Thomas Kyd: Two New Books

Thomas Kyd: A Dramatist Restored, by Brian Vickers, Princeton, Princeton University Press, 2024, xxviii + 368 pp., £35 (hardcover), ISBN 9780691211602
Shakespeare's Tutor: The Influence of Thomas Kyd, by Darren Freebury-Jones, Manchester, Manchester University Press, 2022, x + 238 pp., £80 (hardcover), ISBN 9781526164742

Brian Vickers is the General Editor and Darren Freebury-Jones is the Associate Editor of a new *Collected Works of Thomas Kyd* edition, the first volume of which was published by Boydell and Brewer in 2024. Until recently the widely accepted dramatic canon of Kyd comprised just one play, *The Spanish Tragedy*, and only Kyd's most ardent fans would look beyond it to his minor poetical works and his play and prose translations. From the Stationers' Register entry of the book in 1594 it has always been clear that Kyd translated Robert Garnier's French-language play *Cornélie* into English as *Cornelia*, but the claim that Kyd wrote the anonymously published *Solimon and Perseda* – the story of which forms the play-within-the-play near the end of *The Spanish Tragedy* – gained significant critical traction only with the publication of Lukas Erne's book *Beyond 'The Spanish Tragedy'* in 2001.

The Kyd canon that Vickers and Freebury-Jones present is much larger, comprising *Verses of Praise and Joy*, the prose works *The Householder's Philosophy* (a translation from Torquato Tasso's Italian) and *Two Letters to John Puckering*, the play translation *Cornelia*, the two sole-authored plays widely accepted as Kyd's (*The Spanish Tragedy* and *Solimon and Perseda*), plus three anonymously published plays (*King Leir*, *Arden of Faversham*, and *Fair Em*) and two co-authored plays (*1 Henry VI* with Shakespeare and Thomas Nashe, and *Edward III* with Shakespeare). The primary aim of the two books reviewed here is to convince readers that there are good reasons to expand the Kyd canon to include these five additional plays: *King Leir*, *Arden of Faversham*, *Fair Em*, *1 Henry VI*, and *Edward III*. The chief conclusion of this review is that there are not.

The evidence for these authorship attributions consists of linguistic features found in the known writings of Kyd and in the newly attributed works. These features are habits of verse style – principally the use of feminine endings and the positioning of pauses within the line – and the presence of certain words and phrases. Vickers and

Freebury-Jones cite others' scholarship on these features and add new evidence consisting primarily of words and phrases shared by the known Kyd works and the newly attributed works.

The foundational book on feminine endings is Philip Timberlake's *The Feminine Ending in English Blank Verse* of 1931, which counted this feature in plays produced up to the year 1595. In his preface, Vickers misreports Timberlake's findings, remarking that Marlowe 'used a regular iambic line with a low incidence of feminine endings ranging from 0.5 percent to 3' of all lines, which he thinks is 'a statistic that removes any possibility of Marlowe as co-author of the *Henry VI* plays' (xvii). But looking at Timberlake's book shows this not to be true. On pages 39–41, Timberlake tabulated his figures for each scene in Marlowe's plays *Doctor Faustus*, *The Jew of Malta*, and *Edward II*, and then summarised them. In a short scene, even just a few feminine endings will give a high percentage values, as when the three such endings in the 31 lines of scene 4.4 of *The Jew of Malta* mean that 9.6% of the lines in this scene have this feature.

Timberlake separated out the percentages for scenes that have at least 100 lines, since these larger scenes are more representative of Marlowe's overall habit. Using what he called his strict counts (removing all questionable occurrences where the metre is uncertain), Timberlake reported that across all scenes of over 100 lines Marlowe's range in *Doctor Faustus* is '0.0–4.5 per cent', in *The Jew of Malta* it is '1.1–5.9 per cent', and in *Edward II* it is '0.5–8.0 per cent'. These ranges are far greater than the 0.5% to 3% Marlowe figures that Vickers reports from Timberlake's book. Vickers goes on to summarise Timberlake's figures for Kyd as follows: '*Soliman and Perseda* (1588) 10.2 percent; *King Leir* (1589) 10.8; *Arden of Faversham* (1590) 6.2; *Fair Em* (1590) 6.5; *Cornelia* (1594) 9.5' (xviii). These numbers come from the tables that Timberlake gives on his pages 46, 61, and 63, but Vickers omits to mention that the table on Timberlake's page 46 begins '*The Spanish Tragedy* 1.2'. That is, the one play that we all agree is by Kyd has, by Timberlake's counting, far less frequent use of feminine endings than the questionable ones that Vickers wants to add to the Kyd canon. Misrepresentations of Timberlake's work recur throughout Vickers's and Freebury-Jones's books, as we shall see.

The primary work on pause patterns in verse lines is Ants Oras's book *Pause Patterns in Elizabethan and Jacobean Drama* of 1960. Pauses can be of different kinds, from the weakest, where punctuation may or may not represent where the dramatist intended a pause, to the strongest, where a verse line is split between two speakers. A pause might occur between the first and second syllable, or between the second and third, and so on. Oras recorded each line's pause by the syllable, from first to ninth, that it occurs after, so that for each play he could state the proportion of lines containing a pause after the first syllable, the proportion containing a pause after the second syllable, and so on. Then he constructed for each play an x/y graph in which the horizontal axis runs from 1 to 9 (for the syllable after which the pause occurs) and the vertical axis runs from 0% to 100% showing what percentage of the lines of the play have a pause at that position.

