

Shanshan Liu

s5336651

Dr.Federico Pianzaola & Dr.Lucas van der Deijl

Shakespeare Authorship problem

Introduction

Mainstream Shakespeare scholars have no debate on his identity, the William Shakespeare of Stratford-upon-Avon, but since the 16th century there has been a voice that Shakespeare was a pseudonym. The Shakespeare authorship debate is the contention that works attributed to WShakespeare were written by someone other than Shakespeare. Whether there was “true author” behind this literary master has always been one of the focal points of debate in the literary world, and many people do not believe that an English squire could write many monumental works. As early as centuries ago, questions were raised because Shakespeare was a world-renowned playwright who had many works celebrated by future generations, but there are still many unknown aspects of his life, and besides, he did not personally leave such texts. Thus, there is very little material about Shakespeare's life, which makes Shakespeare a veil of mystery.

There are many hypotheses of this question: some support the true author as Fancis Bacon, Christopher Marlowe, Edward de Vere, illegitimate son of Queen Victoria, even Queen Victoria herself.

The Bacon hypothesis of Shakespeare authorship asserts that the plays attributed to William Shakespeare were written by Sir Francis Bacon, a philosopher, writer, and physicist. Various theories have been advanced for this supposed deception, the most popular being that Bacon's climb to high rank would have been hampered if it became known that he authored plays for the public theater. The theory was first put forth in the mid-nineteenth century, based on perceived correspondences between the philosophical ideas found in Bacon's writings and the works of Shakespeare

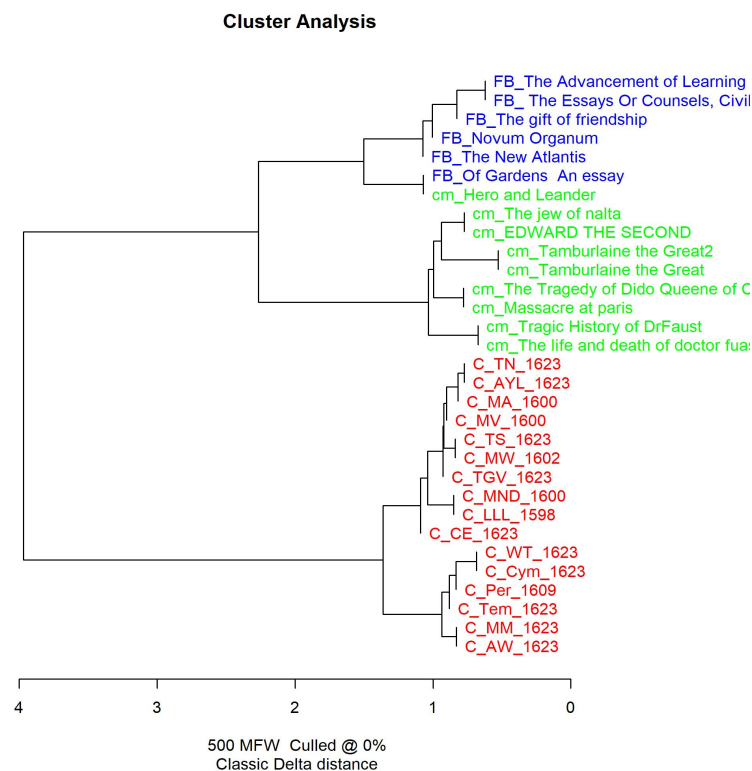
The Marlowe theory: In 1819, an anonymous writer for The Monthly Review, or Literary Journal, argued that 'Christopher Marlowe' was a pseudonym used by Shakespeare for a while. In the late nineteenth century, scholars studying contemporaneous playwrights: Christopher Marlowe and Shakespeare.They found a subtle consistency in their diction, style, and point of view, and came up with the hypothesis that the two were the same person.

