

Maria Frey  
LING 492B  
4 March 2021

## Lab 2 Questions

### Part 2 Questions:

The s value that I found to work the best was 0.6. This gave me a correlation of 0.67. One example of a wug that had different similarity ratings and human ratings is 'ghez'; the similarity rating was 0.43 which is pretty low, and the human rating was 4.19 which is relatively high. 'Skride' also had a lower similarity rating (0.44) in comparison to its human rating (4.05). This divergence may be because humans are taking into account other factors such as how the words would actually sound when they are phonetically produced. Or they may be biased in terms of the English words that the wug is compared to, and so they might find the wug to be good even if it's not similar to many English words.

### Part 3 Questions:

The accuracy of the model on the forced choice task is 0.71. In the table below I listed all the words that the model got wrong:

WUG	Response	Prediction
drice	NULL -> t	NULL -> d
rife	NULL -> t	NULL -> d
drit	NULL	NULL -> Id
fleep	iX -> EXt	NULL -> t
gleed	i -> E	NULL -> Id
glit	NULL	NULL -> Id
queed	i -> E	NULL -> Id
skride	2 -> 5	NULL -> d
gude	NULL	NULL -> d
blafe	NULL -> t	NULL -> d
tesh	NULL -> t	NULL -> d
nold	NULL	NULL -> Id

In the case of 'drit' and 'glit' the model might be considering as neighbors words like 'admit,' or the word 'fit' which were listed in the training file as NULL -> Id. In the case of 'queed' and 'gleed', the model may be considering plead and speed which are listed as NULL -> Id. Neighbors of 'nold' may be 'fold' and 'scold' which are NULL -> Id.

A couple of problems with the model are that it seems to rely more on the regular past tense ending than humans do and that it's classifying some of the voiceless consonant endings as NULL -> d instead of NULL -> t like they should be.

### Part 4 Questions:

I was able to improve the model's accuracy from 0.71 to 0.73 by penalizing the insertion, deletion and substitution of vowels, i.e., adding one point to the cost of inserting, deleting or substituting a vowel. I thought this might help because as humans we judge verbs to be similar by prioritizing the vowels over the consonants and so it might make the model's predictions more in line with the human responses. It did work somewhat to improve the accuracy, but not by much.