

Assignment 4: Phonotactic probability, cont.

Due: Oct. 7

- Keep tweaking your program; in its final state, it should produce files that list the scores for each word, so it is relatively easy to give it a set of nonce words, and browse through the list of predicted probabilities
- It should also calculate probabilities in two different ways (one based on positional probability, and one based on biphones, in some fashion)
- Finally, I would like you to pick one of the questions raised here, and see how it affects the results, e.g.
 - Transitional probability vs cooccurrence probability
 - Different alignments
 - * Different weighting of different parts of the syllable
 - Use of type vs token frequency
 - Training on monosyllabic vs all words
- Specifically, discuss how a different way of modeling phonotactic probability affects the match to the data found in `AlbrightHayes.txt`