Clarisse Taboy
May 30, 2021
Methods II
HW8 Write-Up

1. The whole assignment was quite a challenge, and if it weren't for practicum with Lucas, I would've been very lost. Most of my code is modeled on what we did in practicum, though it's also helpful that this assignment has directions that explains each step so that I am reminded of which part of the code from practicum does what and how it can apply or be adapted to this assignment. But I'll definitely need more practice/review on this topic.

2. To deal with cases when the list index was out of range for feature extraction, i.e. when the homograph is the first/penultimate/ultimate token and when the token is numeric, I wrote a separate function because my initial code became too messy/hard to follow otherwise. Since the last token is presumably always a punctuation mark, I had to change the condition from if x > len(tokens) to if x >= len(tokens), which also got rid of the error I got saying that the index was still out of range.

3. In part 3, I spent most of my time figuring out how to compute the averages. I initially used the sum() function on both 'correct' and 'size' and divided them to get the micro-averaged accuracy. Although when it came to using the mean() function for the macro-averaged accuracy, I got an error saying I needed an iterable as input, not the float that I got in the previous step. So in order to have an iterable, I looked up how to iterate over 'correct' and 'size' simultaneously to get their sum (via the zip function) and to store the accuracy for each homograph in a new list that would then serve as input for mean().