**A New Chronology for Shakespeare’s Plays**

**Douglas Bruster and** **Genevieve Smith**

University of Texas at Austin, United States of America

**Abstract**

It is widely recognized that Shakespeare’s verse lines grew progressively longer as his career unfolded. Scholars have traditionally used this fact, among others, to date the plays. Drawing on both published and original data relating to their syntactical structures, this essay constructs a new chronology for 41 dramatic texts, and parts of texts, by Shakespeare. This chronology is based on a constrained correspondence analysis (CCA) of the plays’ internal pauses, qualified in relation to a principal component analysis (PCA) of other of their verbal features. The result is a more specific ordering of the Shakespeare canon than has previously been available. Employing bootstrapping methodology to suggest probable boundaries for its composition, we locate each text at a specific moment relative to periods during which the Elizabethan playhouses were open. Because recent scholarship has established the Shakespearean authorship of several disputed texts, and has been able to identify the specific portions that Shakespeare contributed to others, our enhanced data set augments previous analyses in scope as well as precision.

**1 Introduction**

It is a commonplace that Shakespeare's lines became longer throughout his career. Nearly as familiar to scholars is how this aspect of his verse, particularly his lines' internal pauses as marked by punctuation, sheds light on his works' chronology. Together with external evidence (such as publication, or records attesting to performance or the availability of a manuscript for printing), changes to Shakespeare’s habits in versification have helped establish our timeline of his works. Major chronologies of the plays and poems, including those of E.K. Chambers (Chambers, 1930), G. Blakemore Evans (Evans, 1974, rev. 1996), and Stanley Wells and Gary Taylor (Wells and Taylor, 1987), have drawn on what we know about the patterns of Shakespeare’s prosody to order his works. These three chronologies agree on the general shape of his literary output, placing *Julius Caesar* and *Henry V* at the midpoint of thirty-eight plays so evaluated. They disagree, however, as to which year or years various works were written, as well as which came before or after others in the canon.

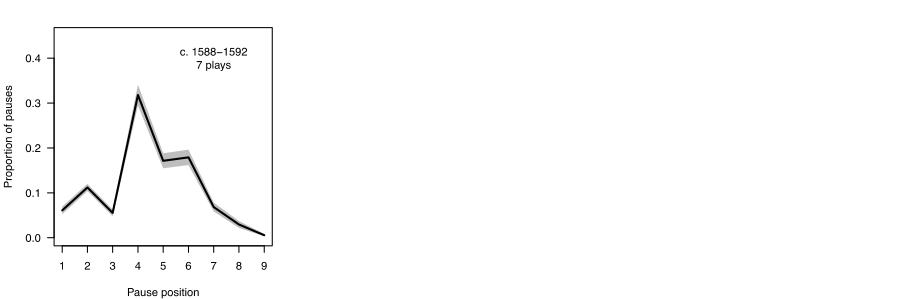
The present study offers a new chronology for Shakespeare’s plays based on an analysis of the most extensive data concerning the structure of Shakespeare’s verse lines: the pause figures collected by Ants Oras (Oras, 1960). We revise some of Oras’s figures in light of new findings concerning attribution, narrowing Shakespeare’s portion of particular plays (*Titus Andronicus*, *Timon of Athens*) and adding to our data set parts of two other texts (*Edward III*, the Additional Passages to the 1602 *Spanish Tragedy*). This enhanced data set is then subjected to both a principal component analysis (PCA) and constrained correspondence analysis (CCA), with various methodological modifications. The latter includes setting a range for Shakespeare’s literary output and fixing selected “anchor” texts for the determination of dates for the remaining plays. These dates are then compared with the figures from an original PCA of new data concerning various linguistic features in Shakespeare’s verse (Tarlinskaja, 2014). We employ a bootstrapping procedure to establish a likely range for the composition of each work. Finally, in light of a theory advanced by J. Leeds Barroll (Barroll, 1991), we construct our timeline of Shakespeare’s plays by coordinating the CCA’s date predictions with periods when the playhouses of Shakespeare’s time were open for business.

**2 Chronology: Background**

Oras's aim was to demonstrate historical and personal patterns in iambic pentameter by tabulating where punctuated pauses fall within its first nine syllables (punctuation after the tenth syllable is not counted). Pauses can be counted in three different ways in Oras's tabulation; he labels these A, B, and C pauses. A pauses are those signaled by punctuation of any kind within a pentameter line (Oras counts short lines, but not their terminal punctuation). B pauses, a sub-group of the A pause, are so-called "strong" pauses within the line: pauses signaled by any punctuation mark other than a comma, including periods, question marks, colons, semi-colons, and dashes. C pauses are comprised of punctuation marks dividing "split-" or "shared" lines. Almost by definition, C pauses are a sub-group of the B pause: the punctuation dividing shared lines is invariably "strong" in nature, rather than a comma. All B and C pauses are also A pauses, with C pauses representing the smallest (because most heavily specified) of the groups.

Oras counted A, B, and C pauses for thirty-eight Shakespeare plays, adding the A and B figures as well for *Venus and Adonis*, *The Rape of Lucrece*, and the *Sonnets*.

Oras presented his figures both numerically and in line graphs that display the percentages of pauses in the first nine syllabic positions of Shakespeare’s pentameter line. Because Shakespeare’s verse is iambic, with an unstressed syllable followed by a stressed one, pauses tend to come after the even syllables, making these graphs a virtual study in pulsation. The changes they reveal have clear significance for the study of chronology. For example, Figure 1 represents the changing pause patterns—the averages for the nine pause positions—in groups of plays drawn from what we believe to be different phases of his career.



**Fig. 1.** Pauses in Shakespeare's plays: early, middle, and late. [cite titles by group.]

As these graphs indicate, in plays written at the beginning of his time as playwright, in the early and mid-1590s, pauses cluster heavily after the fourth syllable. As his career progresses, however, the distribution balances between the fourth and the sixth positions. Finally, toward the end of his time as a dramatist, the most significant proportion of pauses shifts markedly to the sixth position, with a greater number in the second half of the line than the first.

The relevance of such data for chronologies of Shakespeare’s work has long been recognized (Bathurst, 1857). Oras’s research was used for what we will call the “Oxford chronology” (Wells and Taylor 1987, pp. 69-144), which reproduces Oras’s A, B, and C pauses in separate columns opposite an ordered list of plays. In its discussion of various plays, the Oxford chronology refers to Oras’s data for confirmation of an estimated date or range. Yet there is some divergence between the order implied by Oras's figures and the order of the Oxford chronology. That is, 9 of the 38 plays ordered in the Oxford chronology share an exact position with the sequence that Oras's A pauses suggest. These are *Shrew*, *3 Henry VI*, *Richard III*, *King John*, *Julius Caesar*, *Timon*, *Lear*, *Macbeth*, and *Kinsmen*. Fifteen other plays fall within two slots of each other in the Oras A order and Oxford chronology: *1 Henry VI*, *Errors*, *Love's Labor's Lost*, *Richard II*, *Romeo*, *Dream*, *Merry Wives*, *2 Henry IV*, *Much Ado*, *As You Like It*, *Measure*, *Othello*, *Pericles*, *Winter's Tale*, and *Cymbeline*.

