Shakespeare's Vocabulary: Myth and Reality[1]

Hugh Craig

I

Sometime in the second decade of the nineteenth century, Samuel Taylor Coleridge noted in his copy of Shakespeare's *Dramatic Works* that the playwright "can be complimented only by comparison with himself: all other eulogies are either heterogeneous, (ex. gr. in relation to Milton, Spencer, &c) or flat truisms (ex. gr. to prefer him to Racine, Corneille, or even his own immediate Successors, Fletcher, Massinger &c.)."[2] For Coleridge, to compare Shakespeare to anyone else made little sense. Shakespeare scholarship has continued in full spate since Coleridge's time, and a great deal has changed, but studies which put Shakespeare in the context of his peers as a regular member of a collective playwriting enterprise that together created English Renaissance drama are still the exception rather than the rule. It has been more common to regard him, as Coleridge does, as *sui generis*. Gary Taylor remarked in 1989 that Shakespeare's fame had made it hard to see his works and the works of others clearly. In gravitational terms, "cultural space-time" is bent by the black-hole-like singularity of his reputation.[3] Taylor cites publications by leading Shakespeare scholars of the day that reflect this sense of Shakespeare's exceptionality: Kenneth Muir's *Singularity of Shakespeare* and Harry Levin's "Primacy of Shakespeare."[4] More recently, while scholars have viewed the drama of Shakespeare's time much more as a collective enterprise and have questioned the importance of individual authorship in general, the effect has been to downplay all authorial difference and thus to

[1] In their "Shakespeare's Vocabulary: Did It Dwarf All Others?" in *Language and Style in Shakespeare—New Insights*, ed. Mireille Ravassat and Jonathan Culpeper (London: Continuum, forthcoming), Ward E. Y. Elliott and Robert J. Valenza independently come to the same conclusion as I do here, using different texts and methodology. Thanks to Ward Elliott for letting me know of this essay and offering material support for my own parallel endeavors; thanks also to Arthur F. Kinney for help with this essay in its early stages.

[2] Marginal note in *The Dramatic Works of William Shakespeare*, ed. Samuel Ayscough (London, 1807), quoted in James Engell, "Coleridge, Johnson, and Shakespeare: A Critical Drama in Five Acts," *Romanticism* 4 (1998): 22–39, esp. 24.

[3] Gary Taylor, *Reinventing Shakespeare: A Cultural History, From the Restoration to the Present* (New York: Weidenfeld and Nicholson, 1989), 410–11.

[4] Taylor, 374, 377.

pay less attention to the characteristics of one playwright's output against another's.[5] Perhaps the strongest indicator of a new interest in Shakespeare in relation to his peers is a new wave of strictly attributive studies in the past decade.[6]

One aspect of the idea that Shakespeare is somehow beyond comparison is the belief that he has a preternaturally large vocabulary. This notion is still current, having survived both the rethinking of playwriting as a communal activity and modern attribution studies. The best-selling history of the English language, *The Story of English*, estimates that Shakespeare's vocabulary was twice as large as that of an educated person today.[7] In Stephen Greenblatt's introduction to the *Norton Shakespeare*, we read of "Shakespeare's astonishing vocabulary of some 25,000 words." Greenblatt calls this an "immense word hoard" and compares it to the 12,000 of John Milton, "his closest rival among the great English poets of the period."[8]

For the most recent systematic study of this question, we must go back to the work of Alfred Hart, which began appearing in the 1930s. Hart used newly published concordances to establish the number of different words used in each play, how many words were peculiar to each play, and how these measures fluctuated over Shakespeare's career. He divided the plays into four successive chronological groups and found that over time the average vocabulary for each group "shows a small but definite increase in size," on top of the fact that "Shakespeare commenced dramatist with a very large stock of words."[9] He offered some comparisons to plays by Marlowe, Kyd, and Greene and found that Shakespeare exceeds them in vocabulary. One of Hart's measures is the number of what he called "fresh" words—that is, words a playwright uses for the first time in a given play. (These are distinct from "coinages" or "neologisms"—the words a writer

[5] A high-water mark of this trend was *A New History of Early English Drama*, ed. John D. Cox and David Scott Kastan (New York: Columbia UP, 1997), which had no sections devoted to individual authors. In Shakespeare studies, the movement did not go unchallenged, of course; for example, see Richard Levin, "The Poetics and Politics of Bardicide," *PMLA* 105 (1990): 491–504; and Brian Vickers, *Appropriating Shakespeare: Contemporary Critical Quarrels* (New Haven: Yale UP, 1993), 92–162.

