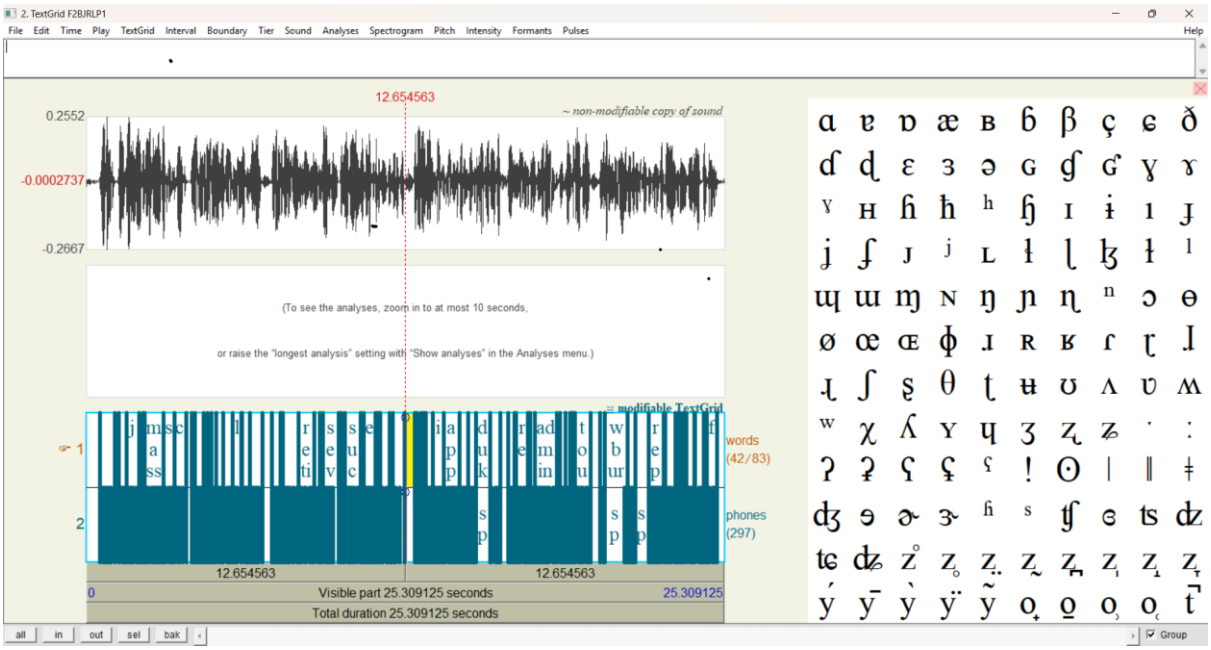
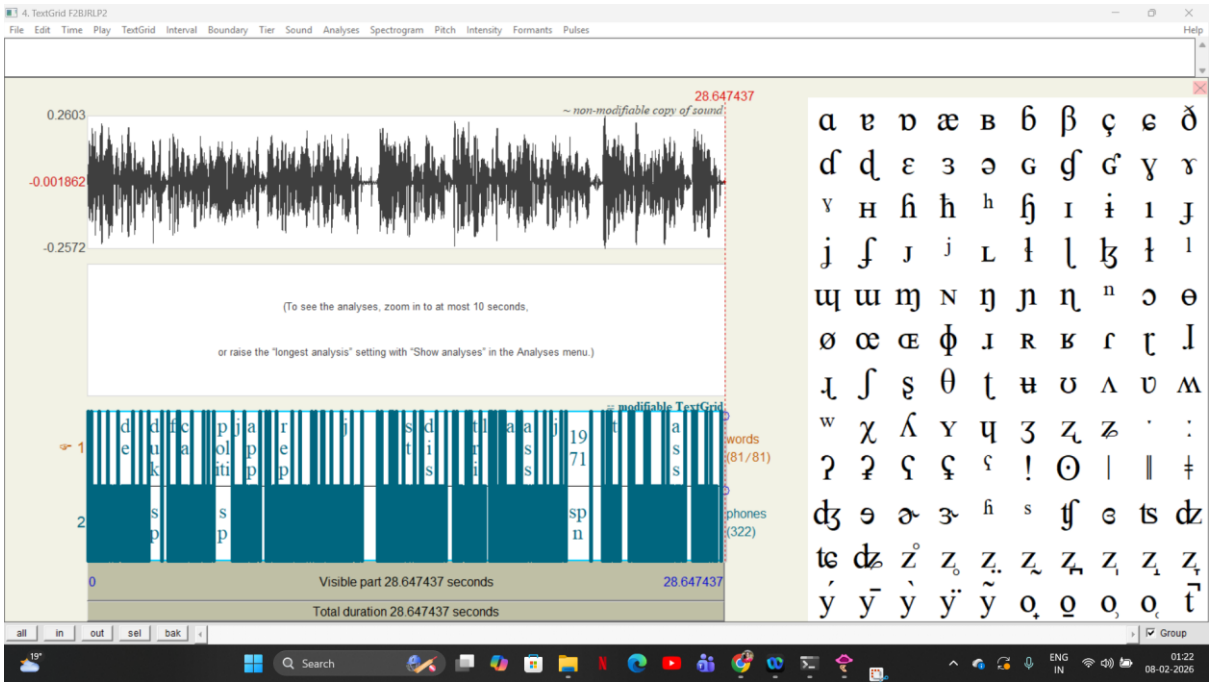
REPORT:-

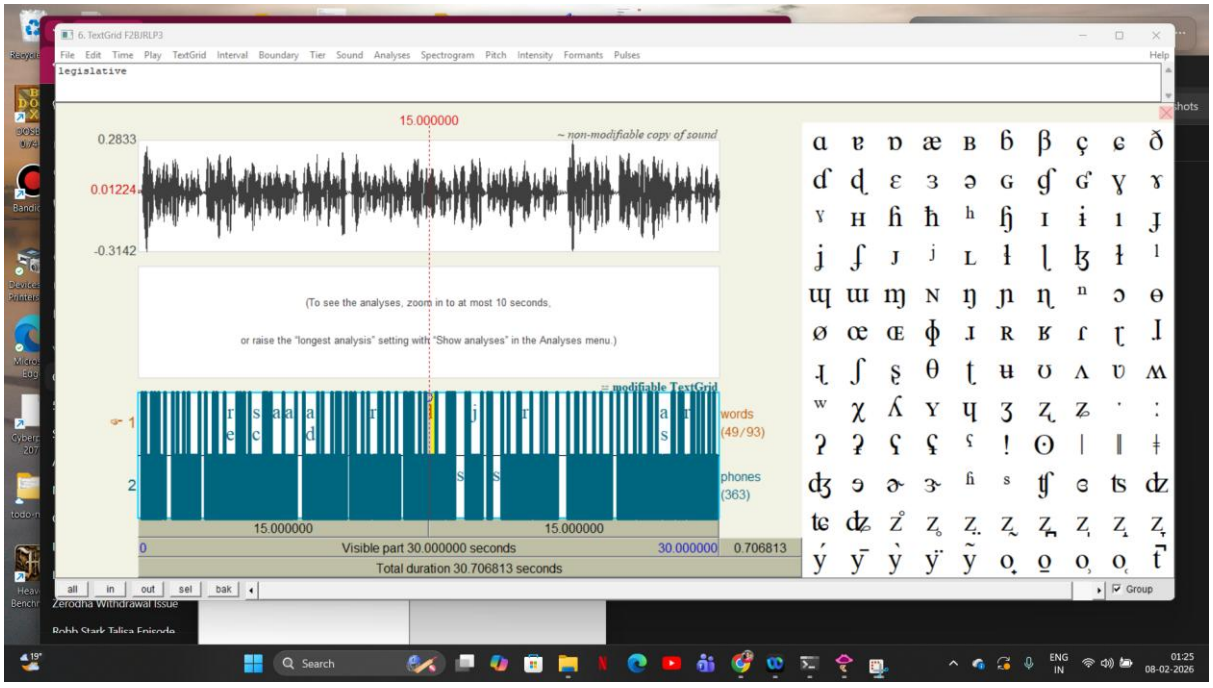
1) Forced alignment performed on speech dataset with MFA.
2) In initial alignment the english_us_arpa acoustic model and dictionary were used.
3) Visualization was done in praat for textgrid files containing word and phone tiers that were generated.
4) Some domain specific words were missing before oov handling.
5) Iiit, Hyderabad and dataset were oov words.
6) With help of MFA G2P model, pronunciations for oov words were generated and it was supplied with custom dictionary.
7) After re-alignment was done, improvement like word coverage, fewer gaps and more consistent boundaries were visible in textgrids.
8) Below screenshots before oov confirms improved quality after oov handling.

OUTPUT SCREENSHOTS:-

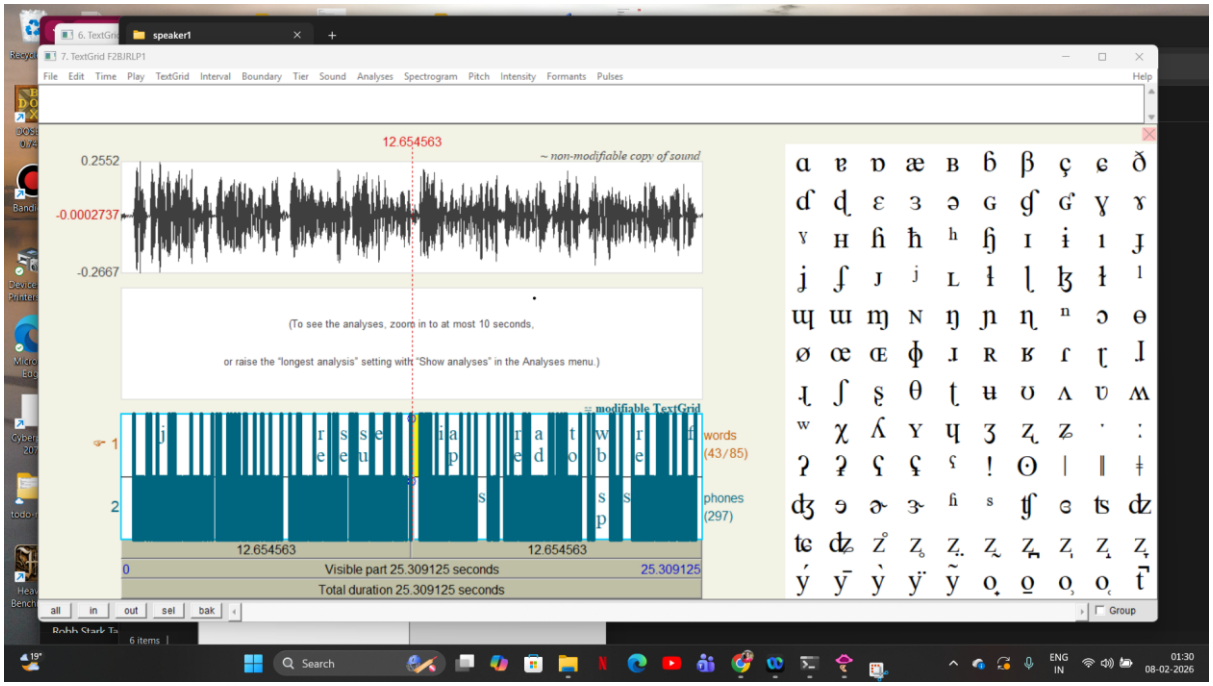
Before OOV

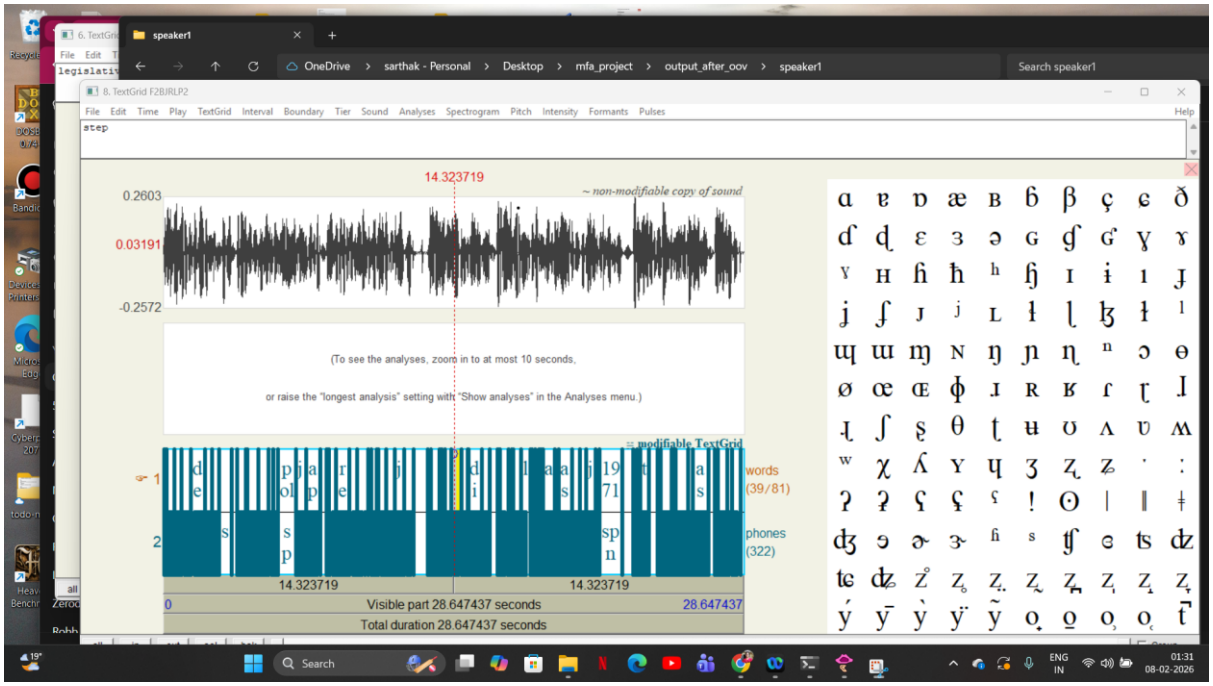


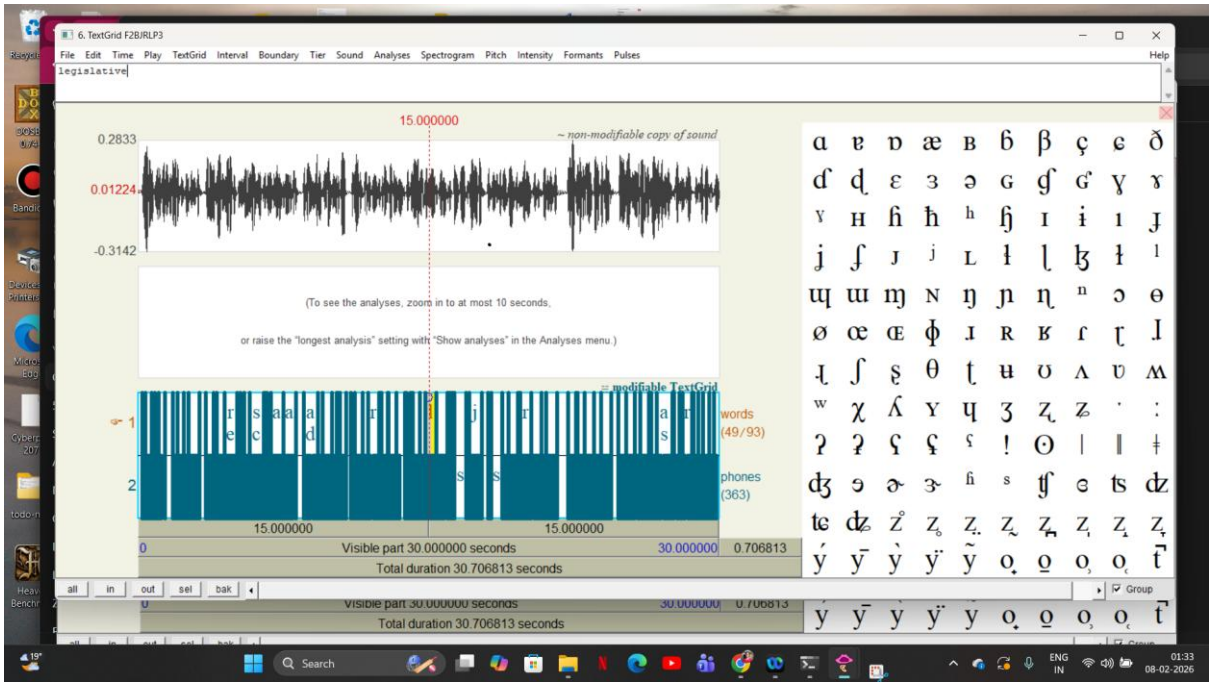

4. TextGrid F2BJRLP2
File Edit Time Play TextGrid Interval Boundary Tier Sound Analyses Spectrogram Pitch Intensity Formants Pulses Help
28.647437
~ non-modifiable copy of sound
0.2603
-0.001862
-0.2572
(To see the analyses, zoom in to at most 10 seconds,
or raise the "longest analysis" setting with "Show analyses" in the Analyses menu.)
modifiable TextGrid
words
(81/81)
phones
(322)
Visible part 28.647437 seconds
Total duration 28.647437 seconds
all in out sel bak
Group


6. TextGrid F2BJRLP3
File Edit Time Play TextGrid Interval Boundary Tier Sound Analyses Spectrogram Pitch Intensity Formants Pulses Help
legislative
15.000000
~ non-modifiable copy of sound
0.2833
0.01224
-0.3142
(To see the analyses, zoom in to at most 10 seconds,