Yet almost as many plays, 14, are separated by three or more places in the two lists, among them *Two Gentlemen of Verona*, *2 Henry VI*, *Titus Andronicus*, *Merchant*, *1 Henry IV*, *Henry V*, *Hamlet*, *Twelfth Night*, *Troilus*, *All's Well*, *Antony*, *Coriolanus*, *Tempest*, and *Henry VIII*. Of course, no test of any linguistic feature--whether run-on lines, feminine endings, or colloquialism in verse--should be expected to produce smooth, universally satisfying results. This is particularly the case because so many factors, extrinsic and intrinsic alike, can affect the makeup of a literary text. At the same time, however, it seems significant that the Oxford chronology and the Oras data disagree to this extent. Some of the plays are quite divergent in their places: Oras's figures for "first half" pauses, for instance, would have us put *Troilus* seven places earlier than Oxford locates it, and *Merchant* seven places later; *Antony* is, by this measure, six places later in Oras, *Titus* six, and both *Two Gentlemen* and *Coriolanus* four places later than in the Oxford chronology. Added to this puzzle is the extremely unlikely positioning, in the Oras data, of *2 Henry IV* before *1 Henry IV*, and of *The Tempest* before *Pericles*--chronological placements with which few if any scholars would be likely to agree.

The Oxford chronology’s use of the Oras data formed the basis of the most sustained examination to date of the relation between syntax and temporal ordering in Shakespeare: MacDonald P. Jackson's "Pause Patterns in Shakespeare's Verse: Canon and Chronology” (Jackson 2002). There Jackson describes Oras's methodology and findings before submitting his A-pause figures to statistical analysis. Jackson uses SPSS (Statistical Package for the Social Sciences) software to compare each play's A pauses with every other play's A pauses, producing 1640 Pearson product moment correlation coefficients to indicate how close each play is, in terms of its pause patterning, to all the other plays. Listing each play separately, Jackson provides its five closest correlations in descending order, and notes that the results tend to confirm the accuracy of the Oxford chronology and support our traditional understanding of the evolution of Shakespeare’s versification.

Because his methodology emphasizes relation and proximity among plays, Jackson does not seek to establish new dates for them. Yet he acknowledges that his analysis produced correlations that diverge significantly from what the Oxford chronology would predict. Six plays come in for particular mention: *The Merchant of Venice* (which his results would place later than Oxford); *Merry Wives* (later than Oxford); *2 Henry IV* (varied, but on the whole earlier than Oxford); *Troilus* (earlier than Oxford); *Othello* (earlier than Oxford); and *All's Well* (later than Oxford). This divergence seems important, not least because such also characterizes Oras’s relation to the 1930 chronology of Chambers. That is, Oras employed Chambers’s chronology but did not revise it, even though his own graphs and numbers contradicted the Chambers order in numerous instances. Like received narratives generally, chronologies can be “sticky” phenomena: something fixed through custom, and hard to dislodge (Kuhn, 1962).

**3 Correspondence Analysis and Principal Component Analysis**

Oras's pause figures suggest a seemingly straightforward procedure for ordering the plays: simply record where a particular texts fits into an ascending or descending row of figures. If straightforward, such ranking simplifies the complexity of the data. To address the differing number of pauses in various texts, for instance, Oras quite understandably made them equal by converting their pauses to percentages. But while this commonsense solution solves one problem, it creates another, as the plays (and parts of plays) vary greatly in the amount of data they offer. Thus treating the shortest Shakespeare text in the sample (in this case, his contributions to the lightly-punctuated *Sir Thomas More*, with a scant 32 pauses) as statistically equivalent to his most pause-heavy text (*Cymbeline*, with 2,735 pauses) emphasizes the former at the expense of the latter. Making their pause data equal 1, that is, imposes a statistical constraint on them both, and implies equal confidence in how representative their information is. Thus the element of the Oxford chronology that looks to Oras for confirmation, and to a certain extent Jackson's more statistically sophisticated analysis of that chronology (Jackson, 2002), rely on artificially constrained data. How, then, to acknowledge the differential weight of the Oras-type data?

A valuable method for comparing compositional data--including multivariate data--is correspondence analysis (or CA) (Greenacre, 2007). At its most basic, correspondence analysis is a computer-intensive, statistical methodology that takes categorical information and looks for associations, and strength of associations, in the relations of rows (where its various objects and assemblages are listed) to columns (where such variables as, for instance, time, location, and income, among others, may be recorded). Correspondence analysis first attempts to identify, and then rank, the most statistically significant variation in that data. Thus the first CA axis will account for the largest amount of variation in the original data, the second axis will account for the next largest portion, and so on. By identifying the most crucial of these variables, researchers can visualize and interpret the bulk of the variability in any original data. For this reason correspondence analysis as well as a related methodology, principal component analysis (PCA), are known as *dimensionality reduction methods*, meaning they reduce the complexity of a set of data by boiling it down to a few axes of major importance.

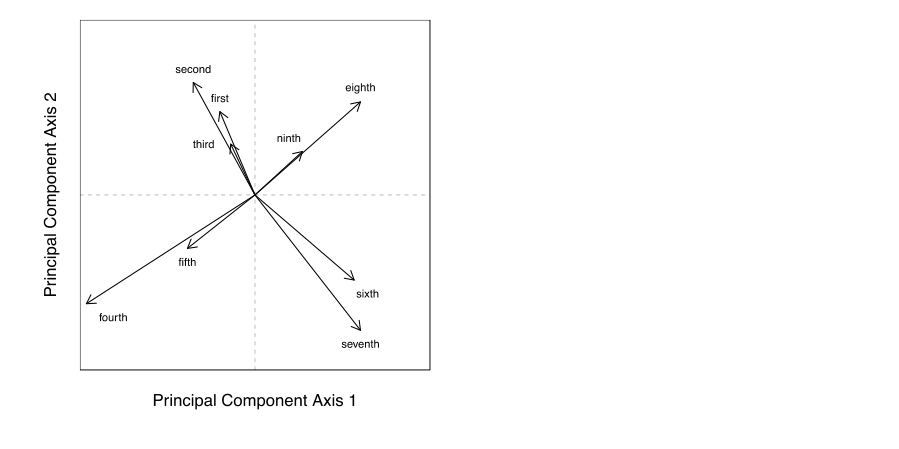
Adopted widely across the natural- and social sciences, correspondence analysis has been employed for a variety of research purposes. Germane to our purposes here is the extensive, and sophisticated, use of correspondence analysis in seriation studies. Seriation, or simply "putting things in order," is usually an exercise in relative dating, employed when an absolute dating method may be unavailable. In the field of archeology, for example, researchers are often confronted with artifacts that only occasionally have information regarding production or use attached to them. When carbon dating or information relating to, say, tree rings or chemical composition is unavailable, archaeologists have refined the statistical bases of correspondence analysis to help them order, and thus date (however approximately) things of uncertain origin. By comparing the known composition of pottery remains, for instance, archeologists may place certain assemblages closer together in time based on how similar they are to one another.

Shakespeare's plays may not strike us as artifacts, of course, but the procedures of correspondence analysis as refined for seriation nevertheless provide a statistically rigorous methodology for examining their material components. On the basis of such an examination, in fact, we can offer a tentative chronology that responds to the differential distribution of features--in this case, pauses as recorded in the verse of early texts--throughout the canon. This chronology should be understood as a provisional timeline of *when the iambic pentameter in the play texts under examination was largely composed*. By emphasizing this last phrase, we mean to call immediate attention to two things. First, Oras’s prose data comes from the pentameter verse in plays that are sometimes—particularly in the late 1590s--made up heavily of prose. Pause pattern analysis is therefore limited by the amount of verse in each play. Second, there is a strong likelihood that a number of Shakespeare's plays were written at one time (even over various times) and revised at another, or others. A play title thus need not connote an event, but may sometimes have been multiple events, even a process. The fixity of any ordering, then, needs to be qualified in the context of diachronic composition.

