

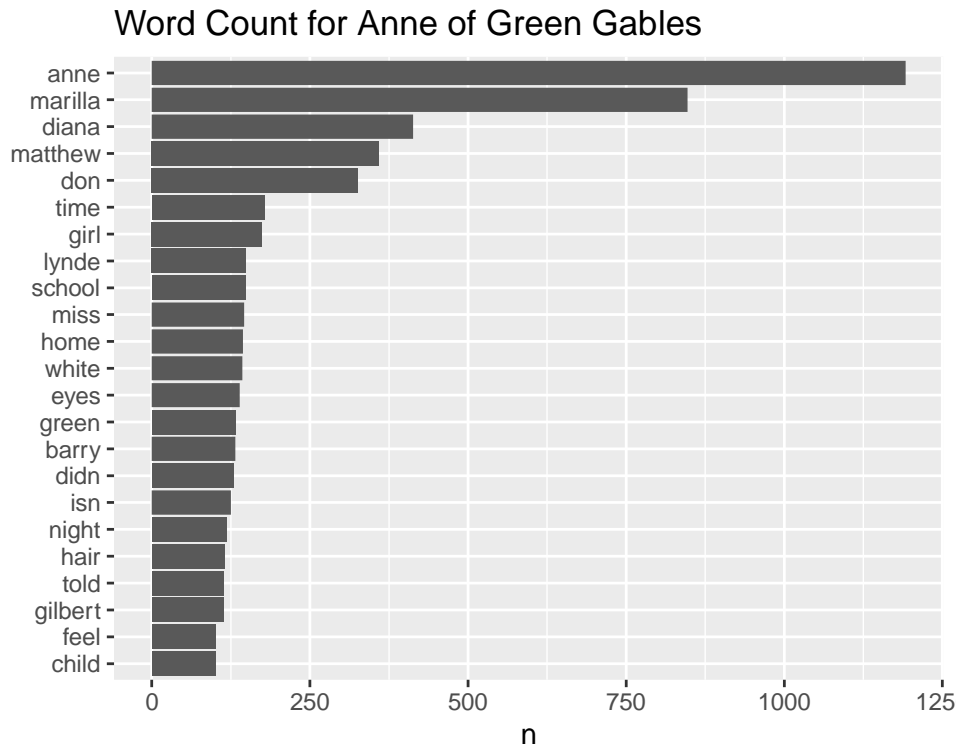
MA615: Text Analysis

Yongrong Chai

11/29/2021

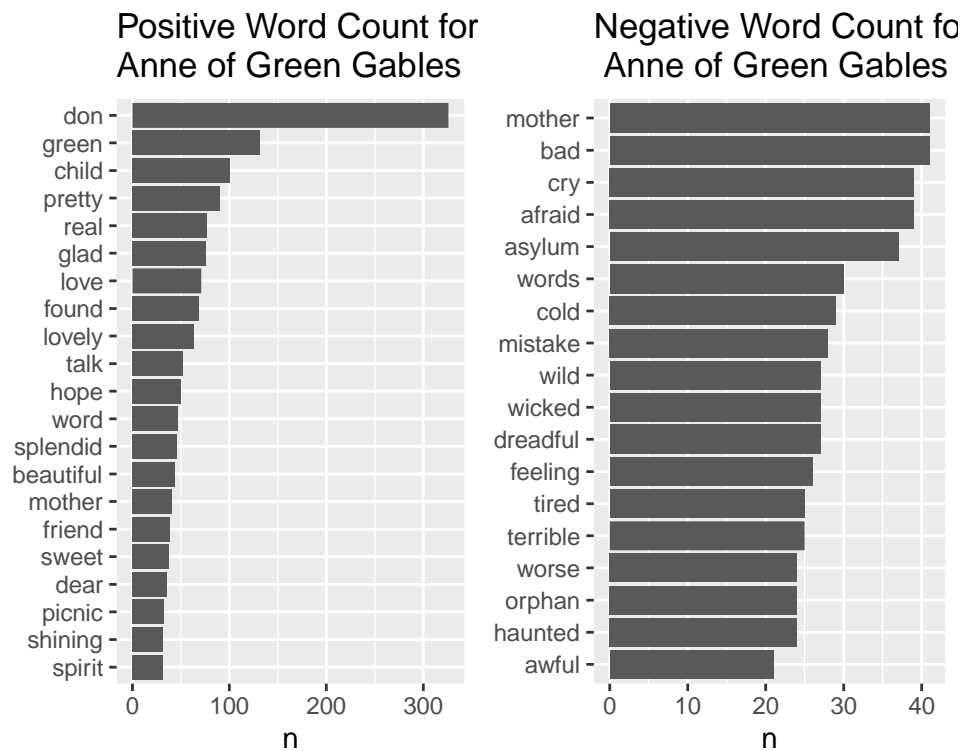
This is a sentiment analysis of Anne of Green Gables by Lucy Maud Montgomery.