As part of his argument for the anonymously published play *King Leir* being Kyd's, Vickers reproduces two graphs showing the pause patterns in Kyd's *The Spanish Tragedy* and *King Leir* (190). Two significant problems vitiate Oras's pictorial presentation of his data. The first is that for each play he had exactly nine data values (one for each possible pause position), so the horizontal scale is, in reality, discrete and categorical not continuous. That is, Oras had no data for the syllabic positions 1½ or 2¼ because there are no such syllabic positions. For this reason, his pictures should be

bar charts. But Oras instead put a dot at the vertical position representing the percentage of lines for each syllabic position and then joined the dots with straight lines, as if there were percentage values (on the *y* axis) for all syllable values (on the *x* axis), such as 1½ and 2¼ and every other real number between one and nine. The need to avoid this sort of misrepresentation of discrete categorical data as if it were continuous data is elementary mathematics.

The second problem with Oras's 'graphs' (which should be bar charts) concerns the vertical scaling. The *y* axis is different for different authors. The *y* axis ticks go up in units of 10% and are labelled right up to 80% for the early play *Gorboduc* but with Ben Jonson's works they are labelled only up to 30%. This would be acceptable if the physical distances on the page were scaled the same between authors – if the 10% units were the same height in every picture – but they are not. Thus, it looks like we can overlay one pause-pattern picture on another and visually confirm that the underlying data are alike, but we cannot do this because the pictures are not to the same scale. But at least with Oras's axes being labelled we can read off the data values from his pictures. When Vickers reproduces Oras's pause-pattern pictures for Kyd's *The Spanish Tragedy* and *King Leir* he silently rescales the latter so that its main peak reaches exactly as high up the page of his book as does the former, making them look virtually identical (190). Vickers omits the *y*-axis labels in his reproduction of Oras's pictures, so his rescaling is not obvious unless the reader consults Oras's book to see that the underlying numbers are in fact different, and it is impossible to read off the numbers themselves in Vickers's reproduction of the 'graphs'. Vickers's rescaling and omission of axes misrepresent Oras's data.

The art of finding words and phrases in common between different early modern texts has recently been enhanced by several technological developments. In the 1990s the company Chadwyck-Healey paid for hundreds of canonical works of English Literature to be typed into computers to produce a series of datasets on CD-ROM and magnetic tape that eventually were combined to make the online subscription website called *Literature Online* (*LION*), now owned by the ProQuest corporation and sold as *One Literature*. Around the same time, the non-profit *Text Creation Partnership* (*TCP*) paid for the keyboarding of thousands of texts from the *Early English Books Online* (*EEBO*) set of images of books in the A. W. Pollard and G. R. Redgrave *Short Title Catalogue*. *EEBO* is now a virtually complete set of images of books published in Britain up to the year 1700 (totalling around 145,000 books) and there exist searchable *TCP* transcriptions for around 60,000 of these.

In 2017, the independent scholar Pervez Rizvi compiled his own digital dataset of virtually all the early modern drama that has survived and he ran software on it to find, for each play, all the phrases that it has in common with each of the other plays. He called the website from which his data can be downloaded *Collocations and N-Grams* (*CAN*) and Rizvi has since added to it reports of his own experiments with this data, including explorations of the claims made by Vickers and Freebury-Jones for an expanded Kyd canon. Vickers misreports Rizvi's work too, as we shall see. Vickers also bizarrely misreports how he came to work with Freebury-Jones, recounting his examination of Darren Freebury-Jones's PhD thesis in 2016, in which Freebury-Jones had 'used Rizvi's database more thoroughly than' Vickers had (xxiv). This is impossible: neither of them could at this point have used Rizvi's *CAN* dataset since it was not created until 2017. Moreover, Freebury-Jones's PhD thesis is available for anyone to digitally download from the website of Cardiff University's library, and it is easy to see that it makes no mention of Rizvi or his database.

In his chapter on *Solimon and Perseda*, Vickers surveys the evidence that the play is by Kyd, starting with J. E. Routh's essay on its unusual patterns of rhyme, which Vickers dates to 1909 in his body text on page 117, 1905 in a footnote on that page, and 1903 on page 118. The footnote is correct. Also on page 118 Vickers wrongly dates Timberlake's book on feminine endings to 1936, having correctly given its date as 1931 on page xvii. Continuing this curious run of errors, Vickers writes 'As I explained in chapter 2, the regular blank verse line, the iambic pentameter, has a "masculine ending" with a stress on its final word' (119). There is no such explanation of masculine or feminine endings in his Chapter 2. Repeating the claims from his preface, Vickers gives Timberlake's figures for what percentage of Marlowe's verse lines have feminine endings, but this time using different numbers. In the preface he wrote that the group of University Wits including Marlowe 'used a regular iambic line with a low incidence of feminine endings, ranging from 0.5 percent to 3' (xvii) but here on page 119 he gives Marlowe's peak as '3.7'%, for *The Jew of Malta*. Timberlake took the trouble to break his counts down by scene precisely because the mean average of 3.7% for the whole of *The Jew of Malta* is deflated by seven of the play's scenes having no feminine endings at all. Timberlake's tabulation shows such wide variation between Marlowe's scenes – ranging from zero in many scenes to 12.5% in scenes from *Doctor Faustus* and *The Massacre at Paris* and 11.1% in *Edward II* – that mean averages by play make no sense. Why not use modal averages or medians and why should we average by play?

Any metric used for authorship attribution needs to be one that is characteristic of authorship in two particulars. It needs to give a relatively consistent value across different works by one author and to give distinctly different values for works by different authors. We expect that as we examine smaller and smaller units of writing any metric is likely to vary more and more from the mean value that we get for large units. Thus, for instance, across all his plays Shakespeare uses the word *and* about 280 times in every 10,000 words. In the whole of *As You Like It* he uses it 291 times in every 10,000 words, but if we take just the first act of *As You Like It* the frequency of *and* drops to 222 per 10,000 words. If we look at just the first scene of the play it rises again to 282 occurrences per 10,000 words, and if we look at just the first 500 words it falls again to 240 occurrences per 10,000 words. The smaller the sample, the more likely are we to find fluctuation around the mean derived from the whole Shakespeare canon. But in this case, the frequency of *and*, the fluctuations are not wild.

Because of this relative consistency and because the frequency of *and* is markedly different for other authors – Thomas Middleton's mean across all his plays is 220 per 10,000 words and George Peele's is 393 per 10,000 words – we can use this metric (alongside a set of others that behave the same way) in authorship attribution. It is logically invalid to use as a marker of authorship a measure that does not display this kind of consistency. There is nothing in Timberlake's data to show that mean average use of feminine endings is a consistent metric that is roughly the same across all of one author's works and is significantly different for each author. In other words, nothing Timberlake found makes mean average rates of feminine endings an author-specific marker that we can rely upon in attribution studies.

In his chapter on *Cornelia*, Vickers cites phrases that occur in this play and known Kyd works (such as his *The Householder's Philosophy* and *The Spanish Tragedy*) and the claimed Kyd works (*Solimon and Perseda* and *Arden of Faversham*), adducing these phrases as evidence of common authorship. But many of these phrases are common to thousands of other works too. For instance, it is true that *to him to* appears in *Cornelia* and *Solimon and Perseda* (140), but it also appears 17,903 times across all the books in *EEBO-TCP*. The essential question to be asked is how often we should

expect to find such a match between works by different authors and how often between works by the same author. Vickers offers nothing about this.

The phrase *once for all* occurs in *Cornelia* and *Solimon and Perseda* (144), but also over 4000 times in other works in *EEBO-TCP*. Likewise, *than the most* appears in *Cornelia* and *The Householder's Philosophy* (138) but also over 2300 times elsewhere. It is baffling that Vickers cites over 50 such shared phrases without quantifying anything about them. If Vickers could show that we should not expect as many as 50 such matches of commonplace phrases unless common authorship were the cause then he might be able to construct an attribution argument from this evidence, but he makes no attempt to do this. And, in this case, it would be pointless anyway, since Kyd's responsibility for *Cornelia* is already established by the Stationers' Register and the first printed edition naming him. Presumably, his reason for taking this approach is that when Vickers comes to use it again for plays that have not traditionally been attributed to Kyd the reader is supposed to see the likeness to the case for *Cornelia* and transfer her acceptance of one attribution to another.

Vickers's method for establishing Kyd's authorship of *King Leir* is the same as before: he looks for phrases in this play and in the accepted Kyd works. Some are relatively rare but others obviously not. Of the former kind is *never can be* followed by one word and then *but by* (187) which occurs in *King Leir* and *Solimon and Perseda* and only seven other books in *EEBO-TCP*. Of the latter kind is *exceed* followed within two words by *bounds* (185), which has over 1200 occurrences in *EEBO-TCP*. Vickers is quite right that 'First to the heavens, next, thanks to you, my sonne' found in *King Leir* is like 'First, thanks to heaven; and next to Brusor's valour' in *Solimon and Perseda* (178). But Vickers does not mention that it is also like 'first thanks to heaven, next to my forward countrymen' in the anonymous play *The True Tragedy of Richard the Third* published in 1594 and also like 'First thanks to heaven, and next to thee my friend' in the 1594 quarto of the play *The Contention of York and Lancaster*, usually attributed to Shakespeare and identified as an alternative version of his *2 Henry VI* published in the 1623 Folio collection. Again Vickers is right that 'And add fresh vigour to my willing limbs' found in *King Leir* is like 'And add fresh courage to my fainting limbs' in *Solimon and Perseda*, but he does not mention that it is also like 'add fresh strength to these my withered limbs' in the anonymous play *Guy of Warwick* first performed in the early 1590s, and like 'add fresh vigour to thy feeble limbs' from Robert Armin's play *The Valiant Welshman* first performed in the early 1610s. In the absence of a complete description of his experiments it is impossible to weigh the evidentiary value of Vickers's phrase matching results. But we can say that he is not presenting all the evidence for how often and where the phrases he adduces are to be found in early modern plays.

To attribute *Fair Em*, Vickers undertakes the same phrase-matching procedures as before, and as before some examples really do seem unique, for instance *in hope* followed by one word followed by *oath is true* in *Fair Em* and *The Spanish Tragedy* (218). But others are utterly common, for instance *the wrath of* which occurs in *Fair Em* and Kyd's *The Householder's Philosophy*, but also occurs 27,400 times in other books in *EEBO-TCP*. That Vickers thinks *the wrath of* is a 'truly unusual collocation' (221) indicates that he has not searched for it across the rest of early modern writing. It occurs, of course, many times in the Bible. The same thing happens with Vickers's searches for shared phrases to attribute *Arden of Faversham*. Some are rarities, such as *be it* followed by *spoke* or *spoken* followed by *in secret here*, and also *link* or *linked* followed by *in liking* (271); both of these seem genuinely unique to *Arden of Faversham* and *Solimon and Perseda*. But Vickers also lists commonplaces, such as *sit in* followed by one word and then *seat* (271), which appears over 900 times in

EEBO-TCP, and *cannon* or *cannon's* followed by up to five words and then *discharge* or *discharged*, which appears 184 times.

Vickers provides a weblink to his personal website where many more phrase matches between *Arden of Faversham* and Kyd's plays are promised (269n45), but there we again find that the useful ones are mixed with examples such as *is this the*, found twice in *Arden of Faversham* and twice in *The Spanish Tragedy*, but also found over 11,000 times in *EEBO-TCP*, and often twice or more in one work. This part of Vickers's book seems to have been printed from an incompletely revised typescript, since the phrasal matches discussed on pages 266–267 are discussed again, as if for the first time, on page 272. Similarly, this chapter's discursive footnote 31 is identical to its footnote 53.

In Chapter 8, 'Denying Kyd', Vickers addresses what he sees as MacDonald P. Jackson's mistaken and long-held adherence to the idea that Shakespeare wrote part of *Arden of Faversham*, which belief started with Jackson's Bachelor of Letters degree in 1963. In his thesis for that degree, Jackson considered and rejected the attributions made by Charles Crawford in 1903. Scandalously, Vickers suggests that Jackson's degree was improperly examined: ' … given the marginal status that attribution studies have had in English departments, Crawford's work was probably unknown, and it might have been asking too much of the examiners of a B. Litt. thesis to have checked whether Jackson had given a reliable account of his work' (281). Vickers's prime objection is that although Jackson knew of the prior scholarship on which he (Vickers) builds his lists of phrases found in *Arden of Faversham* and Kyd's plays, he (Jackson) dismissed them as largely drawn from 'a common stock of dramatic diction' (Jackson quoted in Vickers, 282) rather than being evidence of shared authorship. My own discoveries for this review of how common are many of the phrases cited by Vickers put me in essentially the same position as Jackson: I find that Vickers's evidence does not clinch the argument.

To counter Jackson's demonstration that there are many more rare phrasal matches between the middle act of *Arden of Faversham* and Shakespeare's plays than between that middle act and Kyd's plays, Vickers introduces a new explanation: one dramatist imitated the phrases of another. If this is true of Jackson's matches, it must also be a possible explanation for the phrasal matches that Vickers previously discussed as sure signs of Kyd's authorship. Vickers makes some objections to Jackson's experiments using *Literature Online* (*LION*), now called *One Literature*, and *EEBO-TCP*, and if these are accepted then they should apply equally to his own experiments. Jackson used rare collocations where up to 10 unrelated words are allowed to intervene between the words searched for. Vickers cites the correct limit as 'a maximum of four words intervening' (John Sinclair quoted in Vickers, 304). But a skim through Vickers's own collocation evidence shows him accepting longer gaps too, as when he treats *What ails you, woman … a sudden qualm* from *Arden of Faversham* as a collocation (266) although seven words intervene where I have put ellipsis. Likewise, Vickers uses *nips me … the … tender blossoms* from *Arden of Faversham* as a collocation (270) although again seven words go in the first ellipsis. Likewise, Vickers's collocation *lead … to the slaughterhouse* from *Arden of Faversham* (274) has eight words intervening, as does *earth … hope* from *The Spanish Tragedy* (276).

Vickers's understanding of the field of computational stylistics is weak. Introducing what he considers to be some recent correctives to the misuse of computational methods, Vickers claims that 'Computers … can identify a word as a unique sequence of letters or characters, a graphological unit, but they are unable to distinguish the various senses that it can have' (305). Experts in the field know this claim to be

untrue: what is called morphosyntactic tagging of texts by computers – for instance using the MorphAdorner software developed at Northwestern University – has for some time been able reliably to distinguish the semantic difference between sequences of letters based on their contexts. Thus, computers do now know when the sequence *r-o-w* means the noun for an argument rather than the noun for the opposite of a column or the verb for propelling a boat or the verb for having an argument. Vickers is not only ignorant of the studies he condemns, but also confidently asserts the impossibility of methods that are now established as the state of the art.

Vickers attempts to explain John Burrows's innovations in computational stylistics called Delta and Zeta, and it is clear that he does not understand what he is describing when he writes that 'Both basically counted the frequencies with which selected groups of words were repeated, with procedures added to promote accuracy' (305). The promotion of accuracy has nothing to do with it, since the one thing computers have always been able to do is count word frequencies perfectly. According to Vickers, Burrows and his successor in this approach Hugh Craig 'were unable to understand the mathematics' of the procedures they developed (306) and he extols what he sees as the correctives provided by David Auerbach, Nan Z. Da, and Rizvi, which he describes. Auerbach is quoted by Vickers rejecting the assumption 'that single word frequencies are sufficient to establish a high degree of confidence in authorial attribution' (308). In fact, no one need make this assumption since the proposition has been empirically tested and found to be true. We can 'blindly' apply authorship attribution methods to the many cases where we know the authorship of particular works and see how often each method picks the correct author. Multiple such 'blinded' studies have shown that methods based on single word frequencies achieve high levels of accuracy and these have subsequently been used in authorship attributions as varied as the Latin *Consolatio* attributed to Marcus Tullius Cicero, the Book of Mormon, and the anonymised judgements of the US Supreme Court. Experts in the field are familiar with this extensive area of much-replicated experimentation.

In Chapter 9, 'Kyd's Restored Canon', Vickers presents what he considers to be corroboration of his attribution claims provided in studies by Martin Mueller and Rizvi. Describing Mueller's work in creating a morpho-syntactically tagged set of digital transcriptions of early modern drama, Vickers reveals his ignorance of the subject. Vickers reports that using transcriptions in modern spelling is 'essential for high-speed data analysis' and that the tagging 'made it possible to extract all the repeated phrases, or *n*-grams, as they are known, extending from two words (a bigram) to seven words (a heptagram) that were repeated at least once' (312). In fact, computers are no slower at finding words in original spelling than words in modern spelling and Mueller's tagging had nothing to do with the algorithms – available since the 1950s – for finding recurrent *n*-grams. Mueller's tagging differentiates the multiple words and multiple senses of words that attach to a single string of letters such as *r-o-w* – the feat that Vickers declared impossible in Chapter 8 – and it lemmatises inflected forms (as in *rows*, *rowed*, *rowing*) so that they can be counted as occurrences of a single dictionary headword (here, *row*).

The work of Mueller that Vickers cites is a blog posting of 1450 words in August 2009 entitled 'Vickers is right about Kyd'. Vickers calls this an 'essay' and he reports the parts of it that support his Kyd attributions, while omitting to mention the parts that undermine them. Mueller used a method called Linear Discriminant Analysis (LDA) applied to his counts of the 56 three-word phrases that occur at least 500 times across his dataset of 318 early modern plays, such as the phrases *I will not* and *what do you*. The results of this count give each play a set of 56 numbers, one for

how often each of the 56 phrases occurs in that play. This string of 56 numbers is then treated as the coordinates of a point in 56-dimensional space and each point is labelled with the name of the author of the play it represents. Applying LDA is analogous to finding a new plane that slices through that 56-dimensional space such that when every point is projected onto this new plane the different authors are represented by distinct clusters of points on that plane, where each cluster is as tightly bound as possible while being as far away as possible from the other clusters.

The LDA process generates a value called 'confidence' which expresses how successfully the plane that LDA produces divides the points into the categories that we provided when we labelled each point as representing a 'play by Jonson', a 'play by Shakespeare' and so on. In his blog posting, Mueller misunderstood the concept of 'confidence' in LDA and thought it was a comment on how accurately he had labelled the points (that is, whether each play represented by each point really has the author he assigned to it). Hence Mueller mistakenly thought that the notion of 'confidence' in LDA could be used to comment on how much confidence we should place in our actions if we label *Cornelia*, *Solimon and Perseda*, *Arden of Faversham*, *King Leir*, and *Fair Em* as Kyd's. Mueller's results should have told him that something was wrong with his understanding and application of LDA since it gave only 80% and 85% 'confidence' respectively that *Cornelia* and *Solimon and Perseda* are Kyd's while having over 99% 'confidence' that *King Leir* and *Fair Em* are by Kyd. Even before applying the method to Kyd, Mueller's preliminary results should have rung alarm bells, since his method found that 'there is a 92% chance that John Lyly wrote *Love's Labor's Lost* as opposed to an 8% chance for Shakespeare'. Vickers does not mention this result. Mueller is candid about his own understanding of what he is doing, writing of LDA that 'The math is well beyond me'.

Vickers's second expert witness in defence of his new Kyd attributions is Rizvi, who created the *Collocations and N-Grams* (*CAN*) website of datasets and short research papers. Rizvi used matches of phrases across all the early modern plays to test the authorship of *The Spanish Tragedy*, *Solimon and Perseda*, *Cornelia*, *Fair Em*, *Arden of Faversham*, *Edward III*, *1 Henry VI*, and *King Leir*. For this he had to tentatively attribute them all to Kyd in his dataset because when counting the phrase matches across all early modern plays it makes a difference to each play's counts (and hence the resultant rank ordering of matching authorial canons) whether each of the other plays is assigned to a named author or to the vast canon of 'anonymous'. Rizvi tentatively attributed these eight plays to Kyd and then relied on this method to disprove the attribution if it was wrong.

Rizvi found that his method did not disprove the attribution to Kyd of the first five of these eight plays, but it did disprove the attribution to him of *Edward III*, *1 Henry VI*, and *King Leir*; these last three were now, on account of the new Kyd attributions, attributed to Marlowe. By Rizvi's method, we can have the first five of Vickers's Kyd attributions (three of which were already uncontentious) only at the price of attributing *King Leir* to Marlowe. Worse, tentatively attributing the eight plays to Kyd affects the counts for other plays so that now Marlowe's *The Jew of Malta* and *Edward II* are, by Rizvi's method, attributed to Kyd too. In other words, by Rizvi's method the only way to extend the Kyd canon as Vickers wishes to do is to extend it so absurdly far that it takes in plays that even Vickers could not countenance giving to Kyd. Vickers makes no mention of this result of Rizvi's experiments and declares himself satisfied with what he characterises as corroboration of his work.

Vickers ends the chapter with an account of Jackson's response to Rizvi's experiments and then Freebury-Jones's four-page reply to Jackson's response that was

published in *Notes & Queries*. Freebury-Jones sliced Rizvi's data in new ways and Vickers quotes his conclusion: ' … with the exception of the spreadsheet for *King Leir*, every play that Vickers ascribes solely to Kyd has *at least one other* play in the "enlarged" [Kyd] canon featuring in *the top dozen* for unique trigrams and/or tetragrams' (Freebury-Jones quoted in Vickers, 323, my emphasis). Notice how weak this claim is. Freebury-Jones's threshold for success does not require that Kyd's plays dominate the top of the rank order of other plays that have most matches with each putative Kyd play. Rather, Freebury-Jones considers his experiment to have successfully demonstrated Kyd's authorship if 'at least one other' Kyd play makes it into 'the top dozen' places in the rank order of plays with matching phrases. And, by Freebury-Jones's admission, *King Leir* still fails to pass this test for likeness to Kyd's style.

Freebury-Jones's own book necessarily duplicates much of the argument and evidence in Vickers's book, since although it also covers how Kyd influenced Shakespeare it is largely concerned with establishing the case for the enlarged Kyd canon. It is a pity that Vickers and Freebury-Jones did not coordinate their efforts, since a reader who buys only one or other of their books will not get the full story but if she buys both she will pay to read a lot of the same arguments and evidence. The most important topic unique to Freebury-Jones's book is the evidence for Kyd's hand in *1 Henry VI* and *Edward III*.

