Ngram.py will take two arguments, a training file and a value of n (2 or 3).

To take these arguments sys module is needed

Using the training data as a base, ngram.py will produce 25 random words.

Import random?

For my prototype I am going to create some fake training doc with less words. For each phoneme, I need a probability that it follows another phoneme.

For this, a dictionary of all phonemes that are followed by another will be created. Each value will be another dictionary with all phonemes and their frequency.

Then when the program knows a probability for each phoneme following another (or beginning a word), it can produce new words.