[I think we need to include a few lines here about our updating of Oras’ counts (how you did it, a reference to the table of values in the appendix, a list of plays for which we used only Shakespeare’s portion, etc.

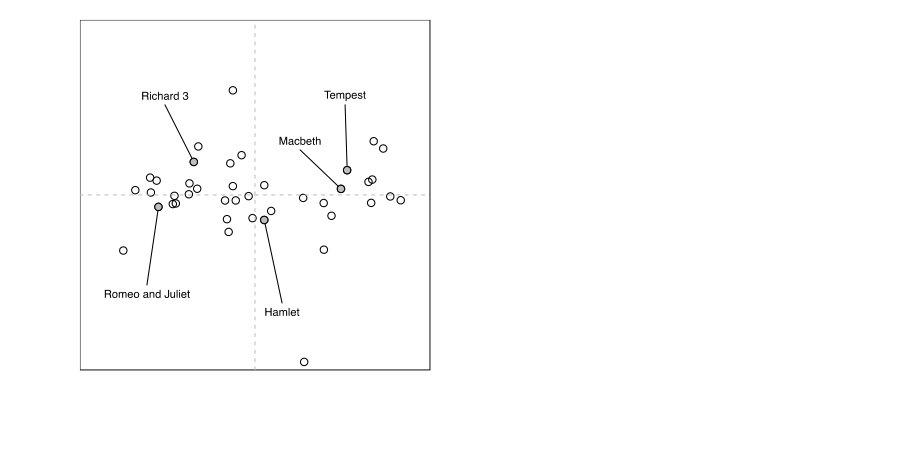
Our next two illustrations, Figures 2 and 3, introduce the results of two statistical procedures performed on our data set (see Appendix B). These are, respectively, a principal component analysis (PCA) and a correspondence analysis (CA), with results from both procedures displayed, left to right, in each graph. Given our nine discrete categories of pauses, it would conceivably be possible to map out the position of each play text (and, in the case of collaborative work, of each relevant part of a play text) in nine dimensions. For ease of interpretation, the two pairs of biplots below (in Figures 2 and 3) map the Euclidian distances among the data in two dimensions. The panels of Figure 2 plot the pause data as arrows in relation to the first two axes of variation, with the most important variation on the horizontal axis (the dashed line bisecting the square horizontally), and the second most important variation on the vertical axis (the vertical dashed line). The left panel represents principal component analysis (PCA). One could note that the foremost variation, labeled PC1 on the x-axis, is mostly capturing variation in 4th, 6th, 7th, and 8th position pauses (looking at how far each arrow travels left-right--that is, along PC1--those arrows reach the farthest). For its part, PC2 (labeled on the y-axis) is mainly capturing variation in 4th, 7th, 2nd, and 8th positions; again, one could notice far up or down each arrow goes to gauge its contribution.

The patterns are broadly similar in the correspondence analysis (CA) panel on the right-hand side of Figure 2. Yet now the 4th position arrow, although slightly shorter, is more closely aligned with the horizontal plane. This means that changes to pauses in the 4th position contribute a great deal to the primary axis of variation (CA1) in our correspondence analysis, but hardly at all to the second most important (CA2). Similarly, 6th position pauses are contributing somewhat to CA1, but mostly to CA2.



Our next illustration maps the data derived from the principal component analysis (PCA) and correspondence analysis (CA) described by the previous figure. Figure 3 indicates the position of each text and relative distance from other texts in a two-dimensional rendering of Euclidian space.

**Fig. 2.** Plots illustrating the contribution of each pause type to the first two axes of variation for principal component analysis (PCA) (left) and correspondence analysis (CA) (right). The direction of each arrow indicates to which axes each pause position contributes, and in what direction, while the lengths of the arrows indicate the strength of each pause type's contribution.



**Fig. 3.** Plots illustrating the first two principal component analysis (PCA) scores (left) and correspondence analysis (CA) scores (right) for 41 Shakespeare texts. Each point represents an individual text; distances among texts reflect differences in pause position composition. In the CA plot (right), the size of each point represents the weighting assigned to each text in the analysis. The same five titles have been highlighted in each plot.

One can interpret the position of the play points by looking back at the axes of variation in Figure 2. Texts are positioned in relation to where pauses occur in their pentameter lines, and in what proportion. For example, plays like *Romeo and Juliet* and *Richard III* have relatively more pauses early on in lines (i.e., at positions 1 through 5), while plays like *Macbeth* and *Tempest* have more late-position pauses (6th through 9th). The circles in the right-hand panel in Figure 3 (the results of the correspondence analysis) have been scaled to reflect the relative weights of the data: plays with low total pause counts appear as slightly smaller circles; those with more pauses appear as larger circles. This is one advantage of correspondence analysis: as opposed to principal component analysis, correspondence analysis allows one to recognize differences in the amount of data contributed.

**4 Bootstrap Methodology**

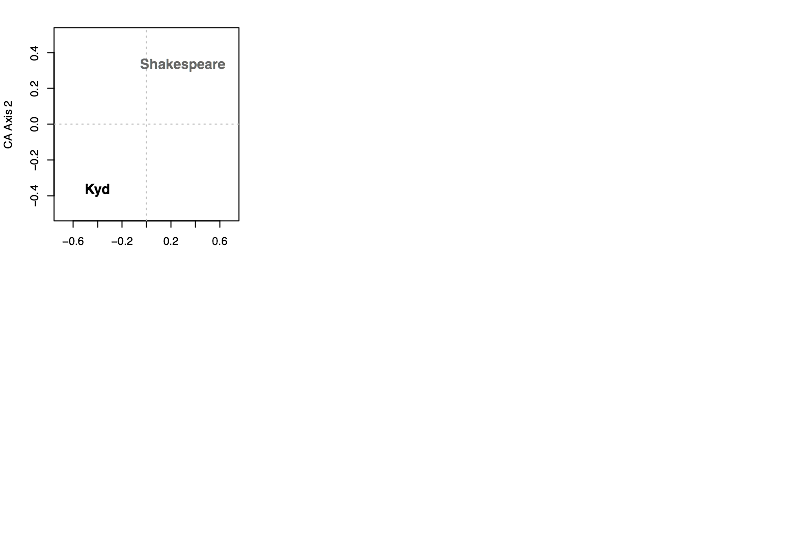
In most statistical analyses it is desirable to describe not only average patterns, but also variability. When researchers can return to a population and resample it, they may gain powerful insights. Or, when sampling from populations assumed to follow a normal distribution, they can rely on measures such as standard deviation or standard error to measure confidence in their estimates. However, we have only one set of counts for each play, just as an archaeologist may only have one set of pottery fragments from a given location. To understand how small differences in our observed data may influence the outcome of our analysis, therefore, we employ a "bootstrap" method. This too has a long history of use in archeological studies (Ringrose, 1992). The bootstrapping procedure is essentially a method of resampling. When one cannot resample from the population of interest, one resamples from the data available. In our case, that means taking random samples of pauses from each play and re-running our correspondence analysis using the new values (Lockyear, 2012; Peeples and Schachner, 2012). We retain the same sample size for each play, but we sample with replacement (meaning some pauses from the original data will be sampled more than once, and others not at all); this results in new counts of each pause type that vary slightly from the original counts. We repeat this resampling 1000 times, performing the correspondence analysis with each new set of counts. This affords us some measure of uncertainty for our correspondence analysis scores--"uncertainty" in this case being a desirable thing, as it can suggest confidence intervals for the results.

