00:00, 71
7.04it/s]
INFO - sat_3 - Iteration 18 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:05<00:00, 70
1.38it/s]
INFO - sat_3 - Iteration 19 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:05<00:00, 68
0.01it/s]
INFO - sat_3 - Iteration 20 of 35
INFO - Generating alignments...
100%|████████████████████████████████████████| 4065/4069 [00:14<00:00, 28
7.18it/s]
WARNING - No files were aligned, this likely indicates serious problems
with the
                    aligner.
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:06<00:00, 66
8.87it/s]
INFO - sat_3 - Iteration 21 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:06<00:00, 65
0.25it/s]
INFO - sat_3 - Iteration 22 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:06<00:00, 64
2.29it/s]
INFO - sat_3 - Iteration 23 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:06<00:00, 61
5.89it/s]
INFO - sat_3 - Iteration 24 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████████| 4069/4069 [00:06<00:00, 60
```

```
8.41it/s]
INFO - sat_3 - Iteration 25 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:06<00:00, 58
7.31it/s]
INFO - sat_3 - Iteration 26 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:07<00:00, 57
3.18it/s]
INFO - sat_3 - Iteration 27 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:07<00:00, 54
3.15it/s]
INFO - sat_3 - Iteration 28 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:07<00:00, 55
9.36it/s]
INFO - sat_3 - Iteration 29 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:07<00:00, 56
1.02it/s]
INFO - sat_3 - Iteration 30 of 35
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:16<00:00, 24
2.28it/s]
WARNING - No files were aligned, this likely indicates serious problems
with the
                    aligner.
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:07<00:00, 55
6.42it/s]
INFO - sat_3 - Iteration 31 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:07<00:00, 55
7.87it/s]
INFO - sat_3 - Iteration 32 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:07<00:00, 56
3.27it/s]
INFO - sat_3 - Iteration 33 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:07<00:00, 55
8.56it/s]
INFO - sat_3 - Iteration 34 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:07<00:00, 55
1.28it/s]
INFO - sat_3 - Iteration 35 of 35
INFO - Accumulating statistics...
100%|████████████████████████████████████| 4069/4069 [00:07<00:00, 55
4.91it/s]
INFO - Creating alignment model for speaker-independent features...
100%|████████████████████████████████████| 4065/4069 [00:08<00:00, 46
7.06it/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%|████████████████████████████████████| 4069/4069 [00:05<00:00, 74
```

```
0.67it/s]
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:16<00:00, 24
0.66it/s]
INFO - Calculating fMLLR for speaker adaptation...
100%|████████████████████████████████████████| 60/60 [00:12<00:00,
4.71it/s]
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:16<00:00, 23
9.77it/s]
INFO - Accumulating transition stats...
100%|████████████████████████████████████| 4065/4069 [00:01<00:00, 282
1.36it/s]
INFO - Finished accumulating transition stats!
  0%|                                                  | 0/4069 [00:00
<?, ?it/s]INFO - Collecting phone and word alignments from sat_3_ali la
ttices...
100%|████████████████████████████████████| 4065/4069 [00:09<00:00, 42
4.49it/s]
INFO - Beginning phone LM training...
INFO - Collecting training data...
100%|████████████████████████████████████| 4065/4069 [00:01<00:00, 305
0.01it/s]
INFO - Training model...
INFO - Completed training in 2310.6078288555145 seconds!
INFO - Saved model to /home/philipp/Workspace/thchs30_model.zip
INFO - Done! Everything took 2416.310 seconds
```

Since the training of the acoustic model is complete, we can align the data:

In [4]:
```
! mfa align --clean ~/Workspace/corpora/thchs30_part/ mandarin_china_mf
```