In his preface to Robert Greene's *Menaphon* (1589), Nashe refers to 'the Kidde in Aesop' who has left 'the trade of noverint' (meaning scribe) and now meddles 'with Italian translations' as Kyd had done in translating Torquato Tasso's *Padre di famiglia* as *The Household Philosophy* (1588). Greene ridiculed the author of *Fair Em*, but Freebury-Jones has to really stretch to make this remark apply to Kyd. The remark attacks the misuse of the Bible (and Kyd planned a poem about the conversion of St Paul) and where Greene seems to mock a churchman Freebury-Jones hears a mock of Kyd's father, who was a churchwarden. (The obvious questions are how this constitutes an attack on Kyd, and would Greene even have known about Kyd's father?) Greene refers to 'Saint Giles without Cripplegate' and Freebury-Jones gets from there to Kyd by noting that the historical Saint Giles was wounded by a hunter's arrow that was aimed at a young deer, that is, at a kid. Freebury-Jones finds plot parallels between *Fair Em* and 'the newly attributed *King Leir*' such as ' … the characters in both plays realise that they have been deceived' (23) and there is in both 'the illicit opening of a letter, or a box' (23). I should say that such links are found in scores of other plays too.

When Freebury-Jones turns to unusual phrasing in stage directions new matter emerges. *Arden of Faversham* has five stage directions that start '*Then they* …' and *Solimon and Perseda* has six, and Freebury-Jones claims that 'No other publicly performed play of the Elizabethan period matches' this (52). He is right, since the phrase does appear in other plays' stage directions but not so frequently: once each in *Captain Thomas Stukely*, *A Warning for Fair Women*, and *Mucedorus*, and once each in Peele's *The Battle of Alcazar* and Nicholas Udall's *Ralph Roister Doister*. Freebury-Jones observes that ' … *Arden of Faversham* and *Fair Em* share sixty-eight examples of stage directions featuring the relatively rare two-word unit "*Here enters*" (there are sixty-four instances in *Arden of Faversham* and four in *Fair Em* by my count)' (53). They do, but the anonymous play *Common Conditions* has 34 instances of '*Here enter*' and '*Here entereth*' stage directions, and *Appius and Virginia*, anonymously published in 1575 and perhaps by Richard Bower, has 16. The number of occurrences of a shared phrase does not on its own tell us what to make of them.

Freebury-Jones writes brief summaries of work done by Thomas Merriam and Albert Yang on counting single word frequencies which he says corroborate new

claims for Kyd's authorship. These summaries give no detail of the methodologies, but since Vickers has repeatedly insisted that simply counting word frequencies cannot shed light on authorship, it may be that Freebury-Jones does not want to dwell on this topic.

Freebury-Jones next turns to the topic of Shakespeare's plays echoing phrases from Kyd's plays, using a spreadsheet created by Martin Mueller called 'SHCSharedTetragramsPlus'. Freebury-Jones says this is available on his (Freebury-Jones's) website at < https://darrenfj.wordpress.com/2017111/> but at the time of writing (3 January 2025) this URL returns a '404. Page Not Found' error. This is a recurrent problem with Freebury-Jones's published work: it is impossible to follow up his references because of broken URLs. In the present case there is little excuse as the URL points to a location on Freebury-Jones's own website. It is to be hoped that reading this review encourages Freebury-Jones to reinstate the spreadsheet on his website in order to repair this deficiency. A related problem happens when, on page 61, Freebury-Jones supports his authorship claims with a reference to documents on Mueller's website called *Scalable Reading* at < https://scalablereading.northwestern.edu/?p = 312 > . At the time of writing, this website requires login credentials so the general public cannot see what is in it. But even without access to Mueller's data it is clear that Freebury-Jones's claims about it are untenable.

Freebury-Jones writes that when he had Queen Margaret say 'I will not hence' in *3 Henry VI* Shakespeare was recalling Amurath saying 'I would not hence' in *Solimon and Perseda* (63). Maybe he was, but Kyd was not the only writer to use this phrase: *EEBO-TCP* has 26 occurrences, including in William Warner's *Albion's England* (written in the late 1580s), which we know Shakespeare read. Mueller's dataset comprises only plays, but Shakespeare could as easily pick up phrases from other written sources. Other phrases that Freebury-Jones finds in Kyd and in what he considers to be the early plays of Shakespeare genuinely are rare or even unique, but Freebury-Jones seems not to notice that this undermines his claims elsewhere in the book that rare and unique shared phrasing proves common authorship not imitation or recollection.

How then are we to distinguish the effects of common authorship from instances of one writer merely echoing another? This question forms the title of one of Freebury-Jones's chapters ('Authorship versus Influence'), but he has no discernible answer for it. The only way to address it is systematically to quantify the phenomenon and look for cases where the quantity of shared phrasing far exceeds what is normal for work by different authors. What counts as normal will, however, depend at least in part on the size of the two texts that share the phrasing. We would expect to find more rare phrases in common between Charles Dickens's *Bleak House* (over 350,000 words) and Samuel Richardson's *Clarissa* (about a million words) than between any two of Dickens's short stories (typically 2000–7000 words), simply because large texts have more phrases of all kinds. Can we 'weight' the matches we find by the sizes of the texts and/or authorial canons they come from in order to control for this effect? No one has yet come up with a weighting formula that demonstrably levels this playing field, although Rizvi and Mueller have offered different formulas that they hope will do so. Freebury-Jones attempts no systematic investigation of any of this and for most of his discussion of phrase matches he eschews even counting what he finds. The closest he comes is the remark that 'The number of matches shared between two plays is divided by the combined word count of that play pair' (67) when discussing Rizvi's approach, but in fact that is not how Rizvi does it. (Rizvi's website contains an essay that describes his formula, which he revised in 2018 because he was dissatisfied with his first attempt.)