Overall Word Count



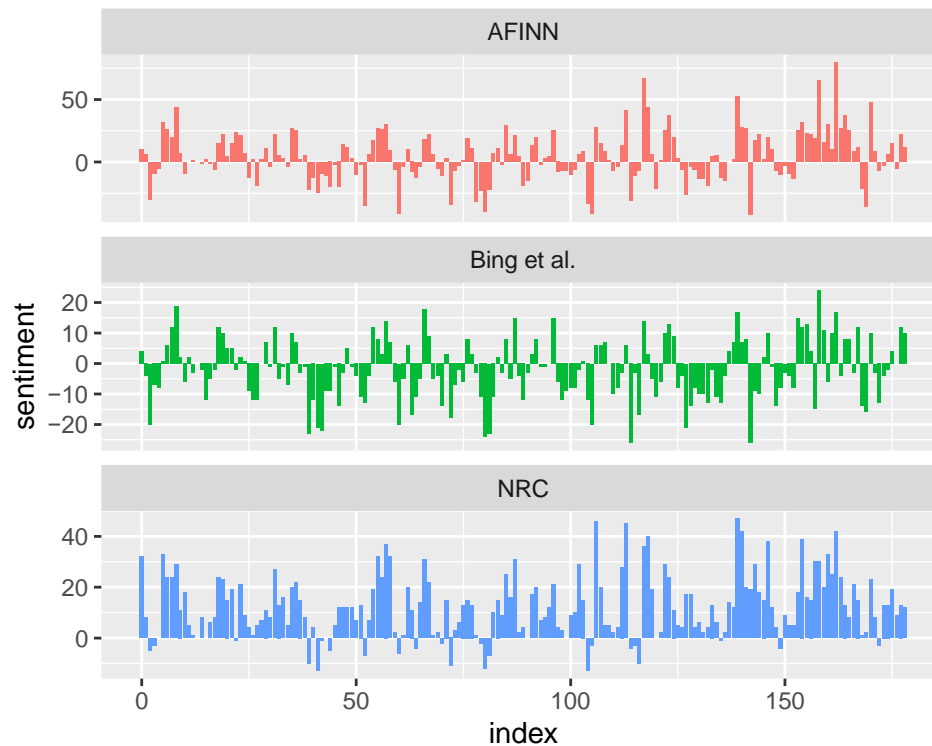
Top 4 words in book are Anne: protagonist is Anne Marilla: Although she is not the protagonist's biological mother, she is like Anne's mother. Diana: Anne's best friend Matthew: Marilla's brother, he is like Anne's father.

Positive word count V.S. negative word count



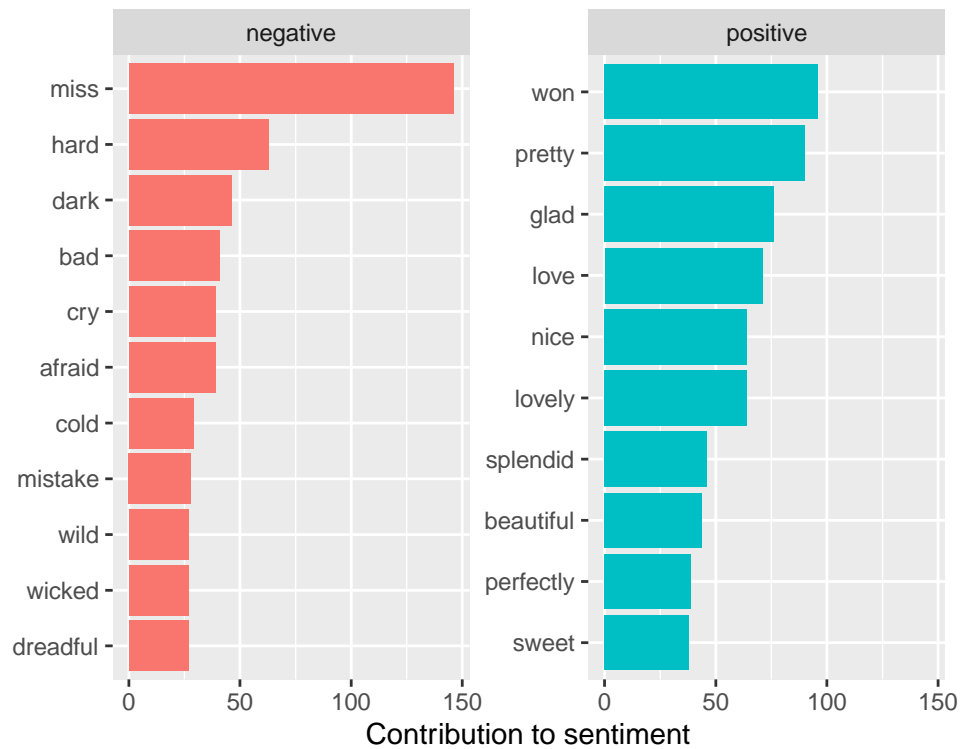
The main plot of this book: Anne is an orphan who is adopted by Matthew and Marilla and grows up on their farm, Green Gables. Positive words here show the joy of growing up as a little girl. Cold words here shows Anne, as an orphan, just arrived in a new environment of loneliness, exclusion, re-establish the world view of helplessness at the beginning.

Comparing the three sentiment dictionaries



Here are three different lexicons for sentiment analysis. We can compare the proportion of positive and negative words between them. As you can see, the last one has more the least negative words.

Most common positive and negative words



Word Clouds

[illegible]

negative



positive

Positive words are obviously much larger than negative words. As a children book, the whole book is about family, friendship and childhood, which is why there are more positive words.

Reference:

<https://www.r-bloggers.com/2020/04/sentiment-analysis-in-r-with-sentimentr-that-handles-negation-valence-shifters/>

<https://www.gutenberg.org/ebooks/45>

https://learn.bu.edu/ultra/courses/_80585_1/cl/outline

<https://www.tidytextmining.com/sentiment.html>