Wordvariability Project Behavior data Workflow

1- For each subject, you will find the following files (take subject RS for example here):

DSP_OUT_2020-08-20,10;10;17.wav: audio record of the response in the scanner;

RS_RESPONSE_cohort1.xlsx: spreadsheet for transcribing the audio response;

RS_Phoneme.R: script for extracting phonemes from target words & responses, and to calculate the accuracy in percentage;

cmudict.R, phoneme_mapping.R: scripts called by the RS_Phoneme.R script to help extracting the phonemes;

cmudict-0.7b_no_header.txt: phoneme bank to source from for the phoneme extraction.

2- First, transcribe the audio clip, and write down the single word response into the spreadsheet for each run. The transcription needs to be consistent and follows these rules:

a) Use lower case for all the transcription words;
b) If the subject responded more than one word, put a hyphen between the words, e.g., black-potato;
c) If the subject said "not sure", the transcription is: not-sure;
d) If the subject's response is not word, try write down similar sound/phonemes for that response. e.g., data-lizer;
e) If the response word is definitely wrong, write down 'definitely-wrong';
f) Highlight the transcribed word that meets criteria d) above, and keep a record of these "new words", as you will need them in the future steps

3- Second, copy all the 'movie_list' and 'response' into a new datasheet (named 'Sheet1') inside the same excel file. In Sheet1, there are 7 columns:

movie_list
type (An, AnV, V)
word_speaker
word
word_phonemes
response
response_phonemes

phoneme_accuracy

4- Third, start phoneme extracting R script (RS_Phoneme.R), extract the phonemes for target words and responses. Then calculate the accuracy rate. These results will be stored in a temporary file named 'temp.csv'.
Note:

a) if encountered error similar to this:
> Error in fMRIwords$response_phonemes[index] <- get_phonemes(words[index]) :
  replacement has length zero
In addition: Warning message:
In get_phonemes(words[index]) :
  word not found in dictionary: [SHOWMING] returning NULL

This means that the word 'SHOWMING' is not in the phoneme bank (cmudict-0.7b_no_header.txt), then you can manually add this word to the bank:
SHOWMING SH OW1 M IH0 NG

Then re-run the script.

VERY IMPORTANT!!! Please write down a list of the new words that you have manually added to this bank. In this list you will have two columns: word, phoneme.

b) If the response word is definitely wrong, you can just skip that word by changing the index number in the (RS_Phoneme.R) code.

5- Copy out the results from the temp.csv to RS_RESPONSE_cohort1.xlsx;

6- Calculate the mean/SD for each type;

7- Calculate the number of correct and incorrect responses for each type. And calculate the mean/sd for them. Correct is defined as > 50% of accuracy performance. Name should be like this (1 for correct, 0 for incorrect): e.g.,

An1 –An correct trials

An0 –An incorrect trials