Keywords of Protagonists in Shakespeare’s Tragedies

Cetin Sert
INF 521 (4-6-2)
69120 Heidelberg
cetin.sert@gmail.com
6. Fachsemester
Studiengang Anglistik/Deutsch als Fremdsprachenphilologie
14.11.2006
Contents

1 Introduction .......................................................................................................... 1
2 Corpus Compilation and Markup ......................................................................... 2
  2.1 Design ............................................................................................................ 2
  2.2 Compilation ................................................................................................... 2
  2.3 Speech Prefix Normalizations ....................................................................... 3
  2.4 Markup .......................................................................................................... 3
  2.5 Subcorpus of Protagonists’ Speeches ............................................................ 4
3 Case Study ............................................................................................................ 4
  3.1 Definition of ‘Word’ ..................................................................................... 4
  3.2 Definition of ‘Word List’ ............................................................................. 5
  3.3 Definition of ‘Keyness’ ............................................................................... 5
  3.4 Methodology – ‘Frequency Profiling using Log-likelihood Values’ .......... 5
  3.5 Table of Positive Keywords for Protagonists in Shakespeare’s Tragedies ... 8
4 Conclusion ............................................................................................................ 9
References ............................................................................................................. 11
1 Introduction

“… if words be made of breath, / And breath of life …”

*Hamlet*, 2573 – 2574

With this term paper, I aim to discover the *key* words which are the ‘breath’ and ‘life’ of the protagonists in Shakespeare’s tragedies.

The starting research question for this study was, “Can corpus linguistic methods be applied to protagonists in Shakespeare’s plays to find out about that particular character or even about the theme of the play in which he happens to exist?” This question eventually matured to focusing on only one genre of Shakespeare’s plays: the tragedies. I also preferred to simplify and reformulate the question: “What are the keywords for Shakespeare’s tragic heroes?”

*First Folio* was chosen as the source of the editions of the tragedies to compile the corpus with, due mainly to the fact that there were publicly available electronic versions and partly to the desire to remain as close to the originality of the language of Shakespeare’s time as possible.

After the compilation and a markup process to ready the texts for the selective exclusion of characters not perceived to be the protagonists in their respective plays, the subcorpus produced consisted of 10 texts which included only the speeches made by the leading characters.

Finally, the methods of frequency listing and *frequency profiling* (Rayson & Garside) were applied to the subcorpus to produce a table of keywords.

Results are discussed in the conclusion.

---

1 Spoken by Queen Gertrude in Hamlet, not by Hamlet himself.
2 Corpus Compilation and Markup

2.1 Design

The corpus used for this term paper was designed to include all tragedies by William Shakespeare in their *First Folio* editions. The table below is derived from the article *Chronology* in *The Oxford Companion to Shakespeare*.

Table 1 – Shakespeare’s Tragedy Plays in Supposed\(^2\) Chronological Order of Composition:

<table>
<thead>
<tr>
<th>Year</th>
<th>Play</th>
</tr>
</thead>
<tbody>
<tr>
<td>1592 – 3</td>
<td><em>Titus Andronicus</em></td>
</tr>
<tr>
<td>1595</td>
<td><em>Romeo and Juliet</em></td>
</tr>
<tr>
<td>1599</td>
<td><em>Julius Caesar</em></td>
</tr>
<tr>
<td>1600 – 1</td>
<td><em>Hamlet</em></td>
</tr>
<tr>
<td>1603 – 4</td>
<td><em>Othello</em></td>
</tr>
<tr>
<td>1605</td>
<td><em>Timon of Athens</em></td>
</tr>
<tr>
<td>1605 – 6</td>
<td><em>King Lear</em></td>
</tr>
<tr>
<td>1606(^3)</td>
<td><em>Macbeth</em></td>
</tr>
<tr>
<td>1606</td>
<td><em>Antony and Cleopatra</em></td>
</tr>
<tr>
<td>1608</td>
<td><em>Coriolanus</em></td>
</tr>
</tbody>
</table>

(Dobson & Wells 78)

2.2 Compilation

The corpus was compiled from the machine-readable versions of the *First Folio* editions of the tragedies named in Table 1, as they are publicly available in the archive of the *Electronic Text Center of the University of Virginia Library*\(^4\). The headers of the machine-readable versions cite solely *The Norton Facsimile* as source.