or raise the "longest analysis" setting with "Show analyses" in the Analyses menu.)
modifiable TextGrid
words
(49/93)
phones
(363)
15.000000
15.000000
Visible part 30.000000 seconds
30.000000
0.706813
Total duration 30.706813 seconds
all in out sel bak
Group

After OOV


speaker1
7. TextGrid F2BJRLP1
File Edit Time Play TextGrid Interval Boundary Tier Sound Analyses Spectrogram Pitch Intensity Formants Pulses Help
12.654563
~ non-modifiable copy of sound
0.2552
-0.0002737
-0.2667
(To see the analyses, zoom in to at most 10 seconds,
or raise the "longest analysis" setting with "Show analyses" in the Analyses menu.)
modifiable TextGrid
words
(43/85)
phones
(297)
12.654563
12.654563
Visible part 25.309125 seconds
Total duration 25.309125 seconds
all
in
out
sel
bak
Group


speaker1
OneDrive › sarthak - Personal › Desktop › mfa_project › output_after_oov › speaker1
Search speaker1
8. TextGrid F2BJRLP2
File Edit Time Play TextGrid Interval Boundary Tier Sound Analyses Spectrogram Pitch Intensity Formants Pulses Help
step
14.323719
~ non-modifiable copy of sound
0.2603
0.03191
-0.2572
(To see the analyses, zoom in to at most 10 seconds,
or raise the "longest analysis" setting with "Show analyses" in the Analyses menu.)
modifiable TextGrid
words
(39/81)
phones
(322)
14.323719
14.323719
Visible part 28.647437 seconds
Total duration 28.647437 seconds


6. TextGrid F2BJRLP3
File Edit Time Play TextGrid Interval Boundary Tier Sound Analyses Spectrogram Pitch Intensity Formants Pulses Help
legislative
15.000000
0.2833
0.01224
-0.3142
~ non-modifiable copy of sound
(To see the analyses, zoom in to at most 10 seconds,
or raise the "longest analysis" setting with "Show analyses" in the Analyses menu.)
= modifiable TextGrid
words
(49/93)
phones
(363)
15.000000
15.000000
0
Visible part 30.000000 seconds
30.000000
0.706813
Total duration 30.706813 seconds
all in out sel bak
Group
Search
ENG
IN
01:33
08-02-2026