

## Annotation Guidelines for

## An Outlier Analysis of Vowel Formants from a Corpus Phonetics Pipeline

Ahn, Levow, Wright, & Chodroff. 2023. INTERSPEECH.

## Setup

1. Download
  - a. Praat
  - b. Praat script:  
    check\_wild\_formants.praat
  - c. TextGrid folder (1 per lang):  
    TextGrids/{lang}\_errors\_100/
  - d. corresponding audio folder (1 per lang):  
    Audio/{lang}\_errors\_100/
  - e. raw CSV file (1 per lang):  
    Annotations/{lang}\_errors\_100.csv
2. Change path names in Praat script: check\_wild\_formants.praat
  - a. Make the 4 variables at the top point to where downloaded files are located on your machine
3. Run Praat script either by:
  - a. Double-click on praat file from File Manager (e.g. “Finder” on Mac)
  - b. Open Praat. At the top, Praat > Open Praat Script > (select praat file)
    - i. Click “Run”
4. Record your annotations on Google Sheets
  - a. Find your sheet under:  
    Annotations > {name} > errors\_100\_{name}
  - b. Select a category for each row
  - c. If relevant, use “Describe” column to clarify category info
  - d. Use “Notes” column for your own thoughts/hunches/etc
    - i. Can also mark here if you’re not sure of this annotation!
5. Misc
  - a. If a phone (IPA symbol) doesn’t look like IPA, my conversion script (XSAMPA → IPA) must have done weird things.
    - i. Type in that symbol to this [site](#) to see what the IPA is.

# Categories

## 1. Alignment

### a. Target Overlap

- i. The MIDPOINT of the window is not capturing the targeted vowel
- ii. Boundaries for that vowel includes vowel but also includes too much more (boundary too wide) or didn't include full vowel (boundary too narrow)

### b. Broad

- i. Anything beyond Target Overlap

## 2. Transcript

### a. Extra sounds

- i. Extra phones/syllables/words were spoken but not transcribed
- ii. IF only a few phones, try to write them in IPA under "Describe"

### b. Extra transcript

- i. Extra phones/syllables/words were transcribed but not spoken
- ii. IF only a few phones, try to write them in IPA under "Describe"
- iii. \*this won't be used if only the target vowel is not uttered
  - 1. IF there were extra preceding/following phones transcribed but not spoken

### c. Broad

- i. Phone sequence just doesn't match the audio.
  - 1. Good way to check this is by going to the start of the recording and seeing if the first handful of syllables are even matching

## 3. Linguistic

### a. Deletion

- i. Surrounding phones are present, but only Target vowel is absent

### b. Change

- i. A different vowel is pronounced.
- ii. Try to estimate what the other vowel is under the "Describe" column
- iii. If you hypothesize that a neighboring phoneme is affecting this vowel (i.e. Coarticulation), you can mention that in "Coarticulation"

## 4. Formant

- a. The measured formant value does not reflect the frequency of the relevant energy band in the vowel.

## 5. Fine

- a. The boundaries are good, it sounds like the transcribed vowel; everything's fine! But somehow the formants are outlying

## 6. Extra Annotations (~hypotheses)

- a. [These are NOT direct error categories, but useful observations and hunches. To be noted IN ADDITION to previous categories]
  - i. We can add to this as we go
- b. Audio [you can select multiple]
  - i. Background Music
  - ii. Background Other
  - iii. Poor Audio (Quality)
    - 1. No background sounds per se, but the recording quality is really bad...
      - a. E.g. Hard to see spectrogram
    - 2. Might see these more in the CommonVoice data
- c. Coarticulation
  - i. Write in following format:
    - 1. E.g. “prec:[s]” if preceding [s] phoneme is affecting target
    - 2. E.g. “foll:[j]” if following [j] phoneme is affecting target
- d. High Pitch Voice
  - i. YES if speaker has high pitch