The De Vere theory: as earl of Oxford, many of Shakespeare's plays resemble the life of de Vere, who was well educated and social enough to write at length about the classics, the law, foreign countries, and languages, and who traveled frequently,

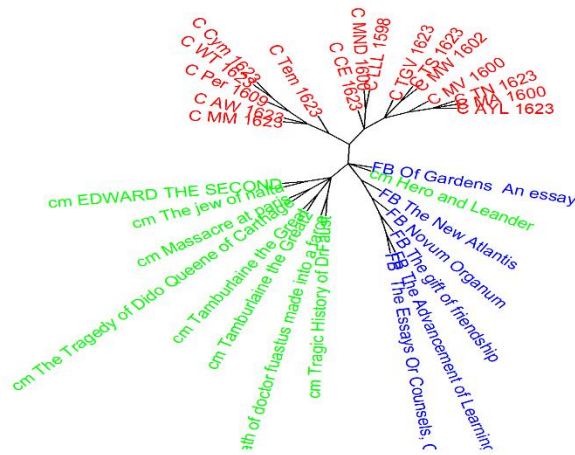
spending most of 1575 in Italy, while 14 of Shakespeare's plays have Italian settings, and seem more convincing than an English country squire.

All the evidence suggests that the most suspicious person may be among these three. I decided to download all the books contributed by the four writers(including Shakespeare) in the format of text from the Gutenberg project. However, it was difficult to find anything of Edward de Vere except the name, so I changed it to compare the three authors by stylo with 31 Shakespeare works, 9 Marlowe works and 6 Bacon works.

The texts of Christopher Marlowe, Francis Bacon, and Shakespeare's works are compared and analyzed by using stylo, where the text beginning with 'cm_' is Christopher Marlowe's work, the text beginning with 'FB_' is Francis Bacon's work, and the text beginning with 'C_' is Shakespeare's work. The text beginning with 'FB_' is the work of Francis Bacon, and the text beginning with 'C_' is the work of Shakespeare. All the text is read with the load.corpus function, and then the xml data is processed with the delete.markup function. Then use the txt.to.words.ext function to split the text sentence into words and convert them all to lowercase. The delete.stop.words function is then used to remove deactivated words to avoid the impact of common words on the text analysis. The 3000 high-frequency words are extracted using the make.frequency.list function. Based on the previous processing, the following graph was generated using cluster analysis, setting the minimum value of MFW to 300 and the maximum value to 1000. From the graph, it can be seen that Francis Bacon and Christopher Marlowe's works are close, and the words used in Christopher Marlowe's Hero and Leander's works are similar to Francis Bacon's Of Gardens An Essay has a higher degree of wording similarity.

