**TO DO:**

* **Add comments to all code for part 1**
* **Update descriptions here**

**------------------------------------------------------------------------------------------------------------------------------------------**

**Folder structure**

**grids\_phrase\_corrected\_not\_FAVE\_ready** : handsegmented .TextGrid from experiment 1

**FAVE\_formatted**: folder, contains *folders* with output .TextGrid files from adapting\_toFave\_TextGrids\_formant.m

**FAVE\_formatted/FAVE\_giver:** folder, output .TextGrid for giver only.

**FAVE\_formatted/FAVE\_rcver:** folder, output .TextGrid for receiver only.

**tables\_generated\_by\_Sam** : data tables generated by Sam, downloaded from the lab’s server.

**Part 1. Formant reading and main tables preparation**

1 - Checking for typos in segmentation and adapting .TextGrids to FAVE format:

**adapting\_toFave\_TextGrids\_formant.m** : reads .TextGrids as the result from handsegmentation, and turns them into a format expected by FAVE.

Adds the subjects judgment on each other.

*folders*

**grids\_phrase\_corrected\_not\_FAVE\_ready**: folder, handsegmented .TextGrid from experiment 1.

**FAVE\_formatted**: folder, contains *folders* with output .TextGrid files from adapting\_toFave\_TextGrids\_formant.m

**FAVE\_formatted/FAVE\_giver:** folder, output .TextGrid for giver only.

**FAVE\_formatted/FAVE\_rcver:** folder, output .TextGrid for receiver only.

*Functions*

ST\_read\_praat\_textgrid.m

ST\_write\_praat\_textgrid.m

FAVE\_modify\_text.m

Dictionary.m

**subject\_data\_entering.m** : aid in entering the questionnaires responses directly to a table.

2 - Use FAVE (in command line) to read all .TextGrid files with their associated audio, and produce output files with formant and duration measurements.

Instructions for reading the text grids and audio files automatically in a batch from the command line.

Modify file names as appropriate.

# There are three folders:

- audio: contains audio files .wav

- grid: contains textgrid files .TextGrid

- output\_files: empty folder ready for output

\*\* Important \*\*

- names of files in audio and textgrid must be the SAME up to the extension.

- run in terminal with Python **2**.

grid="../textgrids/FAVE\_formatted/FAVE\_giver"

ls $grid

audio="../audio/audio\_giver"

ls $audio

mkdir output\_files\_giver

output="output\_files\_giver"

for file in $grid/\*.TextGrid ; do name=$(basename -s .TextGrid "$file") ; python bin/extractFormants.py --config=config.txt "$audio/$name.wav" "$grid/$name.TextGrid" "$output/$name.output.txt" ; done

3 – Putting the content of FAVE output files into a Matlab table.

**read\_FAVE\_results.m** : reads .txt from FAVE-extract\_Toolkit/output\_files both \*output.txt and \*output\_norm.txt. Creates a table for raw and one table for normalized readings. Tables are saved as VOW.mat and VOW\_norm.mat in d/tables/ .

* Adds and fills columns: player, role, session, day, week, team, vow\_id, to each table.
* Adds and fills column with any labels (repeated, disf, etc).

*Functions*

grid\_annotations.m:

*add\_timers.m*

*outliers.m*

**add\_timers.m**: adds columns day\_timer, week\_timer, and all\_timer. Called from within

read\_FAVE\_results.m

**outliers.m**: loads VOW\_norm.mat from d/tables/ . and determines outliers. Creates is\_outlier column

with 0 for good data points and 1 for outliers. Saves VOW\_norm.mat to d/tables/ .

*Only* for VOW\_norm.mat and VOW\_norm.csv. Not done on nonnormalized table. Prints out

summary of % of outliers by instance, with total of data and amount on data that is outliers.

Called from within read\_FAVE\_results.m.

* Adds syll[able] column.
* Adds is\_outlier column.

Outlier exclusion criteria: tokens where the F1 or F2 was more than three standard deviations away from the mean. This was based on the group mean for each formant for each word. In disyllabic words, the procedure was applied based on the syllable to which the vowel belonged.

**Part 2. Data analysis and further tables preparation**

**Teammates.m**: get table created by *adapting\_toFave\_TextGrids\_formant.m* and for each subject, add columns F1\_teammate, F2\_teammate, duration\_teammate, with those values from the corresponding teammate.

Since teammates are out of step in giver round number, it is necessary to interpolate.

Add column that signals whether that data point was interpolated or original.

The original table used has the average by each round the speaker used.

* Table with interpolated rounds
* Table with matched rounds ([1 3 4 7; 2 5 6 8] for example)

Use tables from Tableau from moving average of week\_timer, day\_timer, and all\_timer.

To create difference between teammates, and between self at week 1 and week 2.

Match rows if each teammate has same number, if not, take randomly from the one with more points so as to make both teammates have the same number of rows.

**word\_table.m**: for each word, which words precede it and which follow it.

**Teammates\_1.m**

**Expected\_distribution\_of\_autocorrelation.m**:

Objective: see what happens when you shift time-step by one and correlate it with itself.

Expected: a given time-step is correlated with the previous time-step.

Fundamental time-series analysis, autocorrelation function of lag-1. If the value of the correlation is positive, it means that it is greater than chance that each sample is not independent of each other.

**Tables:**

Round\_averages\_both.csv : uses the average of each round for /ae/ for both cad and cat, and /iy/ for both qeed and qeet. Copied from Tableau.

Round\_averages\_d.csv: uses the average of each round for /ae/ cad, and /iy/ for qeed . Copied from Tableau.

Round\_averages\_t.csv: uses the average of each round for /ae/ cat, and /iy/ for qeet . Copied from Tableau.

**Moving\_avg\_weektimer\_5\_both**: uses moving average of week\_timer for /ae/ for both cad and cat, and /iy/ for both qeed and qeet. Copied from Tableau.

**Moving\_avg\_weektimer\_5\_d**: uses moving average of week\_timer for /ae/ cad, and /iy/ for qeed . Copied from Tableau.

**Moving\_avg\_weektimer\_5\_t**: uses moving average of week\_timer for /ae/ cat, and /iy/ for qeet. Copied from Tableau.

**Moving\_avg\_daytimer\_5\_both**: uses moving average of day\_timer for /ae/ for both cad and cat, and /iy/ for both qeed and qeet. Copied from Tableau.

**Moving\_avg\_daytimer\_5\_d**: uses moving average of day\_timer for /ae/ cad, and /iy/ for qeed . Copied from Tableau.

**Moving\_avg\_daytimer\_5\_t**: uses moving average of day\_timer for /ae/ cat, and /iy/ for qeet. Copied from Tableau.

Moving\_avg\_alltimer\_5\_both: uses moving average of all\_timer for /ae/ for both cad and cat, and /iy/ for both qeed and qeet. Copied from Tableau.

Moving\_avg\_alltimer\_5\_d: uses moving average of all\_timer for /ae/ cad, and /iy/ for qeed . Copied from Tableau.

Moving\_avg\_alltimer\_5\_t: uses moving average of all\_timer for /ae/ cat, and /iy/ for qeet. Copied from Tableau.

**teammates\_moving\_avg\_weektimer\_5\_both**: result of teammates\_1.m, table moving\_avg\_weektimer\_5\_both but with corresponding teammate information added.