The plays compiled into the corpus were first retrieved from the archive and saved as plain text, using the browser Mozilla Firefox 1.5.

\(^2\) “… most of the dates are necessarily provisional rather than definitive” (Dobson and Wells 78)

\(^3\) Corrected according to the first paragraph of the article *Macbeth*. “Internal evidence, moreover, particularly the play’s metre, also suggests that *Macbeth* was composed in 1606, after *King Lear* but just before *Antony and Cleopatra …*” (Dobson and Wells 271)

\(^4\) [http://etext.virginia.edu/shakespeare/folio/](http://etext.virginia.edu/shakespeare/folio/)
2.3 Speech Prefix Normalizations

A problem with the First Folio editions of the play texts constituted a major obstacle in the way to dividing the plays into speech units: the same character was referred to in many inconsistent ways.

The problem was solved by normalizing the speech prefixes checking each speech in the First Folio editions of the play texts against its counterpart in The Complete Works of Shakespeare, edited by David Bevington.

Table 2 – An Example of Speech Prefix Normalization:

<table>
<thead>
<tr>
<th>Norton Line Number</th>
<th>Speech Prefix for Romeo</th>
<th>Normalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>164</td>
<td>Rom.</td>
<td>Rom.&gt;romeo</td>
</tr>
<tr>
<td>169</td>
<td>Ro.</td>
<td>Ro.&gt;romeo</td>
</tr>
<tr>
<td>171</td>
<td>Romeo.</td>
<td>Romeo.&gt;romeo</td>
</tr>
<tr>
<td>960</td>
<td>Rome.</td>
<td>Rome.&gt;romeo</td>
</tr>
</tbody>
</table>

To minimize typographical errors, a program was developed and a normalization was fed into this program at the first occurrence of a graphically unique speech-prefix and the program suggested this stored normalization at the following occurrences of the same speech-prefix. Every automatic suggestion was then manually checked.

2.4 Markup

The whole corpus was tagged in XML using the same program employing a small subset of the elements suggested for performance texts in the guidelines of the TEI Consortium.

A unique problem posed the protagonist Marcius Coriolanus in the play Coriolanus. In Bevington’s modernized text of the play, act 1, scene 9, line 67, the protagonist, which is prefixed Marcius up to that point in the play, gets assigned a new speech prefix: Coriolanus. This may be in compliance with the

---

5 Tagger.exe
6 Text Encoding Initiative
fact that he received a title, but there were now two identifiers for the one and the same person: *Marcius* and *Coriolanus*. Consulting the TEI guidelines, which read:

“The who attribute and the `<speaker>` element are both used to indicate the speaker or speakers of a speech, but in rather different ways. The `<speaker>` element is used to encode the word or phrase actually used within the source text to indicate the speaker: it may contain any string or prefix, and may be thought of as a highly specialized form of stage direction. The value of the who attribute however is a unique code, probably made up by the transcriber, which will unambiguously identify the character to whom the speech is assigned” ("The TEI Guidelines")

*Coriolanus* was chosen as the “unique code” to “unambiguously identify the character” ("The TEI Guidelines"). So the resulting XMLs have the same who attribute value before and after the awarding of the title.

```
<sp n="223" c="223" who="coriolanus">
  <speaker ntln="824">Martius.</speaker>
  <l ntln="824">I will goe wash:</l>
...
</sp>
...
<sp n="7" c="467" who="coriolanus">
  <speaker ntln="1687">Corio.</speaker>
  <l ntln="1687">Spoke he of me?</l>
</sp>
```

### 2.5 Subcorpus of Protagonists’ Speeches

As protagonists are “The chief personage in a drama; hence, the principal character in the plot of … a play …” ("OED Online") and the research question was to discover the keywords for the protagonists only, it was needed to create a subcorpus which only contained their speeches. Stage directions and speech prefixes (because the speeches are all of the same protagonist) were excluded.

The following case study was done solely on this subcorpus of 62,304 tokens.

### 3 Case Study

#### 3.1 Definition of ‘Word’

In compliance with WordSmith Tools 4.0 help files, the word is defined as:
“… a sequence of valid characters with a word separator at each end. Valid characters include all the letters from A to Z, plus all accented characters …, plus any user-defined acceptable characters to be included within a word (such as the apostrophe or hyphen).” (Scott)

The implementation of this definition in both WordSmith Tools⁷ and Tenka Text⁸ catch the following as words.

- ‘tis
  - Apostrophe or hyphen at the beginning does not count as part of the word.
- th
  - Apostrophe or hyphen at the end does not count as part of the word.
- o’re-powr’d ⁹
  - An apostrophe or hyphen in the middle of two letters counts as part of the word.

### 3.2 Definition of ‘Word List’

Basically speaking, a word list is a list of frequencies where for each word type the number of its tokens – its frequency – is stored. (Kennedy 245)

They can be likened to and are implemented in corpus analysis programs as word-frequency dictionaries in which one looks up a word type to see how many tokens of it occur in a text.

### 3.3 Definition of ‘Keyness’

Key, in terms of this research, is a word in a source speech text whose “relative frequency” is significantly greater than that of the same word in a reference speech text. (Rayson & Garside) In this study, log-likelihood values were used to determine the key words.

### 3.4 Methodology – ‘Frequency Profiling using Log-likelihood Values’

The method described here and used to compare the texts of the protagonists was taken from the paper “Comparing Corpora using Frequency Profiling”¹⁰ presented by Rayson & Garside in proceedings of the workshop on comparing corpora, held in conjunction with the 38th annual meeting of the Association for

---

⁷ WordSmith Tools 4.0 (with “hyphens do not separate words” option)
⁸ Tenka Text pre-alpha 2006-11-14
⁹ Antony and Cleopatra, 987
Computational Linguistics (ACL 2000). I found the indirect link to this paper in the help files of WordSmith Tools 4.0\textsuperscript{11}.

Given two corpora we wish to compare, we produce a frequency list\textsuperscript{12} for each corpus. … For each word in the two frequency lists we calculate the log-likelihood (henceforth LL) statistic. This is performed by constructing a contingency table as in Table 1\textsuperscript{13}.

<table>
<thead>
<tr>
<th></th>
<th>CORPUS ONE</th>
<th>CORPUS TWO</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freq of word</td>
<td>O1</td>
<td>O2</td>
<td>O1+O2</td>
</tr>
<tr>
<td>Freq of other words</td>
<td>N1-O1</td>
<td>N2-O2</td>
<td>N1+N2-O1-O2</td>
</tr>
<tr>
<td>TOTAL</td>
<td>N1</td>
<td>N2</td>
<td>N1+N2</td>
</tr>
</tbody>
</table>

…

We need to calculate the expected values (E) according to the following formula:

\[ E_i = \frac{N_i \times O_i}{\sum_i N_i} \]

… So … E1 = N1*(O1+O2) / (N1+N2) and E2 = N2*(O1+O2) / (N1+N2). The calculation for the expected values takes account of the size of the two corpora, so we do not need to normalize the figures before applying the formula. We can then calculate the log-likelihood value according to this formula:

\[ -2 \ln \lambda = 2 \sum_i O_i \ln \left( \frac{O_i}{E_i} \right) \]

This equates to calculating the LL as follows:

\[ LL = 2*((O1*\log(O1/E1)) + (O2*\log(O2/E2))) \]

The word frequency list is then sorted by the resulting LL values. This gives the effect of placing the largest LL value at the top of the list representing the word which has the most significant relative frequency difference between the two corpora. In this way, we can see the words most indicative (or characteristic) of one corpus, as compared to the other corpus, at the top of the list.

(Rayson & Garside)

\textsuperscript{11} To find the same help page: Help > Index > KeyWords: Calculation
\textsuperscript{12} Word frequency list
\textsuperscript{13} Variable names have been changed to the initials of what they stand for: O = observed frequency, N = number of tokens.
This method was implemented in the software developed for this term paper and the text of one protagonist was compared against the texts of the other nine protagonists. The comparison was binary, i.e. one text was compared with another at a time. (E.g. the text of Romeo was compared against that of Titus, and then that of Brutus, and then that of Hamlet all the way to Coriolanus.) After the list of keywords for each protagonist was computed, any keyword which was not key for one character against at least seven others was removed. This condition was built-in to get only those few words which are truly exclusive for each protagonist. The results were sorted on the log-likelihood values.

Excerpt from the resulting positive keys output file for Romeo:

Key frequency list for 2.e.romeo

... 
LOUE, logLikelihood=50,91, referenceName=1.e.titus, ... index=1
LOUE, logLikelihood=40,11, referenceName=4.x.hamlet, ... index=2
LOUE, logLikelihood=39,21, referenceName=10.l.coriolanus, ...
index=3
LOUE, logLikelihood=27,54, referenceName=8.x.macbeth, ... index=4
LOUE, logLikelihood=24,42, referenceName=6.x.timon, ... index=5
LOUE, logLikelihood=23,90, referenceName=7.x.lear, ... index=6
LOUE, logLikelihood=22,23, referenceName=9.x.antony, ... index=7
LOUE, logLikelihood=20,54, referenceName=3.x.brutus, ... index=8
LOUE, logLikelihood=11,78, referenceName=5.x.othello, ... index=9
...
IULET, ...

The data from these output files\textsuperscript{14} were used to construct the table of positive keywords.

\textsuperscript{14} At \url{http://ucrel.lancs.ac.uk/llwizard.html} is an online log-likelihood calculator where one might confirm the accuracy of computation of the log-likelihood values computed with the software developed for this term paper.
### 3.5 Table of Positive Keywords for Protagonists in Shakespeare’s Tragedies