```
INFO - Setting up corpus information...
INFO - Loading corpus from source files...
4069it [00:02, 1858.79it/s]
INFO - Found 60 speakers across 4069 files, average number of utterance
s per
                speaker: 67.81666666666666
INFO - Initializing multiprocessing jobs...
INFO - Normalizing text...
100%|████████████████████████████████████| 4069/4069 [00:06<00:00, 58
5.68it/s]
INFO - Creating corpus split for feature generation...
100%|████████████████████████████████████| 8138/8138 [00:01<00:00, 685
3.68it/s]
INFO - Generating MFCCs...
4080it [00:47, 86.78it/s]
INFO - Calculating CMVN...
INFO - Generating final features...
100%|████████████████████████████████████| 4069/4069 [00:03<00:00, 134
7.18it/s]
INFO - Creating corpus split with features...
100%|████████████████████████████████████| 4069/4069 [00:01<00:00, 223
6.32it/s]
INFO - Compiling training graphs...
100%|████████████████████████████████████| 4069/4069 [00:05<00:00, 73
5.41it/s]
INFO - Performing first-pass alignment...
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:17<00:00, 22
9.59it/s]
INFO - Calculating fMLLR for speaker adaptation...
100%|██████████████████████████████████████████████| 60/60 [00:12<00:00,
4.62it/s]
INFO - Performing second-pass alignment...
INFO - Generating alignments...
100%|████████████████████████████████████| 4065/4069 [00:17<00:00, 23
6.43it/s]
  0%|                                                  | 0/4069 [00:00
<?, ?it/s]INFO - Collecting phone and word alignments from alignment la
ttices...
100%|████████████████████████████████████| 4065/4069 [00:09<00:00, 41
6.64it/s]
INFO - Exporting alignment TextGrids to
                /home/philipp/Workspace/corpora/thchs_alignments...
100%|████████████████████████████████████| 4069/4069 [00:13<00:00, 29
2.20it/s]
INFO - Finished exporting TextGrids to
                /home/philipp/Workspace/corpora/thchs_alignments!
INFO - Done! Everything took 178.709 seconds
```

The mfa provides also an acoustic model of Mandarin. We can compare the alignments of our model with the ones generated from the pre-trained model. Note that the pre-trained model had a lot more training data than our model. Download it using the following command:

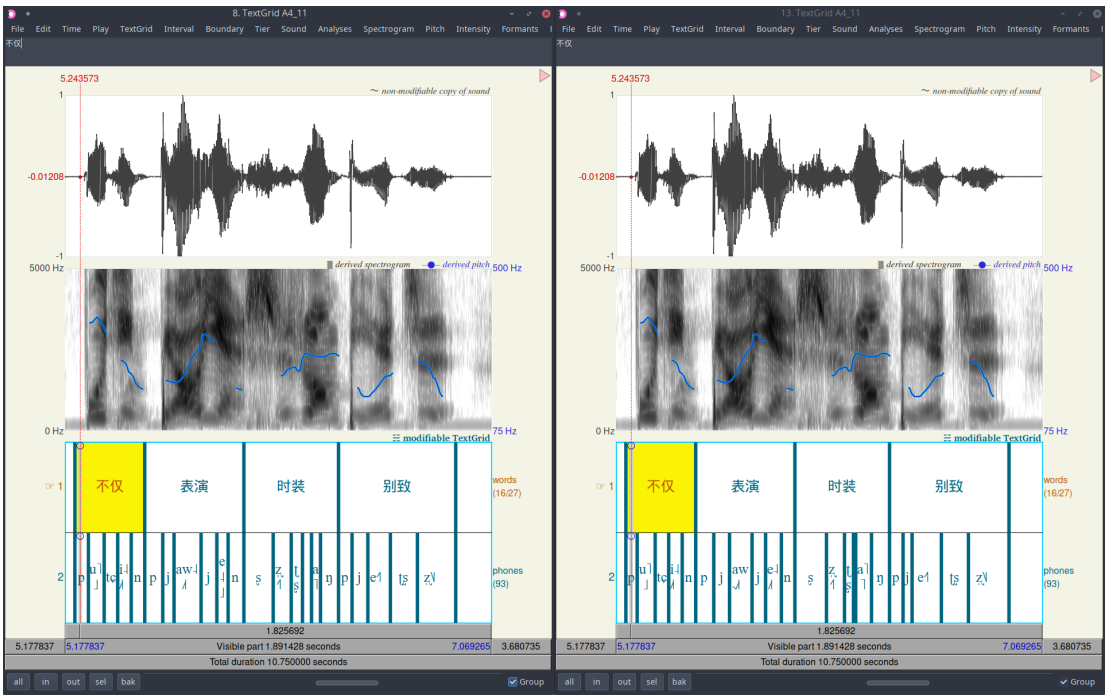
In [7]:
```
! mfa model download acoustic mandarin_mfa
```

```
INFO - Saved model to
                /home/philipp/Documents/MFA/pretrained_models/acoustic/
mandarin_
                mfa.zip, you can now use mandarin_mfa in place of acous
tic paths
                in mfa commands.
```

We align now the dataset using the pre-trained model:

In [8]:
```
! mfa align --clean ~/Workspace/corpora/thchs30_part/ mandarin_china_mf
INFO - Setting up corpus information...
INFO - Loading corpus from source files...
4069it [00:02, 1864.29it/s]
INFO - Found 60 speakers across 4069 files, average number of utterance
s per
                speaker: 67.81666666666666
INFO - Initializing multiprocessing jobs...
INFO - Normalizing text...
100%|████████████████████████████████████████| 4069/4069 [00:07<00:00, 51
8.26it/s]
INFO - Creating corpus split for feature generation...
100%|████████████████████████████████████████| 8138/8138 [00:01<00:00, 691
5.99it/s]
INFO - Generating MFCCs...
4080it [01:02, 65.39it/s]
INFO - Calculating CMVN...
INFO - Generating final features...
100%|████████████████████████████████████████| 4069/4069 [00:03<00:00, 120
7.47it/s]
INFO - Creating corpus split with features...
100%|████████████████████████████████████████| 4069/4069 [00:01<00:00, 219
7.58it/s]
INFO - Compiling training graphs...
100%|████████████████████████████████████████| 4069/4069 [00:05<00:00, 71
3.21it/s]
INFO - Performing first-pass alignment...
INFO - Generating alignments...
 98%|███████████████████████████████████████ | 3987/4069 [00:50<00:01, 7
8.46it/s]
INFO - Calculating fMLLR for speaker adaptation...
100%|████████████████████████████████████████████| 60/60 [00:12<00:00,
4.98it/s]
INFO - Performing second-pass alignment...
INFO - Generating alignments...
100%|████████████████████████████████████████| 4067/4069 [00:20<00:00, 19
7.95it/s]
  0%|                                                    | 0/4069 [00:00
<?, ?it/s]INFO - Collecting phone and word alignments from alignment la
ttices...
100%|████████████████████████████████████████| 4067/4069 [00:09<00:00, 41
3.33it/s]
INFO - Exporting alignment TextGrids to
                /home/philipp/Workspace/corpora/thchs_mfa_alignments...
100%|████████████████████████████████████████| 4069/4069 [00:13<00:00, 30
5.71it/s]
INFO - Finished exporting TextGrids to
                /home/philipp/Workspace/corpora/thchs_mfa_alignments!
INFO - Done! Everything took 232.942 seconds
```

We can compare now the alignments of our acoustic model with the ones from the pre-trained MFA model. Word-alignments are good overall and the phone aligments are comparable.



In [ ]: