Opposition and Change in *Othello*

Nathan Bickel

1/28/2022

Introduction

In Shakespeare's *Othello*, one prevalent aspect throughout the play is that so many of Othello's character traits and relationships are flipped on their head in such a short time. In the course of three days, Othello goes from being in love with Desdemona to murdering her, from choosing Cassio as his lieutenant over Iago to letting Iago convince him Cassio should be killed, and ultimately from trusting his own abilities and judgement to killing himself. This report aims to see if this pattern of change from one extreme to another is reflected in the distribution of words throughout the play. We will first investigate which words appear more often in *Othello* than others of Shakespeare's plays, and then explore the connection between these words and how they relate to Othello's dramatic character arc.

Data

Three different matrices will be used as the foundation for this analysis. The first is P, a 26952×37 matrix where $P_{i,j}$ represents the frequency a given word i occurs in a given play j. The second is S, a 26952×31066 matrix where $S_{i,j}$ represents the frequency a given word i occurs in a given speech j. The last is C, a 26952×1449 matrix where $C_{i,j}$ represents the frequency a given word i is said by a given character j. Additionally, S and C have corresponding matrices matrices $\mathbf{speech\_metadata}$ and $\mathbf{character\_metadata}$ that contain information about each of the speeches in S and characters in C.

The first way this data were combined was creating a vector $\mathbf{words}$ that contains all the words w in P for which $P_{w,*}$ takes its maximum in $P_{*,Othello}$ and for which this maximum is greater than 10. It initially contained 56 elements, and for reasons discussed in the analysis section, 27 of these elements were then removed. 12 of the elements from this list were then hand-picked to create a 2×6 matrix called $\mathbf{opposites}$, where each column represents a pair of words from $\mathbf{words}$. It is structured so that $\mathbf{opposites}_{i,j}$ refers to a word w that is the i^{th} part of the j^{th} pair of opposites, and will refer to the frequency at $P_{w,Othello}$.

Looking at these pairs more closely, a 2×5 matrix called $\mathbf{appearances}$ was defined for a given pair a, b from $\mathbf{appearances}$. The matrix has 5 columns because *Othello* has 5 acts, so $\mathbf{appearances}_{i,j}$ represents the number of times the word i (with 1 corresponding to a and 2 corresponding to b) appears in Act j. This information is derived from S.

Finally, for reasons explained in the analysis section, another word was picked from $\mathbf{words}$: "handkerchief". Two vectors, $\mathbf{handkerchief\_hits}$ and $\mathbf{handkerchief\_names}$, both of length 6, were created. $\mathbf{handkerchief\_hits}$ contains the character ID from C of every character in Othello that says the word "handkerchief", and $\mathbf{handkerchief\_name}[i]$ contains the name of the character at $\mathbf{handkerchief\_hits}[i]$.

Analysis

When looking for the themes that Shakespeare was trying to communicate in a play such as *Othello*, a reasonable place to begin is to consider the words that he used in the play more than any of the other 36. However, given that there are almost 27,000 words used throughout Shakespeare's work, there are sure to be hundreds in each play that are used the most, so it also makes sense to have a threshold where the word will not be included if it occurs too few times.

To this end, **words** was created to find the words that occur the most in *Othello*, and as expected, without a threshold it had too many elements to closely analyze at 604. Many of these words are used only a couple of times, and even though they technically occur more in *Othello* than any other play, many are unlikely to reveal a substantial amount about the plot or its themes. Thus, **words** was restricted so that words would only be in the list if they appeared more than 10 times throughout the play (this threshold of 10 was picked somewhat arbitrarily, and there would be advantages and disadvantages associating with changing it in either direction).

```
words = vector()
for (word in c(1:26952)) {
     if (which.max(P[word,])==24 & P[word,24] > 10) {
         words = append(words,rownames(P)[word])
     }
}
print(words)
```

```
##  [1] "'tis"         "heaven"       "nor"          "not"          "such"
##  [6] "ha"           "i"            "some"         "think"        "again"
## [11] "yourself"     "after"        "do"           "purse"        "am"
## [16] "yet"          "if"           "hell"         "tonight"      "that's"
## [21] "foul"         "found"        "lost"         "patience"     "sure"
## [26] "gave"         "dost"         "drunk"        "even"         "put"
## [31] "loves"        "'twas"        "look"         "ho"           "content"
## [36] "help"         "killed"       "alas"         "honest"       "confess"
## [41] "lieutenant"   "michael"      "free"         "suit"         "whore"
## [46] "sense"        "jealous"      "cyprus"       "handkerchief" "emilia"
## [51] "moor"         "iago"         "cassio"       "roderigo"     "desdemona"
## [56] "othello"
```

Some of these words are connective words that don't carry much weight, or character names that obviously appear here more than any other play, so both of these categories are removed.

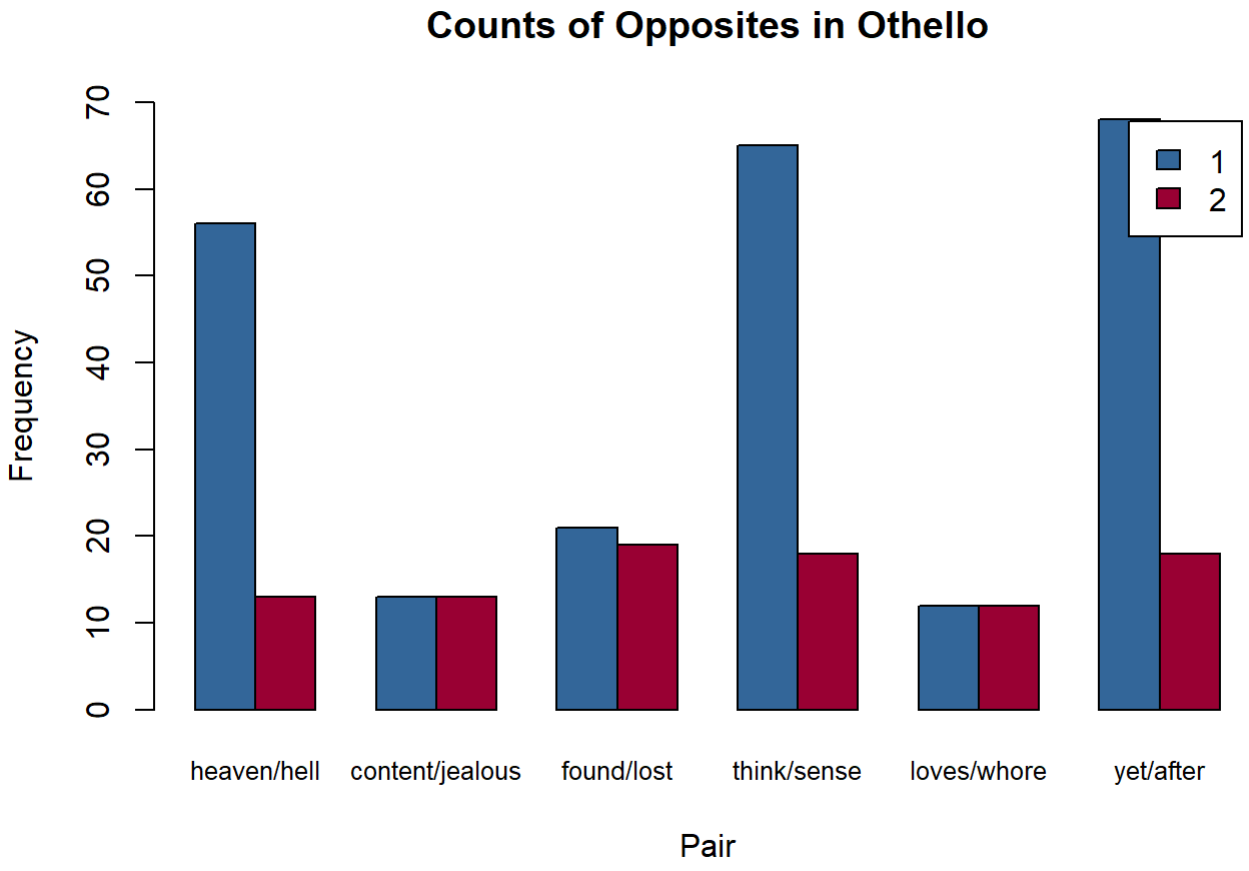
```
words = c("heaven","think","again","yourself","after","purse","yet","hell","tonight","foul","fou
nd","lost","patience","sure","gave","drunk","loves","look","content","help","killed","alas","hon
est","confess","free","whore","sense","jealous","handkerchief")
print(words)
```

```
##  [1] "heaven"       "think"        "again"        "yourself"     "after"
##  [6] "purse"        "yet"          "hell"         "tonight"      "foul"
## [11] "found"        "lost"         "patience"     "sure"         "gave"
## [16] "drunk"        "loves"        "look"         "content"      "help"
## [21] "killed"       "alas"         "honest"       "confess"      "free"
## [26] "whore"        "sense"        "jealous"      "handkerchief"
```

Something striking about the remaining words is that some of them have direct opposites in the list, such as "heaven" and "hell", as well as "found" and "lost". A number of other pairs can also be formed that are not strictly opposites, but have connotations that are in opposition such as "whore" and "loves", as well as "think" and "sense". Six of these pairs have been compiled into the bar graph below, showing the number of times each word appears

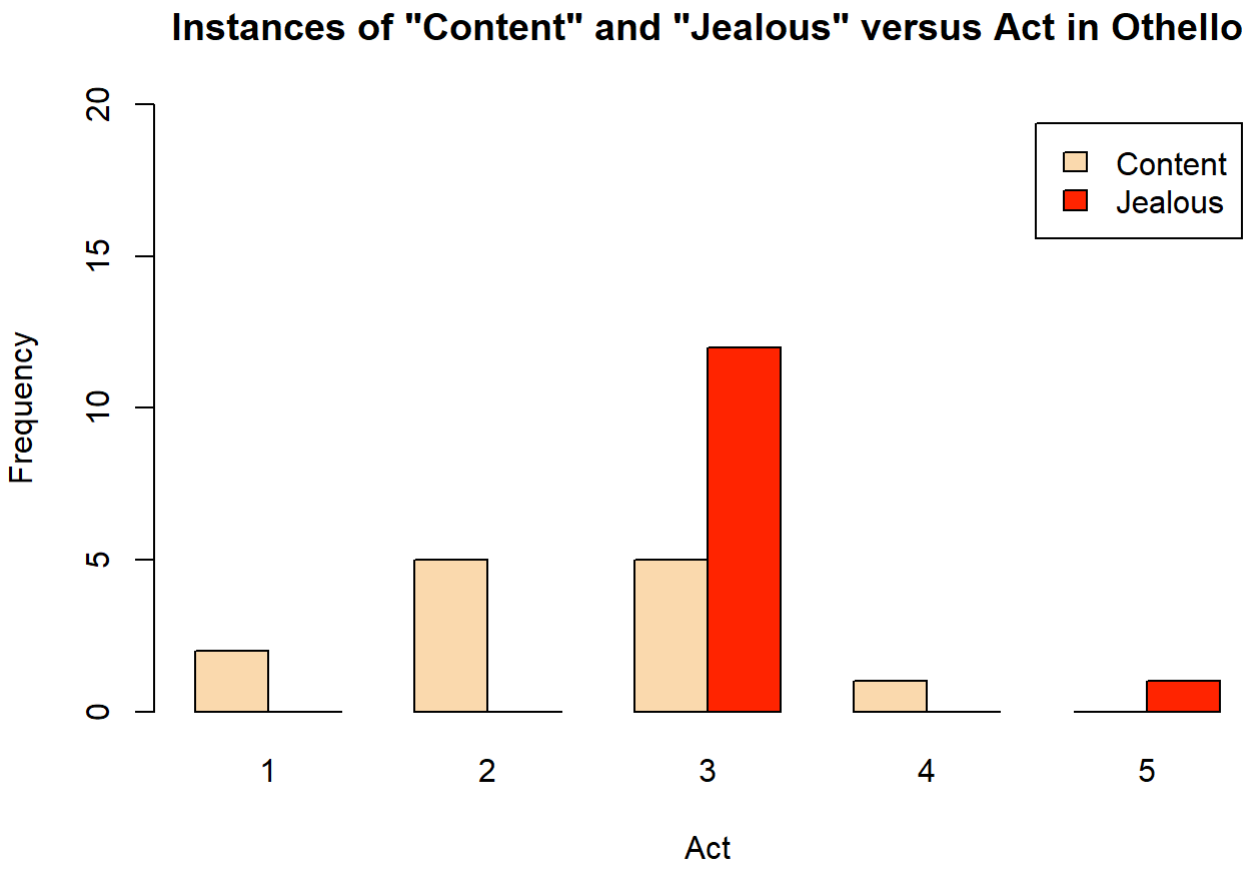
in the play directly next to its counterpart. It is worth noting that the grouping of these opposite pairs is fairly subjective, and it's certainly possible to group them in other ways, or to have fewer or more pairs, but the groupings chosen here are at least plausible.

```
opposites = matrix(
  c(P["heaven",24],P["hell",24],P["content",24],P["jealous",24],P["found",24],P["lost",24],P["th
ink",24],P["sense",24],P["loves",24],P["whore",24],P["yet",24],P["after",24]),
  nrow=2, ncol=6,
  dimnames=list(c(1,2),c("heaven/hell","content/jealous","found/lost","think/sense","loves/whor
e","yet/after")))
barplot(opposites, legend.text=c(1,2), beside = TRUE, col=c("#336699","#990033"), main="Counts o
f Opposites in Othello", xlab="Pair", ylab="Frequency", ylim=c(0,70),cex.names=.8)
```



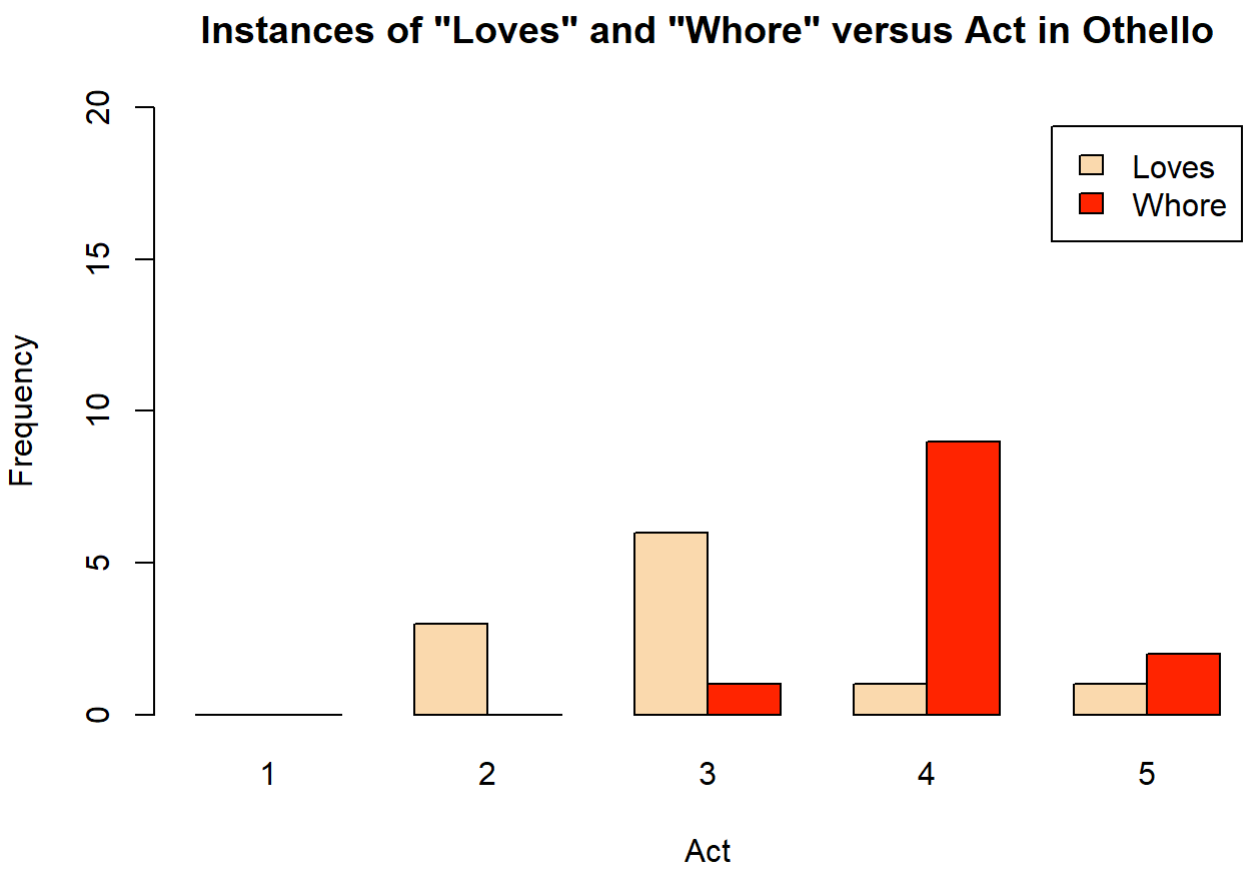
Interestingly, three of the pairs ("content/jealous", "found/lost", and "loves/whore") appear at around the same number of times throughout the play, while the other three pairs appear with very different frequencies. Based on this, one could argue that the pairs occurring with different frequencies are not as good of candidates for opposites, or alternatively that the concepts are not balanced in the play. However, the three that are in equal frequencies suggest either a sort of balance between the opposing concepts in the play, or a shift from one to the other. Two of the pairs subdivided by act are shown below.

```
appearances = matrix(nrow=2,ncol=5,dimnames=list(c("Content","Jealous"),c(1:5)))
for (act in c(1:5)) {
  act_speeches = which(speech_metadata$PLAY=="Oth" & speech_metadata$ACT==act)
  appearances[1,act] = sum(S["content",act_speeches])
  appearances[2,act] = sum(S["jealous",act_speeches])
}
barplot(appearances, legend.text=c("Content","Jealous"), beside = TRUE, col=c("#fad9ad","#ff240
0"), main="Instances of \"Content\" and \"Jealous\" versus Act in Othello", xlab="Act", ylab="Fr
equency", ylim=c(0,20))
```




```
appearances = matrix(nrow=2,ncol=5,dimnames=list(c("Loves","Whore"),c(1:5)))
```

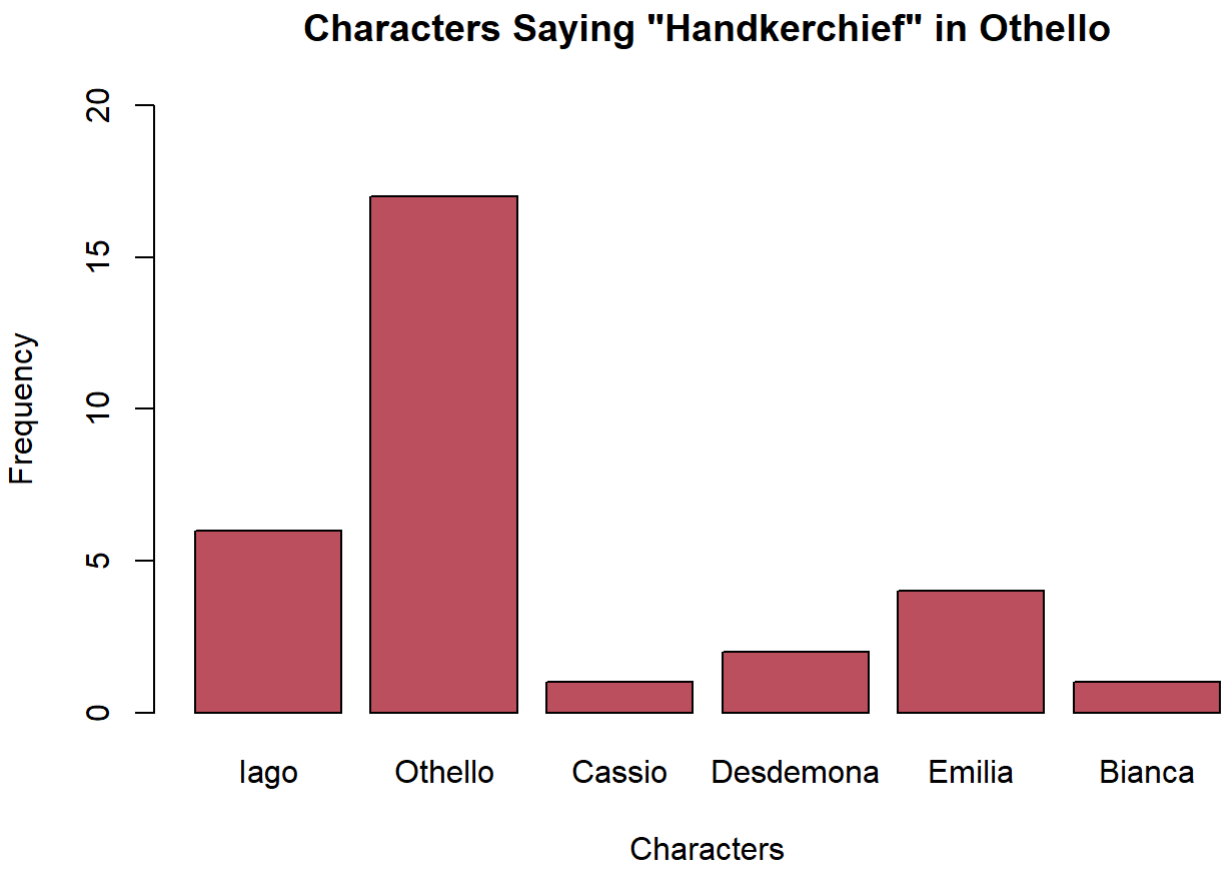
```
for (act in c(1:5)) {
  act_speeches = which(speech_metadata$PLAY=="Oth" & speech_metadata$ACT==act)
  appearances[1,act] = sum(S["loves",act_speeches])
  appearances[2,act] = sum(S["whore",act_speeches])
}
barplot(appearances, legend.text=c("Loves","Whore"), beside = TRUE, col=c("#fad9ad","#ff2400"),
main="Instances of \"Loves\" and \"Whore\" versus Act in Othello", xlab="Act", ylab="Frequency",
ylim=c(0,20))
```



The data show an interesting trend. As the play progresses, “content” and “loves” show up less and less, and “jealous” and “whore” show up more and more. This tracks with Othello’s gradual shift from love for Desdemona to the distrust and jealousy that Iago cultivates in him, and the spike of the word “whore” in Act 4 very likely refers mostly to Desdemona. Thus, the change in Othello’s character arc seems to correlate with the shift from one part of these pairs being used to the other part.

A word that seems a bit out of place in **words** is “handkerchief”, because it clearly has no opposite in the list and is nearly unique in being a physical object. Below, the distribution of who says the word “handkerchief” throughout the play is shown.

```
handkerchief_hits = which(C["handkerchief",] > 0)
handkerchief_hits = handkerchief_hits[which(colnames(C)[handkerchief_hits] %in% character_metada
ta[which(character_metadata$Play.ID=="Oth"),"Character.ID"])]
handkerchief_names = vector()
for (name in c(1:6)) {
     handkerchief_names = append(handkerchief_names, character_metadata[which(character_metadata
[,"Character.ID"]==colnames(C)[handkerchief_hits[name]]),"Name"])
}
barplot(C["handkerchief",handkerchief_hits],names=handkerchief_names,col=c("#bc4f5e"),main="Char
acters Saying \"Handkerchief\" in Othello",xlab="Characters",ylab="Frequency",ylim=c(0,20))
```



As shown, nearly every major character in the play discusses the handkerchief. This is because Othello views it as a token of his love with Desdemona, and Iago, knowing this, uses it as a ploy to further convince Othello of Desdemona's infidelity. Thus, "handkerchief" almost acts as its own opposite, because it means something different to every character and each character's relation to it affects Othello in a different way. Since the handkerchief was initially a private token of their love, it is also indicative of the fact that Othello and Desdemona's love has come to involve other people, such as Iago, that are tearing it apart. In fact, this distribution seems to mirror fairly closely whom Othello turns to for advice about his relationship—he gives much more weight to what Iago says than Desdemona, and it is Emilia in the end of the play who reveals Iago's plot and makes Othello realize his mistake.

Conclusion

Based on the words that Shakespeare uses the most in *Othello*, it seems that many of them connect to others by contrast. Some of the distribution of these pairs of opposing words over time appear to mirror Othello's character arc, and the distribution of characters discussing the handkerchief appears to be connected to the involvement of those characters with Othello's relationship. Thus, these graphs seem to offer some description of Othello's dramatic arc throughout the play, albeit one that requires some context to fully understand. For other plays by Shakespeare, this method of looking for patterns in the words used the most in that particular play could reveal some insight into deeper patterns and themes, and doing so at least offers a reasonable jumping-off point for further investigation.