<table>
<thead>
<tr>
<th>1592 - 3</th>
<th>1595</th>
<th>1599</th>
<th>1600 - 1</th>
<th>1603 - 4</th>
<th>1605</th>
<th>1605 - 6</th>
<th>1606</th>
<th>1606</th>
<th>1608</th>
</tr>
</thead>
<tbody>
<tr>
<td>Titus</td>
<td>Rome</td>
<td>Brutus</td>
<td>Hamlet</td>
<td>Othello</td>
<td>Timon</td>
<td>Lear</td>
<td>Macbeth</td>
<td>Antony</td>
<td>Coriolanus</td>
</tr>
<tr>
<td>MARCUS</td>
<td>LOUE</td>
<td>CASSIUS</td>
<td>KING</td>
<td>IAGO</td>
<td>THOU</td>
<td>REGAN</td>
<td>BANQUO</td>
<td>EROS</td>
<td>VOYCES</td>
</tr>
<tr>
<td>SONNES</td>
<td>JULIET</td>
<td>CAESAR</td>
<td>PLAY</td>
<td>HER</td>
<td>GOLD</td>
<td>DAUGHTERS</td>
<td>FEARE</td>
<td>CAESAR</td>
<td>THEM</td>
</tr>
<tr>
<td>LAUINIA</td>
<td>O</td>
<td>ANTONY</td>
<td>MOTHER</td>
<td>SHE</td>
<td>THEE</td>
<td>DAUGHTER</td>
<td>CAWDOR</td>
<td>QUEENE</td>
<td>THEIR</td>
</tr>
<tr>
<td>ROME</td>
<td>BANISHED</td>
<td>BRUTUS</td>
<td>HORATIO</td>
<td>OH</td>
<td>GODS</td>
<td>NO</td>
<td>THE</td>
<td>POMPEY</td>
<td>EGYPT</td>
</tr>
<tr>
<td>TEARES</td>
<td>HERE</td>
<td>LUCIUS</td>
<td>IT</td>
<td>CASSIO</td>
<td>MEN</td>
<td>OLD</td>
<td>WHICH</td>
<td>EGYPT</td>
<td>MOTHER</td>
</tr>
<tr>
<td>EMPEROUR</td>
<td>SHE</td>
<td>CAESARS</td>
<td>A</td>
<td>HANDKERCHIEFE</td>
<td>WERT</td>
<td>FRANCE</td>
<td>SLEEPE</td>
<td>FIGHT</td>
<td>AUFFIDIUS</td>
</tr>
<tr>
<td>LUCIUS</td>
<td>FAIRE</td>
<td>MESSALA</td>
<td>HIS</td>
<td>HEAUFEN</td>
<td>ATHENS</td>
<td>KING</td>
<td>THANE</td>
<td>CLEOPATRA</td>
<td>YOUR</td>
</tr>
<tr>
<td>SHE</td>
<td>DEARE</td>
<td>HE</td>
<td>OH</td>
<td>HONEST</td>
<td>HONEST</td>
<td>NATURE</td>
<td>WOOD</td>
<td>ANTHONY</td>
<td>THEY</td>
</tr>
<tr>
<td>DOE</td>
<td>LIPS</td>
<td>SHALL</td>
<td>PLAYERS</td>
<td>DESDEMONA</td>
<td>TIMON</td>
<td>CORDELIA</td>
<td>TIME</td>
<td>LAND</td>
<td>HAUE</td>
</tr>
<tr>
<td>AND</td>
<td>IS</td>
<td>LUCIUS</td>
<td>THE</td>
<td>DIUELL</td>
<td>THY</td>
<td>FOOLE</td>
<td>DEED</td>
<td>CAESARS</td>
<td>VOLCES</td>
</tr>
<tr>
<td>TRIBUNES</td>
<td>DEATH</td>
<td>ANY</td>
<td>HAMLET</td>
<td>WIFE</td>
<td>FRIENDS</td>
<td>HA</td>
<td>HIS</td>
<td>GONE</td>
<td>COMMON</td>
</tr>
<tr>
<td>EMPRESSE</td>
<td>MERCUTIO</td>
<td>AWAKE</td>
<td>LORD</td>
<td>SOULE</td>
<td>MANKINDE</td>
<td>DO'S</td>
<td>DARE</td>
<td>COMMAND</td>
<td>A'TH</td>
</tr>
<tr>
<td>BROTHER</td>
<td>NURSE</td>
<td>GOOD</td>
<td>EXCELLENT</td>
<td>DO'ST</td>
<td>APEMANTUS</td>
<td>CALL</td>
<td>BORNE</td>
<td>SEA</td>
<td>TH</td>
</tr>
<tr>
<td>HANDS</td>
<td>LIGHT</td>
<td>WE</td>
<td>GOD</td>
<td>DOST</td>
<td>ALCIBIADES</td>
<td>TILL</td>
<td>SWORD</td>
<td>TULLUS</td>
<td>CONSULL</td>
</tr>
<tr>
<td>THESE</td>
<td>LADY</td>
<td>PORTIA</td>
<td>VERIE</td>
<td>TWAS</td>
<td>THEEUES</td>
<td>BLOOD</td>
<td>MINDE</td>
<td>MORROW</td>
<td>WOUNDS</td>
</tr>
<tr>
<td>LORD</td>
<td>ROMEO</td>
<td>VOLUMNIUS</td>
<td>COULD</td>
<td>GAUE</td>
<td>VILLAINES</td>
<td>MAN</td>
<td>TH</td>
<td>PROFESS</td>
<td>CORNE</td>
</tr>
<tr>
<td>GEE</td>
<td>TYBALT</td>
<td>CAIUS</td>
<td>ENGLAND</td>
<td>WHORE</td>
<td>MAN</td>
<td>WOMAN</td>
<td>ELSE</td>
<td>PEOPLE</td>
<td></td>
</tr>
<tr>
<td>PUBLICUS</td>
<td>SIN</td>
<td>BY</td>
<td>LAERTES</td>
<td>FOUND</td>
<td>THEEFE</td>
<td>PLAGUE</td>
<td>ELSE</td>
<td>PEOPLE</td>
<td></td>
</tr>
<tr>
<td>BOY</td>
<td>RICH</td>
<td>PHILIPPI</td>
<td>MADNESSE</td>
<td>HAD</td>
<td>PROUE</td>
<td>PEOPLE</td>
<td>PEOPLE</td>
<td>PEOPLE</td>
<td></td>
</tr>
<tr>
<td>GOTHES</td>
<td>BEAUTY</td>
<td>ROMANS</td>
<td>SPEECH</td>
<td>HUSBAND</td>
<td>WHOE</td>
<td>WHOE</td>
<td>ELSE</td>
<td>PEOPLE</td>
<td></td>
</tr>
<tr>
<td>SATURNINE</td>
<td>JOY</td>
<td>FOR</td>
<td>MADAM</td>
<td>IST</td>
<td>CONFOUN</td>
<td>THOSE</td>
<td>THOSE</td>
<td>STAND</td>
<td>WHAT'S</td>
</tr>
<tr>
<td>HEEERE</td>
<td>HOLY</td>
<td>ROMAN</td>
<td>LIGHT</td>
<td>FORTUNES</td>
<td>TIS</td>
<td>BEAST</td>
<td>TIS</td>
<td>WHEREIN</td>
<td>BESEECH</td>
</tr>
<tr>
<td>REUENGE</td>
<td>FAREWELL</td>
<td>MARCH</td>
<td>MARCH</td>
<td>LIGHT</td>
<td>FORTUNES</td>
<td>TIS</td>
<td>BEAST</td>
<td>TIS</td>
<td>WHEREIN</td>
</tr>
<tr>
<td>SORROW</td>
<td>DOETH</td>
<td>MARKE</td>
<td>MARKE</td>
<td>THAT'S</td>
<td>SENATORS</td>
<td>WHOM</td>
<td>WHO</td>
<td>WHOM</td>
<td>BESEECH</td>
</tr>
<tr>
<td>MURDER</td>
<td>LOUES</td>
<td>MEN</td>
<td>MEN</td>
<td>WANT</td>
<td>WANT</td>
<td>WANT</td>
<td>WANT</td>
<td>WANT</td>
<td>FIGHT</td>
</tr>
<tr>
<td>WITNESSE</td>
<td>EYE</td>
<td>DEATH</td>
<td>DEATH</td>
<td>DOGGE</td>
<td>DOGGE</td>
<td>DOGGE</td>
<td>DOGGE</td>
<td>DOGGE</td>
<td>FIGHT</td>
</tr>
<tr>
<td>HART</td>
<td>POYSON</td>
<td>STAND</td>
<td>STAND</td>
<td>EACH</td>
<td>EACH</td>
<td>EACH</td>
<td>EACH</td>
<td>EACH</td>
<td>FIGHT</td>
</tr>
<tr>
<td>READ</td>
<td>EYES</td>
<td>EUERY</td>
<td>EUERY</td>
<td>LORD</td>
<td>LORD</td>
<td>LORD</td>
<td>LORD</td>
<td>LORD</td>
<td>FIGHT</td>
</tr>
</tbody>
</table>