Figure 4 displays the results of a bootstrap procedure for 39 of the 41 texts treated in our correspondence analysis. Results for the *Sir Thomas More* pages and the Additional Passages to the 1602 *Spanish Tragedy* have been omitted; because these texts feature so little data, their oversized convex hulls (the shaded polygons representing each title) awkwardly dwarf the chart. As an illustration, we would note that the text with the next fewest amount of pauses, *The Merry Wives of Windsor* (with 240), is responsible for the large polygon in the lower center of the chart: because it has less information to analyze, a bootstrapping procedure repeated 1000 times produces considerably more variation. The polygons for the remaining 38 texts, however, trace a gradual arc up and to the right. We have called out five canonical texts--from *Richard III* and *Romeo and Juliet* through *Hamlet*, *Macbeth*, and *The Tempest*--to show how this data reveals the chronological progression of Shakespeare's works.



**Fig. 4.** Bootstrapped data for Shakespeare's plays. **(**Not shown: *Sir Thomas More* (selections); Additional Passages to *Spanish Tragedy* 1602.)

The "crescent" of Shakespeare's verse pauses also provides a basic template for understanding the syntactical development of his contemporaries' iambic pentameter. Figure 5, for example, charts the pause profiles of six contemporary playwrights over and against those of Shakespeare's works (from Fig. 4, shaded light gray). We can see, in the plays of his contemporaries, the general movement plotted by his verse, from Kyd and Marlowe through Fletcher, with Jonson's less iambic practice (bottom row, center) demarcating him as an exception. We should note that Marston's polygons fall almost entirely within Shakespeare's, a result that is not surprising, given the fact that Marston began and ended his career as a playwright while Shakespeare was still working, and appears to have fashioned his plays (including the *Antonio* plays, and *The Malcontent*) strongly in response to the senior playwright's (Cathcart).

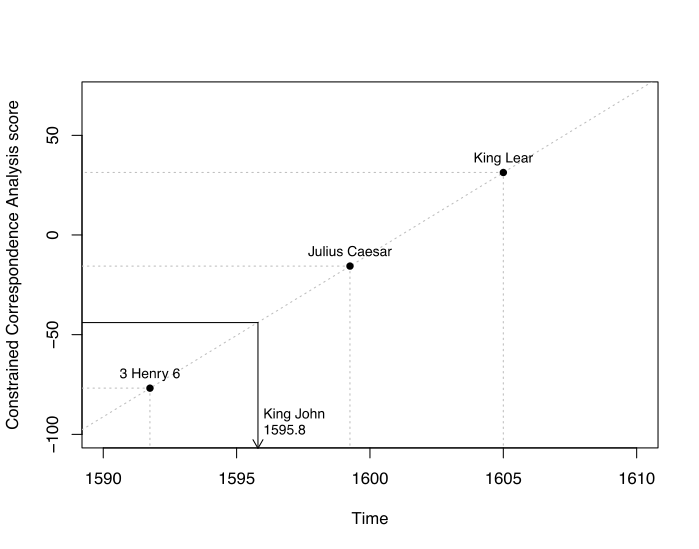


**Fig. 5.** Bootstrapped data for Shakespeare and six contemporaries. See Appendix B for titles and data.

**5 Constrained Correspondence Analysis (CCA) and Shakespeare’s Chronology**

Correspondence analysis, we should point out, provides relative ordering: object X most likely comes before, or follows, object Y, at Z distance. Not being content with simply obtaining a relative chronological order, we were interested in using external evidence about the plays to make specific determinations for each play. An extension of correspondence analysis called "constrained correspondence analysis" (CCA), allows us to do just that (Groenen and Poblome, 2003). By incorporating such information as interval constraints for Shakespeare's career, as well as exact dates for some plays, and upper and lower limits on dates for others, we can constrain the calculation of the correspondence analysis scores (van de Velden et al., 2009). As we saw earlier, correspondence analysis produces only a single score for each play. The same is true of constrained correspondence analysis. We opted, again, to employ a bootstrapping procedure, which allowed us not only to estimate exact dates for each play, but also to generate confidence intervals around those estimates. In this manner we were able to produce a revised chronology of Shakespeare’s plays, using only interval constraints, a few dates, and the pause position data itself.

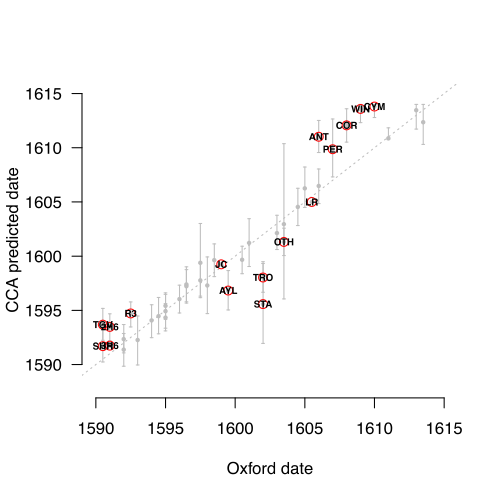
To constrain our correspondence estimates, we assigned numerical values to three plays for which plausible dates could be suggested: 3 *Henry VI*, last quarter of 1591 (= 1591.75); *Julius Caesar*, first quarter of 1599 (= 1599.25); *Pericles*, beginning of (= 160?.0). Scholars could argue over these designations, of course, and other plays and dates could have been employed; these seemed among the more reasonable of our options. In addition, we set upper and lower bounds on the extent of Shakespeare's writing career, demarcating it from 1589.5 to 1614.0. While these boundaries are also open to debate, it seemed to us that they are defensible so long as they are understood to be judgements rather than facts.



**Fig. 6.** Illustration of constrained correspondence analysis (CCA), modified from van de Velden et al. 2009. *King John* is dated using three "anchor" plays (indicated along the diagonal) as constraints. The regression between CCA score and time may be used to predict dates for the remaining plays. For a description of the method see also Groenen and Poblome (2007).

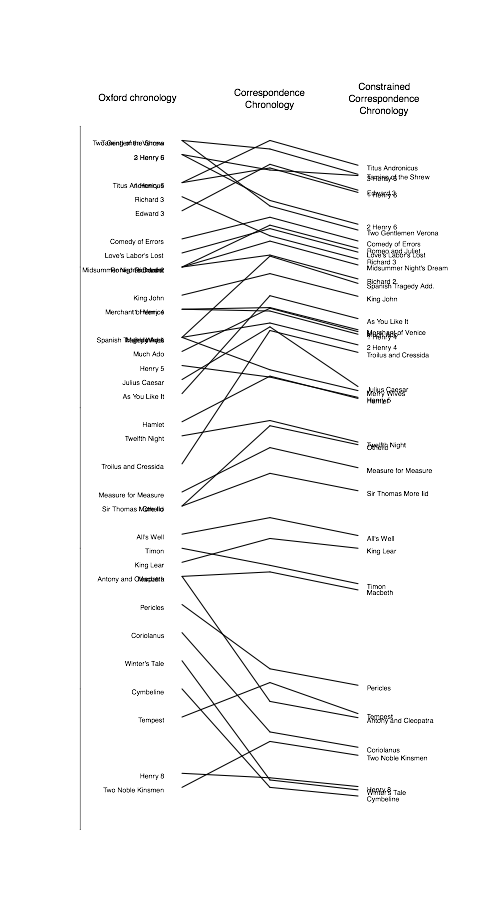
Figure 6 illustrates how constrained correspondence analysis (CCA) works: the relative positioning of an object (in this case, a text) is made concrete through the addition of specific information regarding other objects in the timeline. Thus *King John* is assigned a date of 1595.8, or October of 1595, as a manifestation of its statistical distance from 3 *Henry VI*, *Julius Caesar*, and *Pericles*. Obviously, the accuracy of such a "forcing" method depends in part on the soundness of these anchors, yet the procedure has the advantage of providing a specific date, rather than relative position, for each play.

Using a modified version of code from Poblome et al. (Poblome et al., 2009) we performed the constrained correspondence analysis (CCA) and added a bootstrap analysis to it in MATLAB, a software program dedicated to numerical analysis. We present the results in two ways. First, in Figure 7, we show how the predicted dates fit with the Oxford chronology dates. The diagonal line rising from bottom left to top right indicates the Oxford chronology (which does not include, we should point out, *Edward III* or the Additional Passages to *The Spanish Tragedy*). The circles represent play texts, the lines the confidence interval of the data. To see where a play falls in relation to the timeline, one can draw a horizontal line through the center of its circle back to the diagonal. For some plays, it will be seen, our prediction falls earlier than the Oxford date (e.g. *Troilus and Cressida*, *As You Like It*); in other cases it is later (e.g. *Coriolanus*, *The Winter’s Tale*).



**Fig. 7.** CCA results plotted against Oxford chronology\*.The diagonal dotted line indicates the Oxford chronology (which does not include *Edward III* or The Additional Passages to *The Spanish Tragedy*). Vertical bars represent texts; those above the line suggest a date later than the Oxford chronology; those below it, an earlier date.

Figure 8 illustrates three chronologies: from left to right, the Oxford chronology; the correspondence chronology (CA); and the constrained correspondence chronology (CCA). The latter two are based, once again, on Oras's pause data, and on new pause data provided for collaborative works.



**Fig. 8** Oxford chronology (left); the correspondence analysis chronology (CA) (center); and constrained correspondence chronology (CCA) (right).

During the drafting of this essay, we gained access to new data: the metrical and linguistic tables of Marina Tarlinskaja's monograph, *Shakespeare and the Versification of English Drama, 1561-1642* (Tarlinskaja, 2014). There Tarlinskaja provides a massive tabulation of original prosodic data for numerous playwrights, including Shakespeare. With Tarlinskaja's permission, we ran a principal component analysis (PCA) on the Shakespeare data, which does not include various prose-heavy plays (such as *Merry Wives*, *Much Ado*, *As You Like It*, and *Twelfth Night*), or two collaborative works (*Sir Thomas More*, *Timon*). Further, Tarlinskaja offers only partial data for *Edward III*, so we chose to omit it. That left a total of 34 texts, with over 1,500 pieces of data regarding 45 stylistic and structural categories. These categories include data on: strong and weak syllabic positions, respectively; word boundaries (total and after positions); strong syntactic breaks; run-on lines; proclitic and enclitic stresses; syllabic -ed and -eth; disyllabic -ion; grammatical inversion; meter-sense; syntactical run-ons; feminine endings (total); feminine endings built by simple and compound constructions, respectively; and alliterative lines.

Because Tarlinskaja’s data are rendered largely in percentages, the concerns we have about Oras's artificial constraints on the data obtain here as well. Yet the information struck us as too valuable to go unaddressed. We wish to emphasize that the results are purely our own, rather than Tarlinskaja's, who should not be assumed to endorse any of the datings we adduce from analyzing her data. Further, our analysis of her figures assumes the same intervals for Shakespeare's career (1589.5 to 1614.0) as in the constrained correspondence analysis (CCA). We record a principal component analysis of her figures in the fifth numerical column of Table 1 below, following, respectively, dates from the Oxford chronology, the Riverside chronology, our constrained correspondence analysis of the Oras data, and figures from Brainerd’s multivariate chronology (Brainerd, 1980). Following the principal component analysis of the Tarlinskaja data is a column with our CCA figures adjusted for playhouse closures (as discussed in the following section). The three asterisks in the CCA column flag the dates that we provided (for *3H6*, *JC*, and *PER*, respectively), rather than data generated from analysis.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **TABLE 1: Chronologies Compared** | | | | | | | |
| **title** | **Oxford** | **Riverside** | **Bruster-Smith CCA** | **Brainerd** | **PCA Tarlinskaja** | **Bruster-Smith Adjusted CCA** | **title** |
| **SHR** | 1590-1 | *1593-4* | 1591.3 | 1591.4 | *1594.1* |  | **SHR** |
| **TGV** | 1590-1 | 1594 | 1593.3 | 1592.9 | 1594.9 |  | **TGV** |
| **3H6** | 1591 | 1590-1 | 1591.8\* | 1591.0 | 1591.6 |  | **3H6** |
| **2H6** | 1591 | 1590-1 | *1593.1* | 1591.2 | *1592.4* |  | **2H6** |
| **TA** | 1592 | 1593-4 | *1590.2* | 1592.6 | *1590.5* |  | **TA** |
| **1H6** | 1592 | 1589-90 | 1593.5 | 1592.7 | 1592.3 |  | **1H6** |
| **R3** | 1592-3 | 1592-3 | *1594.4* | *1596.1* | *1595.2* |  | **R3** |
| **E3** | n/a | 1592-5 | 1591.8 | n/a | n/a |  | **E3** |
| **ERR** | 1594 | 1592-4 | 1593.7 | 1594.8 | 1593.5 |  | **ERR** |
| **LLL** | 1594-5 | 1594-5 | 1594.1 | 1601.7 | 1593.2 |  | **LLL** |
| **ROM** | 1595 | 1595-6 | *1594.0* | 1596.3 | *1592.6* |  | **ROM** |
| **MND** | 1595 | 1595-6 | *1594.6* | 1595.9 | *1592.6* |  | **MND** |
| **R2** | 1595 | 1595 | 1595.1 | 1595.9 | 1594.5 |  | **R2** |
| **JN** | 1596 | 1594-6 | 1595.8 | 1599.2 | 1592.9 |  | **JN** |
| **MV** | 1596-7 | 1596-7 | 1597.0 | *1598.3* | *1598.2* |  | **MV** |
| **1H4** | 1596-7 | 1596-7 | 1597.2 | 1596.7 | 1594.1 |  | **1H4** |
| **2H4** | 1597-8 | 1598 | 1597.5 | 1599.5 | 1597.6 |  | **2H4** |
| **WIV** | 1597-8 | 1597 | 1599.2 | 1596.1 | n/a |  | **MW** |
| **SPT** | n/a | n/a | 1595.5 | n/a | 1598.5 |  | **STA** |
| **ADO** | 1598 | 1598-9 | *1597.0* | *1596.9* | n/a |  | **ADO** |
| **H5** | 1598-9 | 1599 | 1599.5 | 1598.0 | 1598.5 |  | **H5** |
| **JC** | 1599 | 1599 | 1599.3\* | 1598.8 | 1599.3 |  | **JC** |
| **AYL** | 1599-00 | 1599 | 1596.5 | 1600.3 | n/a |  | **AYL** |
| **HAM** | 1600-1 | 1600-1 | 1599.5 | *1604.7* | *1602.2* |  | **HAM** |
| **TN** | 1601 | 1601-2 | 1601.1 | 1601.4 | n/a |  | **TN** |
| **TRO** | 1602 | 1601-2 | *1597.8* | *1600.8* | 1601.1 |  | **TRO** |
| **MM** | 1603 | 1604 | 1602.0 | 1604.7 | 1604.3 |  | **MM** |
| **OTH** | 1603-4 | 1604 | 1601.2 | 1603.3 | 1603.2 |  | **OTH** |
| **STM** | 1603-4 | 1594-5 | 1603.2 | n/a | n/a |  | **STM** |
| **AWW** | 1604-5 | 1602-3 | 1604.6 | 1607.2 | 1606.3 |  | **ALL** |
| **TIM** | 1605 | 1607-8 | 1605.2 | 1604.7 | n/a |  | **TIM** |
| **LR** | 1605-6 | 1605 | 1605.0\* | *1606.2* | *1607.0* |  | **LR** |
| **MAC** | 1606 | 1606 | 1606.5 | 1606.1 | 1607.3 |  | **MAC** |
| **ANT** | 1606 | 1606-7 | *1611.0* | *1607.9* | *1608.2* |  | **ANT** |
| **PER** | 1607 | 1607-8 | 1610.0 | 1604.2 | 1605.8 |  | **PER** |
| **COR** | 1608 | 1607-8 | *1612.3* | 1604.8 | *1610.0* |  | **COR** |
| **WIN** | 1609 | 1610-11 | *1614.0* | 1609.4 | *1612.4* |  | **WIN** |
| **CYM** | 1610 | 1609-10 | 1614.0 | 1608.9 | 1610.8 |  | **CYM** |
| **TMP** | 1611 | 1611 | 1610.7 | 1610.0 | 1613.5 |  | **TEM** |
| **H8** | 1613 | 1612-3 | 1614.0 | 1607.4 | 1612.7 |  | **H8** |
| **TNK** | 1613-4 | 1613 | *1612.7* | 1605.5 | *1612.8* |  | **TNK** |

