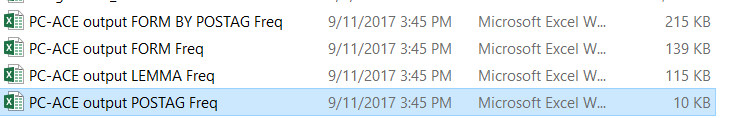
QTM 446W

HW3

Gabriel Wang

Instead of using the whole five books of *A Song of Ice and Fire* (referred as SIF), this time, I did a little bit sampling of my data. I picked 3 chapters for each POV character and did my Stanford Core NLP on it. The result is still very interesting, and the corpus statistics definitely revealed a lot about this corpus.

I used PC-ACE to produce four frequencies table. And the one I found most useful is the PC-ACE output LEMMA Freq, which gave a lot insight on George R. R. Martin’s choices of words. 

The top 43 words displayed are some common words that don’t really reflect the nature of this masterpiece. However, the 44th most common word ‘lord’ is very interesting.



Martin built the word of SIF in the Medieval Era, where the great lords played a significant role. Once I figure out how to play with the Gephi graph and produce a word distance-based co-occurrence graph, I expect to see a lot of interactions between the peasant and the highborn. But right now, I used PC-ACE to create a radar chart to visualize my KWIC table.

I was not surprised to see that most of the words close to ‘lord’ are names, as ‘lord’ is a proper title to add in font of noble’s name when the peasants address a highborn.

Similarly, this lord-peasant relationship typical to Medieval Era is also revealed from the high frequencies of words like ‘prince’, ranked 143th, ‘House’, ranked 236th, ‘command’, ranked 326th, ‘sword’, ranked 136th, etc.

Moretti and Pestre, two major founders and supporters of ‘quantitative linguistic analysis’ will applaud such route of ‘distant reading’ as their study of *‘The Language of World Bank Reports’* indicated. The word occurrence visualization of ‘lord’ closely follow the method of their study of ‘poverty’ and ‘poverty reduction’. And their approach to show that the world bank is ‘overcoming poverty and creating opportunity for people in developing countries’ is definitely worth listening to when I try to reveal the relationship between ‘lords’ and ‘peasants’,

Source:

1. Franzosi, TIP File;
2. Moretti and Pestre (2015) The Language of World Bank Reports;