Data used to construct the table can be found on the CD: [cd:\subcorpus\keywords\](http://tenkatext.sourceforge.net/tp2/subcorpus/keywords/)

4 Conclusion

This paper was written to answer the research question: “What are the keywords for Shakespeare’s tragic heroes?”.

To answer this question with a corpus linguistic approach to the subject, a corpus was compiled and improved with necessary annotations. From this main corpus of complete play texts, a subcorpus of speech texts of protagonists’ was extracted.

Methods of frequency listing and profiling were applied to this subcorpus and a table of keywords was created.

In spite of the general impression that spelling variations make the First Folio a relatively more difficult body of texts to work with, which manifests itself for example in this sentence:

“The spelling, punc[t]uation, and other early modern printing conventions have not been normalized, a factor to be kept in mind when planning a search.” (Shakespeare, "The Shakespeare First Folio").

the table of keywords produced provides an acceptable collection of words which seem to be the most characteristic of the protagonists. Titus speaks mostly of ‘sonnes’ which play key roles in the play, ‘teares’ shed, and of ‘hands’ because ‘Lauinia’s and his are cut. Filled with “sweet reuenge” (2353), he lets the very word show itself in the table of results.

Even more striking are Romeo’s column where we see two words’ childish rush to the top of table to tell us the very substance of the play. Including the column header, these read almost like a sentence: “Romeo loue Iulet”. He speaks of ‘lips’, ‘eyes’, which allow but brief moments of *louely* contact for the lovers. ‘banished’ and ‘poyson’, the two blades of a fatal claw, turn the mood to that of a tragedy. Yet Romeo lets neither his ‘loue’ nor ‘Iulet’ fall between the blades of this claw, they soar in peace above anything that may have crossed his mind or touched his heart.
The results this study has produced can indeed be confirmed as the keywords\textsuperscript{15} for the protagonists. So the question of research seems to have been answered successfully by the application of the intercontrastive frequency profiling method extended to compare more than two texts.

\textsuperscript{15} In the corpus linguistic meaning of the word. (A single word uttered only once but which nonetheless introduces a total turn of events would surely be key for literary scholars but such a word would not make it into keyword lists or tables. (independent of the program used))
References


—. The Shakespeare First Folio. 1 October 2006. 17 July 2006<http://etext.virginia.edu/shakespeare/folio/>.