Directions for speech tagging:

There are 36 Arabic target words (see arabic-word-list.xlsx) per speaker, or 18 word pairs. The word pairs are minimal pairs that differ in either vowel length or consonant length, so it is important to carefully transcribe the correct word for each speech file.

Each .wav file is named with the Arabic word in arabizi transliteration (from the column named: “word-arabizi” with two numbers afterwards corresponding to the trial number and screen number from the Gorilla recording task. In other words, the naming convention is [word]-[trial number]-1.wav You don’t need to worry about the trial or screen numbers when it comes to tagging the words, this is just so you know what the file name means.

Within each recording, the speaker says the word in a carrier phrase, which is repeated three times. Here is the Arabic carrier phrase.

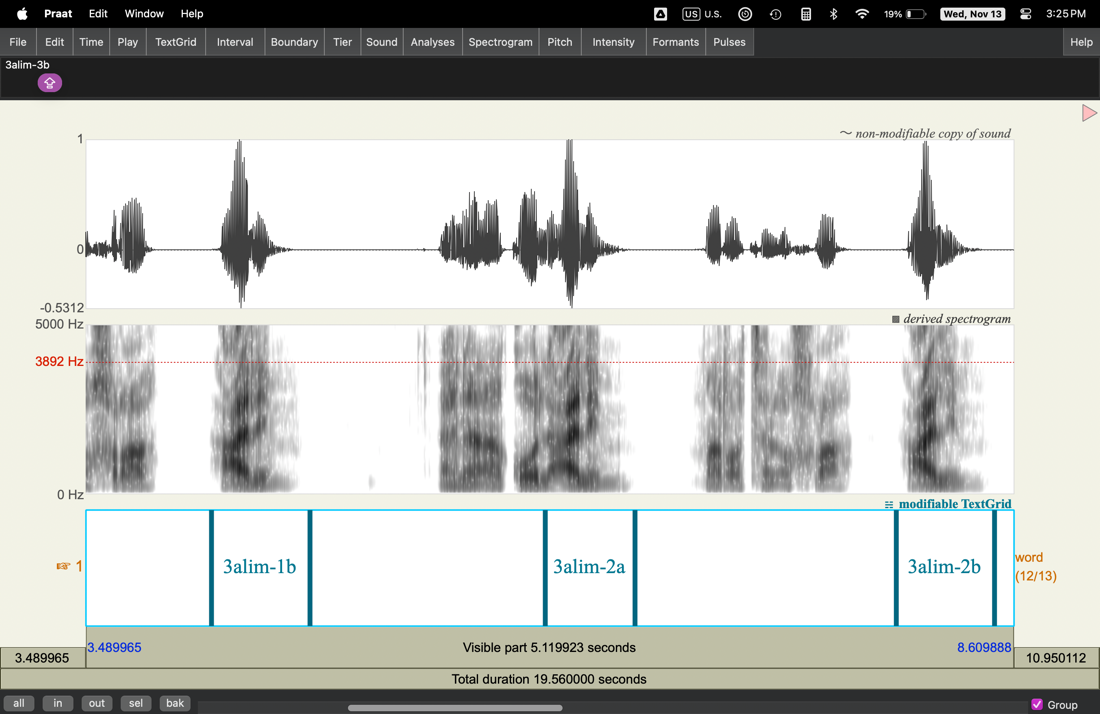
أنا أقول(...) مرة أخرى. (...).

*anaa aquulu (...) marrata ukhraa. (...).*

I say [...] once again. [...].

(((I say [a-position] once again. [b-position].)))

Your job is to tag each utterance of the target word using a Praat text grid. Create one textgrid interval tier named “word.” You don’t need to tag any of the words of the surrounding carrier phrase. Each time the word is said, it will be tagged with the repetition and the position. In other words, the target word in the first position (“I say [...]”) during the first repetition will be labeled as “[word]-1a” and the target word in the second position (“again. [...]”) will be labeled as “[word]-1b,” and then so on for the second and third repetitions. See the screenshot below for an example.



(This audio file and text grid is included in the zip file as a Praat collection “sample-tagging” to show you kind of what I’m looking for.)

Make sure to label the word using the arabizi transliteration (should be the same as the filename, but you can also cross reference with excel file).

After you tag all of the words in a file, save the textgrid with the same filename as the wav file (this should be automatic, but I wanted to make this clear just in case). After saving all of the textgrid files, send me a zip file of just the textgrid files, no need to save and resend the wav files. I will show you how I use a script to cut up the recordings into individual audio files that are saved with the textgrid label as the filename, and from there, I will choose one of the six repetitions to be used in the discrimination task.

Some notes:

* The words uttered in the “a” position might have some coarticulation with the previous word “aquulu” and the following word “marrata.” Do your best to just capture the word, but don’t concern yourself too much, as I likely won’t choose any of those. I am almost certainly going to pick from the “b” position words.
* For the “b” position words, it’s okay to leave some silence around the word, but try to leave out any surrounding breaths or movements. If there is overlapping movement that you cannot avoid, that’s okay, I will just not choose that repetition.
  + Though keep in mind, there are some speech sounds in Arabic that might sound like breaths in to English speakers. For example, in the words ru7 and ruu7, the “7” (arabizi transliteration) is a sound that doesn’t exist in English and kinda sounds like a forced out breath (not sure how else to describe). Basically, try to avoid including breaths in the label that sound like they’re in preparation for the next utterance (hopefully the spectrogram can help distinguish), but let me know if you are ever unsure.
* If the speaker pronounces the word differently than expected, that’s okay, label it anyway with the same spelling as the others.
* I told the speaker that if they mess up or anything, they can restart the sentence and go from there, so if there are more than three repetitions, that’s okay, continue following the labeling conventions (i.e., if they messed up a sentence and end up with four repetitions, label the target words in the fourth repetition as “4a” and “4b”).
  + If the mess up caused the one of the positions to be missing, still follow the labeling conventions by repetition and position (e.g., if they mess up on the first sentence and don’t say the word again in the b-position, then there will only be “[word]-1a” followed by “[word]-2a” and there will be no “[word]-1b”). The position label is just as important as the repetition label so that I know whether to expect coarticulation or not.
* If you’re unsure of anything, feel free to ask me!