[6] See Brian Vickers, *"Counterfeiting" Shakespeare: Evidence, Authorship, and John Ford's "Funerall Elegye"* (Cambridge: Cambridge UP, 2002), and *Shakespeare, Co-Author: A Historical Study of Five Collaborative Plays* (Oxford: Oxford UP, 2002); and MacDonald P. Jackson, *Defining Shakespeare: "Pericles" as Test Case* (Oxford: Oxford UP, 2003).

[7] Robert McCrum, Robert McNeil, and William Cran, *The Story of English*, 3rd ed. (New York: Penguin, 2002), 102.

[8] Greenblatt does not say how he arrived at this estimate, which is higher than those generally quoted. See the discussion of methods for word counting below. See Stephen Greenblatt, "General Introduction," in *The Norton Shakespeare, Based on the Oxford Edition*, 2nd ed. (Norton: New York, 2008), 65.

[9] Alfred Hart, "Vocabularies of Shakespeare's Plays," *Review of English Studies* 19 (1943): 128–40, esp. 133 and 136.

introduces to the language.[10]) Shakespeare's vocabulary, Hart says, is "large and comprehensive," while the number of different words used in Marlowe's plays declines over time "because he did not insert enough fresh words in each of his plays," Kyd has "no super-abundance of words," and Greene has "a meagre stock."[11] In a second article, Hart concedes that "great length and a large vocabulary do not necessarily carry with them high dramatic quality,"[12] but he does often connect the two:

> *Hamlet* is the supreme example of Shakespeare's delight in and command of *fresh* and forceful words. By this time he had written twenty-two plays and all his poems, and could draw upon a vocabulary of 13,765 words, yet to this enormous stock he added another 606 words, all previously unused. How deep and apparently inexhaustible were the wells of his memory and invention, and how marvellous his aptitude for word-coining, will be evident from the addition of 302 peculiar words in writing *Troilus and Cressida*.[13]

Hart wished to prove what might seem natural enough in a playwright regarded as unmatched in English literary history. It stood to reason that such an author would have the most extensive vocabulary. This view fitted a predisposition to think of Shakespeare as a prodigy, and Hart's carefully calculated play-by-play comparisons commanded a respect which was transferred to his estimations of Shakespeare's overall vocabulary as compared to that of other playwrights.[14]

Hart was writing in the 1930s and 1940s, and he did his counting by hand. By the 1970s, a new generation of researchers was approaching quantitative

[10] On Shakespeare's "coinages," see Jürgen Schäfer, *Documentation in the O.E.D.: Shakespeare and Nashe as Test Cases* (Oxford: Clarendon Press, 1980); Bryan A. Garner, "Shakespeare's Latinate Neologisms," *Shakespeare Studies* 15 (1982): 149–70; John Willinsky, *Empire of Words: The Reign of the OED* (Princeton: Princeton UP, 1994); Jeffrey McQuain and Stanley Malless, *Coined by Shakespeare: Words and Meanings First Penned by the Bard* (Springfield, MA: Merriam-Webster, 1998); and MacDonald P. Jackson, "Neologisms and Non-Shakespearean Words in *A Lover's Complaint*," *Archiv für das Studium der Neueren Sprachen und Literaturen* 245 (2008): 288–302.

[11] Hart, "Vocabularies of Shakespeare's Plays," 139.