**6 Adjusted CCA Chronology**

The final numerical column in Table 1 above adjusts the initial CCA results in relation to historical data concerning the disposition of the Elizabethan playhouses. J. Leeds Barroll has posited that Shakespeare would have reduced his dramatic writing or ceased writing new plays altogether when the playhouses were closed owing to the plague. Such a period occasioned the years 1592-94, when, as is well known, Shakespeare penned *Venus and Adonis* and *The Rape of Lucrece* when the public amphitheaters where closed for business. Because a noticeable spread in the data appears after *3 Henry VI* and another prior to *The Merry Wives*—one following our first “anchor” play, another roughly preceding our second—we have elected to adjust the CCA plotting in relation to this early playhouse closure.

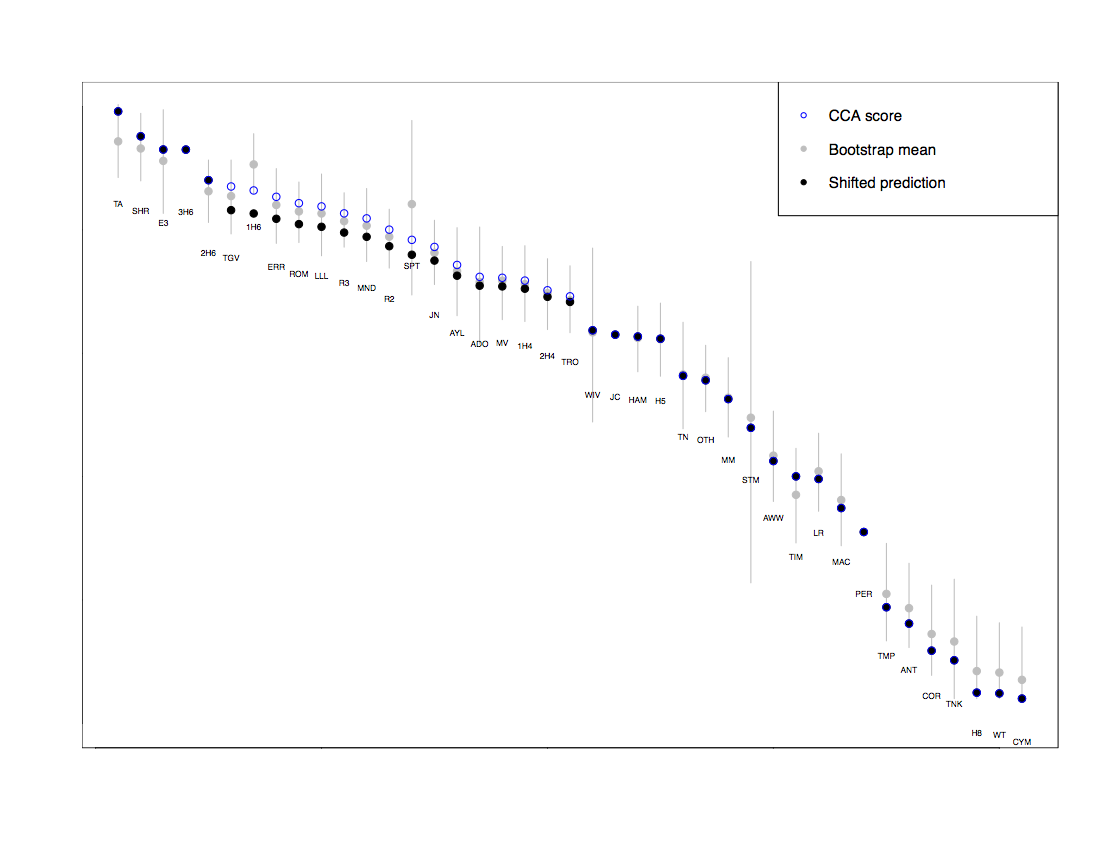


Fig. 9. Timeline of adjusted CCA. The clear circles represent the CCA predictions, the gray circles the bootstrap means, and the dark circles the adjusted CCA predictions, with *The Two Gentlemen of Verona* slotted in after the reopening of the playhouses. The gray bars indicate periods in which the playhouses were closed, largely owing to plague.

**7 Dates and Discussion**

The discussion of our results below employs the following format: title; hypothesized date; bootstrap range. Abbreviations keyed to Table 1 include O = Oxford; R = Riverside; BSCCA = Bruster-Smith Constrained Correspondence Analysis; B = Brainerd (Brainerd, 1980); PCAT = Principal Component Analysis of Tarlinskaja (Tarlinskaja, 2014).

***Titus Andronicus***--March 1590. We believe Shakespeare's first play was this collaboration with George Peele. As with other works, our CCA treats only the portion currently ascribed to Shakespeare (see Appendix [X] for breakdowns). This designation concurs with the PCA of Tarlinskaja's data, as well as with Slater's order. Brainerd falls evenly between Oxford and Riverside.

