Davis Busteed – LING 360 – HW #5

After running my script that created frequency lists for the eight different registers, I looked through the eight .csv files that the script produced. I looked at the top 10 most frequent words for the registers, and saw that there were many similarities between them. For example, the word “one” was found in the top 10 of every register. This word is used in a lot of phrases, so I wasn’t too surprised that it was common across all topics. The words “like,” “would,” and “people” were also quite popular among the registers, appearing in the top 10 of a majority of the lists.

There were some frequently used words that were unique to each register. For example, the word “love” was the second most used word in the song lyrics register, while the word “also” was the second most used in the informational register. This isn’t necessarily new knowledge, but I can infer that the LY register deals more with emotions and feelings, and the IN register is more focused on delivering facts.

I also calculated the total number of different words in each frequency lists. I found that song lyrics and discussions have the lowest “diversity” of word usage, and that informational and opinion blogs have the highest. If I were to make a hypothesis about these figures, I would’ve assumed to see something like this. I don’t think a diverse vocabulary is necessary for everyday conversation, but is more likely to be used when discussing specific, non-ordinary topics.