Assignment 1: Forced Alignment using Montreal Forced Aligner (MFA)

Objective: To perform forced alignment using the Montreal Forced Aligner tool and understand how automatic alignment works between speech audio and phonetic transcription.

1. Environment Setup

- **Software:** Montreal Forced Aligner (v3.3.8)
- **Environment:** Anaconda (Python 3.8)
- **Dictionary Used:** english_us_arpa
- **Acoustic Model:** english_us_arpa
- **Audio Format:** 16 kHz, Mono, 16-bit PCM

2. Dataset Preparation

- A dataset with 6 audio (.wav) files and corresponding text (.txt) transcripts was used.
- Files were organized in the MFA-required format:
• corpus/wav
• corpus/transcripts
- Total duration: ~97 seconds of speech.

3. MFA Commands Executed

The following commands were used during the process:

Step 1: Validate dataset
mfa validate "D:\Assignment\Assignment\corpus" english_us_arpa english_us_arpa

Step 2: Run forced alignment
mfa align "D:\Assignment\Assignment\corpus" english_us_arpa english_us_arpa "D:\Assignment\Assignment\output"

4. Output

- MFA generated TextGrid alignment files in the output directory.
- Each TextGrid corresponds to one audio file.
- Example output:
F2BJRLP1.TextGrid, F2BJRLP2.TextGrid, etc.
- These were successfully loaded into Praat for visualization.

5. Observations

- Forced alignment was successful for all files.
- Minor out-of-vocabulary (OOV) words: 22 total (ignored).
- Word and phoneme boundaries were well aligned.
- Small deviations (~0.05–0.1 sec) possible due to natural speech variation.

6. Visualization Example (to be added)

Include a screenshot here showing alignment visualization in Praat:
View: Waveform + TextGrid (Words + Phonemes tiers)

7. Conclusion

The Montreal Forced Aligner successfully aligned the provided speech dataset with its transcripts. The process demonstrates how acoustic models, pronunciation dictionaries, and text normalization contribute to precise time-aligned phoneme and word-level annotations.