[12] Alfred Hart, "The Growth of Shakespeare's Vocabulary," *Review of English Studies* 19 (1943): 242–54, esp. 242.

[13] Hart, "Growth of Shakespeare's Vocabulary," 254.

[14] MacDonald P. Jackson, "Shakespeare's *A Lover's Complaint*: Its Date and Authenticity," *University of Auckland Bulletin* 72 (1965): 7–39, esp. 13; Mark Dominik, *William Shakespeare and the Birth of Merlin*, rev. ed. (Beaverton, OR: Alioth Press, 1991), 135; Eric Sams, *Shakespeare's Lost Play: "Edmond Ironside"* (London: Fourth Estate, 1985), 346–47; and Vickers, *Shakespeare, Co-Author*, 80, all cite Hart in discussing Shakespeare's exceptionally large vocabulary. For examples of scholars taking up Hart's idea of Shakespeare's fertility in introducing new words in each successive play, see Inga-Stina Ewbank, "*Hamlet* and the Power of Words," *Shakespeare Survey* 30 (1977): 85–102, esp. 89; G. R. Hibbard, ed., *Hamlet* (Oxford: Oxford UP, 1987), 30; and Colin Burrow, ed., *The Complete Sonnets and Poems* by William Shakespeare (Oxford: Oxford UP, 2002), 139.

questions about Shakespeare's vocabulary with machine-readable texts and computerized counts. The main focus was still on internal Shakespeare comparisons, however, and few texts beyond his were prepared, so that large-scale cross-author comparisons were not possible. Shakespeare data, mostly based on Marvin Spevack's concordance,[15] were often used in discussions of ways of finding a measure of vocabulary richness—that is, of the relationship of text length to the number of different words used—that was independent of the length of the texts.[16] Papers estimated the number of words Shakespeare knew but did not use[17] and how Shakespeare's vocabulary compared to the total available English lexicon of his day.[18] Those who did examine the number of different words used in plays, poems, and prose works by other authors in relation to Shakespeare tended to find contradictions of the idea that Shakespeare was exceptional in this regard. Louis Ule was surprised to discover that Nashe's play *Summer's Last Will and Testament* used proportionately more words than any Shakespeare play, for instance. He concluded that vocabulary richness should not be regarded as a measure of literary quality.[19] In a second article, Ule compared Shakespeare's vocabulary to Marlowe's and found that the latter's was larger. He attributed

[15] Marvin Spevack, *A Complete and Systematic Concordance to the Works of Shakespeare*, 9 vols. (Hildesheim: Georg Olms, 1968–80).

16 David A. Ratkowsky, M. H. Halstead, and Linda Hantrais, "Measuring Vocabulary Richness in Literary Works: A New Proposal and a Reassessment of Some Earlier Measures," *Glottometrika* 2 (1980): 125–47; Daniel Dugast, "Le vocabulaire du théâtre de Shakespeare: Sur l'appréciation de la richesse lexicale et le rôle de la fréquence moyenne," *Cahiers de Lexicologie* 55 (1989): 135–65; and Gilbert Youmans, "Measuring Lexical Style and Competence: The Type-Token Vocabulary Curve," *Style* 24 (1990): 584–99, and "A New Tool for Discourse Analysis: The Vocabulary-Management Profile," *Language* 67 (1991): 763–89.

[17] Bradley Efron and Ronald Thisted, "Estimating the Number of Unseen Species: How Many Words Did Shakespeare Know?" in *Biometrika* 63 (1976): 435–47. Efron and Thisted wrote a follow-up paper applying their method to the "Shall I die" poem; see Ronald Thisted and Bradley Efron, "Did Shakespeare Write a Newly-Discovered Poem?" in *Biometrika* 74 (1987): 445–55. The problem of estimating Shakespeare's "passive" vocabulary is also studied in Shahar Boneh, Arnon Boneh, and Richard J. Caron, "Estimating the Prediction Function and the Number of Unseen Species in Sampling with Replacement," *Journal of the American Statistical Association* 93 (1998): 372–79.