When Freebury-Jones wants to deny Shakespeare's contribution in *Arden of Faversham*, its shared phrasing with Shakespeare plays is dismissed by him as Kyd's *influence* on Shakespeare, but when Freebury-Jones wants to assert Kyd's sole-authorship of *Arden of Faversham* and other plays he claims that shared phrasing with known Kyd plays establishes common authorship. In describing his evidence, Freebury-Jones indiscriminantly mixes phrase matches that are truly rare, such as *the shadow of myself*, found only in *King Leir* and Shakespeare's *King John*, with commonplaces such as *entire affection to*, found in *Fair Em* and Shakespeare's *The Taming of the Shrew* but also in dozens of other works.

In Chapter 4, 'Revision', Freebury-Jones attempts to show that *2 Henry VI* and *3 Henry VI* were written solely by Shakespeare and that the play 'Harey the vj' recorded by Philip Henslowe as performed on 3 March 1592 was written by someone else as a prequel to those two plays and then adapted by Shakespeare (at the behest of the Chamberlain's Men) to make a three-part Henry VI cycle. These claims are uncontroversial – they are not far from the position taken in the *New Oxford Shakespeare* – and what has been fought over is who wrote the play that Shakespeare later adapted. That Nashe contributed to it is widely agreed upon, but whether Marlowe or Kyd also did is not. Freebury-Jones starts with what he considers the 'strongest' (115) qualitative evidence for Kyd's hand in the play:

> To be enrolled in the brass leaved book
> Of never wasting perpetuitie
> (*Solimon and Perseda*)

> Deserves an everlasting memory,
> To be inrol'd in Chronicles of fame,
> By never-dying perpetuity
> (*King Leir*)

> Anon from thy insulting tyranny,
> Coupled in bonds of perpetuity,
> Two Talbots winged through the lither sky
> In thy despite shall scape mortality
> (*1 Henry VI*)

It is clear what the first two quotations have in common – the idea of being part of a permanent record and the use of the words *enrolled* and *perpetuity* and the likeness of *never wasting* and *everlasting* and *never dying* – but the links between these two and the third, the one Freebury-Jones is trying to attribute, are not clear. Certainly, the word *perpetuity* is common to all three, but it is far from rare, having over 6800 occurrences in *EEBO-TCP*, and perhaps there is also the general idea of immortality.

Next, Freebury-Jones turns to Timberlake's evidence about rates of feminine endings and Oras's about pause patterns, which Freebury-Jones uses to show that what he thinks is Kyd's part of *1 Henry VI* is not anomalous amongst the other Kyd plays. Of course, if we grant this homogeneity we are saying only that this evidence does not rule out Kyd's authorship of *1 Henry VI*, not that it establishes that authorship. Freebury-Jones appears not to appreciate this distinction. We saw that Freebury-Jones considers '*Here enters*' to be a distinctive phrase in Kyd's stage directions, and in *1 Henry VI* the Bastard says in one of his speeches 'Here entered Pucelle and her practisants'. Freebury-Jones regards this as a 'similar verbal formulation' (117). But if we follow Freebury-Jones in letting lines of dialogue count as matches with stage directions

then Shakespeare's *The Comedy of Errors* has a dialogue occurrence of 'here entered' too. Freebury-Jones also counts as a match the fact that in *Solimon and Perseda* and *Arden of Faversham* a stage direction 'echoes the dialogue' and so too in *1 Henry VI*: Bedford's report that the Dauphin and Joan did 'Leap o'er the walls' echoes the state direction '*The French soldiers leap o'er the walls*' (118). But, of course, this kind of echoing is a recurrent feature in many plays and occurs in the first act of *1 Henry VI*, the part that Freebury-Jones attributes to Nashe: '*Enter on the Walls*' is followed by Joan's 'Advance our waving colours on the walls'.

Freebury-Jones considers that the 'evidential *pièce de resistance*' that seals his claim about *1 Henry VI* 'is the degree to which the verbal fabric of the non-Nashe or -Shakespeare portions is woven out of Kyd's lexical individuality' (118). This evidence is a set of calculations made by Mueller and provided in the form of a spreadsheet that Freebury-Jones says can be 'found … on my website' (141n3). This is the website that, as we noted earlier, returns a '404. Page Not Found' error for the URL that Freebury-Jones gives. It is therefore impossible to explore this evidence that Freebury-Jones describes as the strongest for Kyd's hand in *1 Henry VI*.

When he tries to refute a recent study by Gary Taylor and John V. Nance that found Marlovian word choices predominating in the non-Nashe parts of *1 Henry VI*, Freebury-Jones falls back on the biographical fact that 'Kyd and Marlowe shared a room in 1591, and it is therefore probable that they also shared a reading knowledge of each other's works' (120). Of course, consistent application of this logic would undermine Freebury-Jones's attributions to Kyd based on shared phraseology since we could say that these are cases of Marlowe using Kyd's phrasings, but Freebury-Jones admits this logic only in selected cases and when his argument needs it.

When Freebury-Jones presents the verbal evidence for his belief that Kyd wrote the non-Nashe parts of *1 Henry VI* he has to quote carefully to conceal how tenuous it is. In *King Leir* the Captain tells his men that they have 'To watch in this place, near about the beacon, | And vigilantly have regard' for passing ships. Freebury-Jones finds the same collocation of 'watch … near … vigilant' in *1 Henry VI* in the order 'vigilant … near … watch', and of course the differing order is acceptable since we are dealing here with a claimed collocation not a phrase. However, by quoting the *1 Henry VI* occurrence in two chunks, Freebury-Jones obscures the fact that the first ellipsis (from *vigilant* to *near*) covers seven words while the second (from *near* to *watch*) covers 32 words. That is, the three words *watch*, *near*, and *vigilant* collocate within a 10-word chunk of in *King Leir* but are spread across 42 words (and two speakers) in *1 Henry VI*.

If we allow the notion of collocation to cover words as widely dispersed as this, an extraordinary number of verbal matches can be found between almost any two sizeable texts and such matches prove nothing. Freebury-Jones goes on to present as evidence for common authorship the fact that *Fair Em* and the non-Nashe part of *1 Henry VI* contain the phrase 'support this', which in fact has hundreds of occurrences in books of the period, including many known to have been read by Shakespeare including William Painter's *The Palace of Pleasure*, George Gascoigne's poetry, and Raphael Holished's and John Stowe's Chronicles. Likewise, the phrase 'to none but to' is found in *Fair Em* and the non-Nashe part of *1 Henry VI*, but also found over 1200 times in books of the period, including works by Nashe, Marlowe, Edmund Spenser, and Thomas Lodge. These examples could be multiplied many times – including the phrases 'in the hour of death' and 'dare not speak' – as cases of shared phrasing that Freebury-Jones thinks unusual enough to indicate shared authorship. But they are either just commonplace expressions found in hundreds of books of the period, or else are genuinely rare phrases but are found in other Shakespeare plays, such as

'make this marriage' in *The Spanish Tragedy* and the non-Nashe part of *1 Henry VI* but also (unmentioned by Freebury-Jones) in *Antony and Cleopatra*.

Freebury-Jones next considers Shakespeare's Additions to Kyd's *The Spanish Tragedy* which reached print in 1602 and about which there is now little controversy. Freebury-Jones writes that 'The *New Oxford Shakespeare* team accepts that Shakespeare had a hand in the additions' (140), which suggests the team's reluctance to agree to this claim. In fact, Hugh Craig – advisor to the *New Oxford Shakespeare* – was the first to make this claim and has published repeatedly on it from his independent and mutually buttressing studies.

In his Chapter 5, 'Collaboration', Freebury-Jones sets out the known facts about *Edward III* and claims that Shakespeare's co-author on it was Kyd. The first empirical evidence for Kyd's hand in *Edward III* that Freebury-Jones adduces is that its proportions of various rhyme patterns are like those of *The Spanish Tragedy*, *Solimon and Perseda*, *Cornelia*, *King Leir*, *Arden of Faversham*, *Fair Em*, and parts of *1 Henry VI* (150). Freebury-Jones does not give the corresponding data for these rhyme patterns' proportions in other plays, so it is impossible to judge how alike are the data for the alleged Kyd plays. (And the numbers for the Kyd plays do not seem especially compelling on their own, being a seemingly random scattering from 0 to 30 occurrences across the plays.) Regarding the feminine-ending data, Freebury-Jones builds upon his previous attributions to make this new one, so that ' … we might note that the range for long scenes in these portions of *Edward III* is strikingly close to that obtained for Kyd's scenes in *Henry VI Part One*' (151). Tying new attributions together like this is of course perilous, since it encourages the reader to either accept both new attributions or, just as easily, reject both. Freebury-Jones does the same with the pause-pattern data, relying on his attribution of part of *1 Henry VI* to Kyd to support his attribution of part of *Edward III* to Kyd.

Then Freebury-Jones moves onto the phrasal-matching data from Rizvi, and all the previous objections apply here too. Freebury-Jones is right that ' … the bigram "no issue" appears in Kyd's (by my argument) unhistorical dramatisation of Mortimer's death in *Henry VI Part One*' and in *Edward III* and in *King Leir* (155). But he is wrong to find this significant since there are over 2000 occurrences of *no issue* in other books of the period. The phrase is just another commonplace.

It would be pointless to continue drawing attention to these logical slips in Freebury-Jones's work: they apply everywhere in the book. To be clear, I am not claiming that Freebury-Jones has found no rare verbal parallels between the known works of Kyd and the plays he wants to add to the Kyd canon. He has found some. The problem is that he has diluted this strong evidence with much weak evidence and nowhere undertakes a systematic quantification of what he has discovered.

Having expanded the Kyd canon in his previous chapters, in Chapter 6, 'Kyd's Influence on Shakespeare's Later Plays', Freebury-Jones considers echoes of this expanded canon in Shakespeare's late works. That Shakespeare echoed phrases from existing plays is not news, and the claim here that Kyd wrote some of those plays makes no difference to an analysis of them. That is, Freebury-Jones merely points out the echoes and offers no argument that seeing them all as being by Kyd should alter how we think of them. At best, Freebury-Jones can make only a tautology: 'Acknowledgement of Kyd's "enlarged" canon thus enables us to achieve a better understanding of his enduring influence on Shakespeare, and to recognise a slightly different relationship between their dramas' (170).

As with Vickers's book, Freebury-Jones rightly finds merit in thinking through parallels of phrasing between plays by different authors, and he has uncovered many

interesting ones. But they do not add up to authorship-attribution evidence unless they are handled systematically, and the rare ones are distinguished from the commonplace. All such evidence must be quantified, and it must be shown that the shared phrases that are used to claim shared authorship are genuinely rare and that we should not expect mere chance to put them where we find them and so often. Showing this requires first calculating how often two works that we are sure are by different authors nonetheless share rare phrases, so that we establish a 'baseline' for what is demonstrably attributable to mere chance and against which we can judge when parallels of phrasing are too frequent to be plausibly dismissed as coincidence. A competent investigator using any new authorship attribution method would first apply it to a large number of so-called validation runs. In these runs, works for which we know the author are treated as if we did not know this and the method is allowed to give its various verdicts. Then she would compare the method's answers to our knowledge of the true authorship of these works in order to produce a quantified index of how reliable her method is. Anything short of this scientific rigour should convince no one.

Gabriel Egan
De Montfort University
✉ gegan@dmu.ac.uk

Check for updates