***The Taming of the Shrew***--June 1591. Our CCA concurs with Oxford and Brainerd. We believe this play was written before the closing of the theaters, and preceded *A Shrew*.

***3 Henry VI***--October 1591. Our CCA fixed this title in late 1591 (1591.75), but this position is unusually well supported by Oxford, Riverside, Brainerd, and the PCA of Tarlinskaja's data. Placement before 2 *Henry VI*--a counter-intuitive positioning attested by Brainerd, the Tarlinskaja PCA, Langworthy, and Reinhold--is most likely because 2 *Henry VI*, though written first, was later revised (and perhaps more extensively than 3 *Henry VI*).

***Edward III***--December 1591. Our CCA treats the portions currently assigned to Shakespeare. We see it as written before the closing of the theaters, but Slater's rare-word linkage of "part A" (1.2, 2, 4.4) with *Lucrece* could suggest placement during the hiatus or immediately after the reopening.

**1 *Henry VI***--May 1592. Issues of collaboration make this play an unlikely candidate for accurate placement, but stylistically our bootstrap CCA--in the single instance of this procedure altering rank order--places it just before the closing of the theaters. Regular CCA would place it just after *2H6*, Reinhold tied with it.

**2 *Henry VI***--[March 1594?]. As mentioned above, stylistically 2 *Henry VI* postdates 3 *Henry VI*, most likely owing to revision. (We see such signs most obviously in 2.1.1-4.1.147 and 5.1.1-end, though there are indications of revision throughout). It was probably written prior to the closing of the theaters, and revised later in the 1590s. Its placement here, then, likely averages writing from earlier and later in Shakespeare's career.

***The Two Gentlemen of Verona***--June 1594. Our adjusted CCA (adjusted to recognize the availability of the theaters) is confirmed by the PCA of Tarlinskaja's data. Slater's rare-word list ranks it seventh of the plays, Reinhold's data ninth--just prior to *Love's Labor's Lost*.

***The Comedy of Errors***--October 1594. Our adjusted CCA has a surprisingly tight match with the PCA of Tarlinskaja's data. This date fits well with the recorded performance at Gray's Inn on 28 December 1594.

***Romeo and Juliet***--January 1595. Stylistically, this first of the lyric plays perhaps finishes earlier owing to its formal, amatory verse (our unadjusted CCA places it in 1594). The PCA of Tarlinskaja's data also indicates an earlier date.

***Love's Labor's Lost***--February 1595. Our adjusted CCA squares with Oxford and Riverside. The PCA of Tarlinskaja's data indicates an earlier date.

***Richard III***--April/May 1595. Our hand-fit CCA agrees with the PCA of Tarlinskaja's data and with Brainerd in positing a later date for this text than Oxford or Riverside.

***A Midsummer Night's Dream***--July 1595. Our hand-fit CCA agrees with Oxford and Riverside. As with the lyrical plays generally, the PCA of Tarlinskaja's data indicates an earlier date.

***Richard II***--December 1595. This play was dated 1595 in four of the chronologies, with the PCA of Tarlinskaja's data only a half year earlier.

**The Additional Passages to *The Spanish Tragedy***--April 1596. We believe that these lines were penned after *Dream*, and in the vicinity of *Much Ado* and *Merchant*. The Tarlinskaja PCA puts them later than our dating, in the middle of 1598; there is too little data to be confident about either placement.

***King John***--July 1596. This play has been problematic since Honigmann's argument for an extremely early composition. Our positioning agrees with Oxford and Riverside; Brainerd has a later date, the Tarlinksaja PCA an earlier one, though after *Romeo* and *Dream*.

***As You Like It***--February 1597. This first of the prose-heavy plays (with fewer pauses for analysis) suggests an earlier date than is conventionally accepted. Brainerd's dating better accords with Oxford and Riverside; Tarlinskaja does not have data for it. We would point out only that there have been various arguments for a date earlier than 1599. The adjusted CCA here could reflect multiple dates of composition.

***The Merchant of Venice***--May 1597. Our CCA places this text at the beginning of 1597; both Brainerd and the Tarlinskaja PCA locate it in 1598. Barely distinguishable, statistically, from *Much Ado*, it dates slightly earlier on the bootstrap procedure.

***Much Ado About Nothing***--July 1597. Brainerd's date is very close to the CCA slotting. Our positioning is slightly earlier than Oxford and Riverside.

**1 *Henry IV***--October 1597. Our dating accords closely with Oxford, Riverside, and Brainerd. The Tarlinskaja PCA's earlier date (1594.1) is odd, though this figure would have it follow *Romeo*, *Dream*, *Errors*, *Love's Labor's Lost*, and *John* in that test.

**2 *Henry IV***--January 1598. We believe this play immediately followed 1 *Henry IV*. Our dating accords very closely with Oxford, Riverside, and the Tarlinskaja PCA.

***Troilus and Cressida***--April 1598. This is a significant repositioning of a play typically dated some two to four years later. Brainerd and the Tarlinskaja PCA would place it toward the end of 1601, and Oxford and Riverside a year or so later than that. A 1598 composition would locate it close to the publication of Chapman's Homer, and prior to the War of the Theaters, Gilbert's *De Magnete*, and Nashe's *Summer's Last Will and Testament*, which are sometimes seen as implicated in its language. Langworthy places *Troilus* after *Henry V* and before *Hamlet*, the Reinhold data after *Twelfth Night* but before *Othello* and *Hamlet*.

***The Merry Wives of Windsor***--January 1599. Our dating comes later than Oxford and Riverside, which, like Brainerd, locate the play in the 1596-1598 period. Scant pause data makes placement less certain.

***Julius Caesar***--March 1599. The second of our three "anchor" plays (1599.25), *Julius Caesar* was seen by Platter in September of this year, and apparently quoted in *Every Man Out of His Humor*, which Bednarz dates at . There is good agreement on this position: in addition to Oxford and Riverside, Brainerd has this play at the end of 1598, the Tarlinskaja PCA in April of 1599.

***Hamlet***--June 1599. Statistically indistinguishable from *Henry V*, this play's references to Julius Caesar suggest that it may have followed immediately upon the Roman tragedy. Our dating is slightly earlier than Oxford and Riverside; both Brainerd and the Tarlinskaja PCA have it later than those authorities. Several passages in the Folio text hint at revision during the early 1600s.

***Henry V***--August 1599. Almost certainly composed prior to Essex's disastrous return from Ireland on September 28. A date in late summer would just enable it to inspire certain passages in 1 *Sir John Oldcastle*, completed by 16 October.

***Twelfth Night***--March 1601. This play produced a very close match among our PCA and Riverside, Oxford, and Brainerd.

***Othello***--May 1601. Our CCA locates *Othello* earlier than Oxford, Riverside, Brainerd, and the Tarlinskaja PCA. The latter two tests place it in 1603. Echoes in Q1 *Hamlet*, however, indicate that it was already extent, even familiar, by 1603.

***Measure for Measure***--February 1602. Like *Troilus*, *Measure* finished earlier in the CCA than in all other tests. Commonly thought of as a Jacobean play owing to various of its themes, *Measure* features very little that dates it certainly. Pushing it forward a year or more (i.e. to 1603-04) would imply a correspondingly later date for the texts immediately following it (*Sir Thomas More* additions; *All's Well*; *Lear*). Revisions by Middleton may skew the results here.