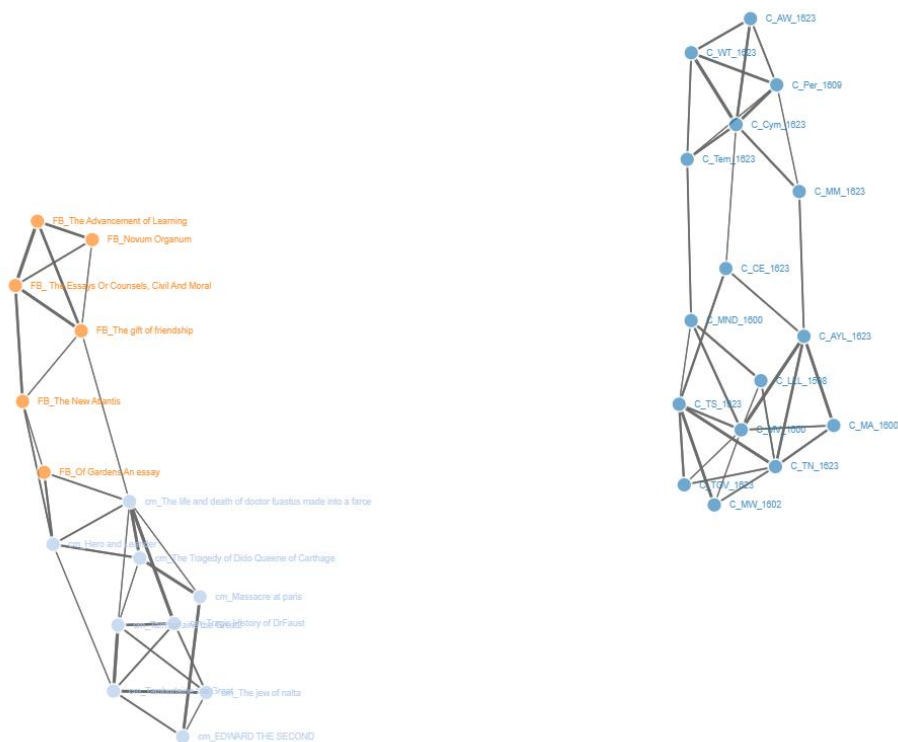


Bootstrap Consensus Tree

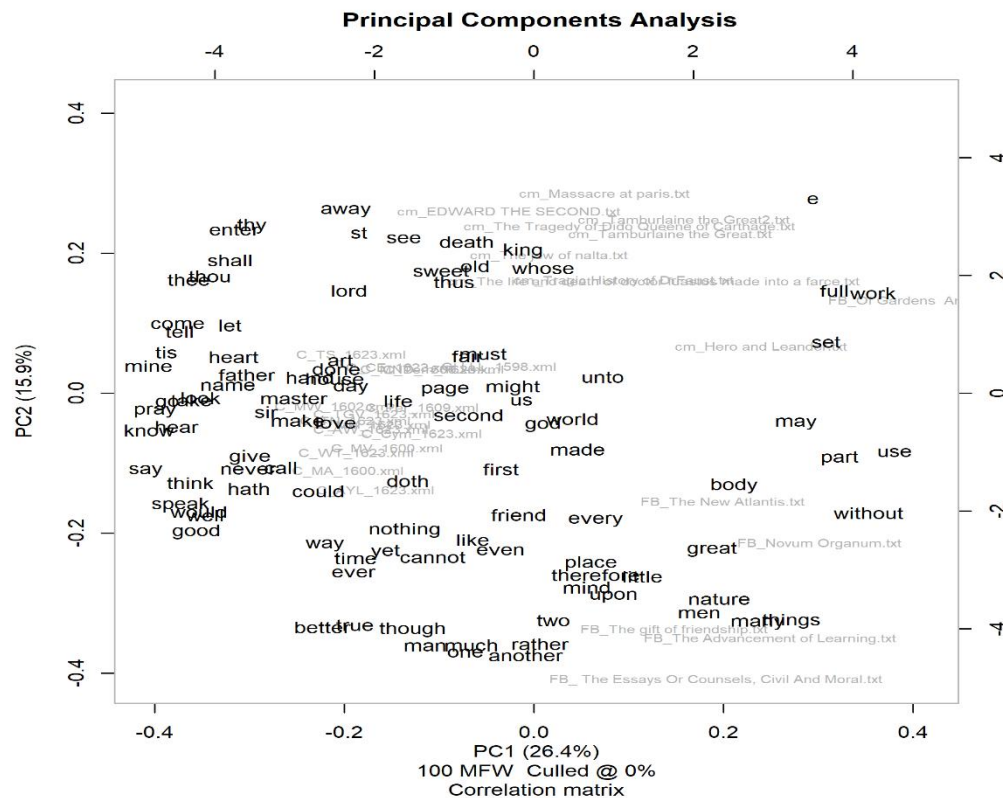


Changing the minimum value of MFW to 100 and the maximum value to 500, the tree diagram shown in the figure below was obtained. From the figure, we can see that the works of the three authors belong to three branches, Hero and Leander and Of Gardens An essay are closer, and their clustering effect is basically consistent with the previous clustering analysis results.

The degree of relevance of each text is represented as a network graph using the stylo.network function, setting both the minimum and maximum MFW values to 300, and obtaining the results shown below. Once again, it is confirmed that the works of Francis Bacon and Christopher Marlowe are highly correlated and distant from Shakespeare's works.



The high-frequency words commonly used by each author were analyzed by the method of Principal Components Analysis. The high frequency words used by Francis Bacon are body, great, nature, men, another, rather, etc. The high frequency words for Christopher Marlowe are full, work, set, king, sweet, etc. and for Shakespeare they are art, master, think, life, love, father, etc.



Conclusion

Unfortunately, I have not demonstrated through stylo's genre analysis of the text that Bacon or Marlowe may be the true identity of Shakespeare. The limitation of Stylo is that it fails to elaborate the exact the similarity and different stands , it can only provide a rough comparison among texts without a more solid method such as Burrows Delta. However, the interesting thing is that I found the amazing similarity in text between Marlowe and Bacon, it may be the reason why they are both the most suspicious candidates for Shakespeare authorship question and it awaits for further exploration.