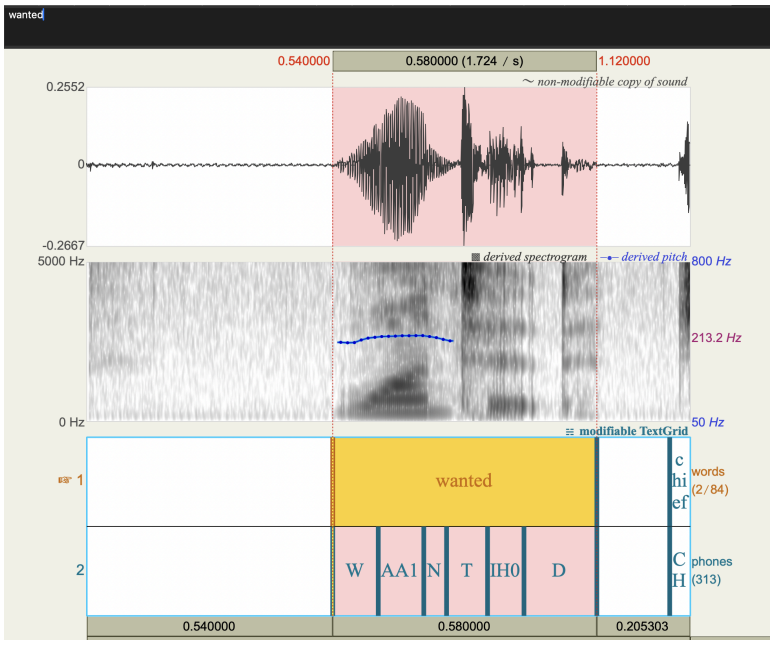
Report

Model/dictionary Used: english_us_arpa
Trained a custom dictionary, to include OOV words
Sample visualization:

1. File name: **F2BJRLP1**



Here we can see, the alignment at both word level and phoneme level.

Key Observation:

1. During alignment visualization in Praat, a few short regions (e.g., breathing sounds, “shh”, “ahh”, or other non-speech noises) were not labeled with phonemes. These segments typically correspond to background or paralinguistic sounds not present in the transcript.
2. The main speech portions were correctly aligned, but these untranscribed noises were ignored by the aligner, which is expected behavior since the dictionary does not contain phonetic representations for such sounds.
3. The alignment for most utterances was accurate at both the word and phoneme levels.
4. In a few cases, a slight **timing mismatch** was observed, for example, one instance where the **word boundary offset was around 0.004s**, while the **phoneme-level offset was about 0.007s**. This was resolved when a g2p model was trained, on the speech

corpus. After training the new alignment is giving good output.