***Sir Thomas More* Additions**--March 1603. This dating squares with Oxford, and has been argued as well by Jackson (Jackson, ????).

***All's Well That Ends Well***--August 1604. Our CCA squares with Oxford but is later than Riverside. Brainerd and the Tarlinskaja PCA suggest a later date (1607.2 and 1606.3, respectively), as does Reinhold's data, which puts this play after *Lear* and before *Macbeth*.

***King Lear***--January 1605. This is the third of the three "anchor" plays on which we fix a constraint (1605.0). Oxford suggests 1605-06, Riverside 1605, Brainerd 1606.2 and the Tarlinskaja PCA 1607. It likely preceded the second edition of *King Leir* in May of 1605.

***Timon of Athens***--March 1605. This collaboration with Thomas Middleton is obviously connected to *King Lear*, with which it has the highest number of rare-word links (Slater). Our CCA puts it just after *Lear*, though the figure from the bootstrap (1606.2) procedure slides it closer to *Macbeth*.

***Macbeth***--July 1606. Our CCA agrees closely with Oxford, Riverside, and Brainerd. (The Tarlinskaja PCA places it almost a year later).

***Pericles***--April/May 1608. Beginning with this text, our CCA almost consistently produces dating later than is conventional. *Pericles* finished at 1610, almost two years later than it was entered in the Stationers' Register, and a year after it was published in quarto. The unusual textual situation with this play may be responsible for the variation in the tests: the Tarlinskaja PCA puts it just before 1606; Brainerd's test (which does not recognize co-authorship) puts it at 1604.2. The date provided here comes just prior to its entry in the Stationers' Register (20 May).

***The Tempest***--October 1610. Most likely based on sources available only from September of 1610, *Tempest* finished slightly earlier (1610.0) in our CCA; it is last on Brainerd's list, and later, and also last, in the Tarlinskaja PCA (1613.5). Slater's rare-word ranking (p. 99) lists it last as well. Our dating just precedes that of Oxford and Riverside.

***Antony and Cleopatra***--March 1611. This title was entered in the Stationers' Register 20 May 1608. The fact that this text dates later in our CCA may indicate a revision--perhaps an expansion--of a script subsequent to its entry. Brainerd and the Tarlinskaja PCA place it around the beginning of 1608. Slater's rare-word catalogue links it with not only *Macbeth* and *Coriolanus*, but also *Cymbeline*, *Winter's Tale*, and *Tempest*.

***Coriolanus***--April 1612. As with the other late plays, our CCA produces a later-than-conventional date. The Tarlinskaja PCA suggests 1610, later than Oxford (1608) or Riverside (1607-08). Slater's rare-word index links it most tightly with *Cymbeline* and *Winter's Tale*. Langworthy places it after *Antony*.

***The Two Noble Kinsmen***--September 1612. This dating is supported almost exactly by the Tarlinskaja PCA, and conforms to both Oxford and Riverside. Brainerd (who does not separate according to collaboration) has it much earlier (1605.4).

***Henry VIII***--May 1613. According to our CCA, it follows *Kinsmen* but precedes two of the romances. Certainly composed prior to late June 1613, when it was performed at the Globe. Of the late plays, it has statistically significant rare-word links with only *Winter's Tale*. Our bootstrap CCA places it at the end of May 1613.

***Winter's Tale***--August 1613. Brainerd's 1609.4 accords with Oxford, and just precedes Riverside. The Tarlinskaja PCA places this play in the middle of 1612. Langworthy has *Winter's Tale* and *Cymbeline* as two of his last three texts (with *Henry VIII*).

***Cymbeline***--October 1613. The Tarlinskaja PCA has this text just before 1611. Rare-word links with *Winter's Tale*, *Tempest*, and *Coriolanus* group it with those plays. Our CCA suggests that it is Shakespeare's final piece of writing.

**8 Conclusion**

**References**

**Barroll, L.** (1991). *Politics, Plague, and Shakespeare’s Theater: The Stuart Years*. Ithaca, NY: Cornell University Press.

Bathurst, C. (1857). *Remarks on the Differences in Shakespeare's Versification in Different Periods of his Life and on the Like Points of Differences in Poetry Generally*. London: John W. Parker and Son.

Brainerd, B. (1980). The Chronology of Shakespeare’s Plays: A Statistical Study. *Computers and the Humanities*. 14.4: 221-30.

Cathcart, C. (1997). John Marston and the Professional Drama, 1598-1608. Ph.D. Thesis, Department of English and American Studies. University of Manchester.

**Evans, G. B., ed.** (1974; rev. 2nd ed. 1996). *The Riverside Shakespeare*. Boston: Houghton Mifflin.

**Greenacre, M., 2007.** *Correspondence Analysis in Practice*. 2nd ed. London: Chapman & Hall/CRC.

Groenen and Poblome (2007) and

**Groenen, P.J.F., and J. Poblome.** 2003. Constrained correspondence analysis for seriation in archaeology applied to Sagalassos ceramic tablewares. In: Schwaiger, M., Opitz, O. (Eds.), *Exploratory Data Analysis in Empirical Research*. Berlin: Springer: 90-97.

**Jackson, MacD. P.** (2002). Pause Patterns in Shakespeare's Verse: Canon and Chronology. *Literary and Linguistic Computing*. **17.1**: 37-46.

**Kuhn, T.** (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

**Lockyear, K.** 2012. Applying bootstrapped Correspondence Analysis to archaeological data. *Journal of Archaeological Science* **xx**: 1-10.

**MATLAB.** 2011. Version 7.13.0. The MathWorks Inc. Natick, Massachusetts, USA.

**Nenandic, O., and M. Greenacre.** 2007. Correspondence analysis in R, with two- and three-dimensional graphics: The ca package. Journal of Statistical Software **20**: 1-13.

**Oras, A.** (1960). *Pause Patterns in Elizabethan and Jacobean Drama*. Gainesville, FL: University of Florida Press.

**Peeples, M.A., and G. Schachner.** 2012. Refining correspondence analysis-based ceramic seriation of regional data sets. *Journal of Archaeological Science* **39**: 2818-2827.

Poblome et al. (2009)

**R Core Team.** 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

**Ringrose, T.J.** 1992. Bootstrapping and Correspondence Analysis in Archaeology. Journal of Archaeological Science **19**: 615-629.

**Saintsbury, George.** (1906). *A History of English Posody: From the Twelfth Century to the Present Day*. 3 vols. London: Macmillan and Co.

**Tarlinskaja, M**. (1976). *Shakespeare’s Verse: Iambic Pentameter and the Poet’s Idiosyncrasies*.

**Tarlinskaja, M.** (2014). *Shakespeare and the Versification of English Drama, 1561-1642*. Burlington, VT: Ashgate.

**van de Velden, M., P.J.F. Groenen, and J. Poblome.** 2009. Seriation by constrained correspondence analysis: A simulation study. Computational Statistics and Data Analysis **53**: 3129-3138.

**Wells, S. and G. Taylor, with J. Jowett and W. Montgomery** (1987). *William Shakespeare: A Textual Companion*. Oxford: Clarendon Press.

**[Appendix 1 Abbreviation of Play Titles]**

**Appendix 2 Pause Data**

**Appendix 3 Source Texts and Play Breakdowns**