[18] Many of these studies arose from a project to create a computerized Shakespeare dictionary; see Thomas H. Finkenstaedt, H. Joachim Neuhaus, and Marvin Spevack, "SHAD: A Shakespeare Dictionary," in *Computers in the Humanities*, ed. J. L. Mitchell (Edinburgh: Edinburgh UP, 1974), 111–23; H. Joachim Neuhaus and Marvin Spevack, "A Shakespeare Dictionary (SHAD): Some Preliminaries for a Semantic Description," *Computers and the Humanities* 9 (1975): 263–70; H. Joachim Neuhaus, "Author Vocabularies Compared with Chronological Dictionaries," *ALLC Bulletin* 6 (1978): 15–20; and H. Joachim Neuhaus, "Lexical Database Design: The Shakespeare Dictionary Model," in *Coling '86: Proceedings of the 11th Conference on Computational Linguistics* (Bonn: Institut für angewandte Kommunikations- und Sprachforschung, 1986), 441–44.

[19] Louis Ule, "Vocabulary Richness in Shakespeare," *ALLC Bulletin* 5 (1977): 174–77.

this unexpected finding to the particular nature of the Marlowe works that have survived.[20] Eliot Slater relied on Hart's counts of Shakespeare words but questioned Hart's conclusions about Shakespeare's "rich vocabulary," finding that the figures for Marlowe were in fact quite similar.[21] Slater's main interest was using statistics for rare words to determine the chronology of Shakespeare's plays and the authorship of *The Raigne of King Edward III*, and he did not pursue the matter of the size of Shakespeare's vocabulary in relation to that of his peers further. Neither did any of the other early computing humanists. They were not able to, after all, because only Shakespeare texts were available to them in any quantity.

To these skeptical voices should be added that of Otto Jespersen, who included a chapter on "Shakespeare and the Language of Poetry" in his classic history of the English language, first published in 1905. He distinguished between the words a writer knows and those he or she actually uses, and between the words used in conversation and in writings, and he corrected extreme views such as Max Müller's about the differences between the vocabularies of educated and uneducated speakers. Jespersen says Shakespeare used around 20,000 words in his works, compared to 8,000 for Milton in his poetry, but considers this to be "a natural consequence of the narrower range of [Milton's] subjects."[22] He associates Shakespeare's large vocabulary not with a hypertrophied language capacity but with the number of topics he covered: "The greatness of Shakespeare's mind is therefore not shown by the fact that he was acquainted with 20,000 words, but by the fact that he wrote about so great a variety of subjects and touched upon so many human facts and relations that he needed this number of words in his writings."[23]

David Crystal recently mounted a specific challenge to what he calls "the quantity myth" of Shakespeare's vocabulary, the belief that Shakespeare had "'the largest vocabulary of any English writer.'"[24] This is not true, Crystal says, simply because the vocabulary of English has increased over the centuries, so that modern writers in general use far more words than Shakespeare. This reminds us that "it is not so much the number of words we have as what we *do* with those words that makes the difference between an ordinary and a brilliant use of language."[25] However, he says, Shakespeare's vocabulary of 20,000 words

[20] Louis Ule, "The Weird Ways of Vocabulary," *ALLC Journal* 6 (1985): 24–28.

[21] Eliot Slater, *The Problem of "The Reign of King Edward III": A Statistical Approach* (Cambridge: Cambridge UP, 1988), 78; see also 64.

[22] Otto Jespersen, *Growth and Structure of the English Language*, 10th ed. (Chicago: U of Chicago P, 1982), 202.

[23] Jespersen, 202.

[24] David Crystal, *"Think on My Words": Exploring Shakespeare's Language* (Cambridge: Cambridge UP, 2008), 2.

[25] Crystal, 3.

"*was* a large vocabulary, in its day." He compares the 6,000 different words used in the 880,000 words of the King James Bible to Shakespeare's 20,000 used in the 884,647 words of the *Riverside* edition. He continues, "Why is Shakespeare's vocabulary so large? Partly because he wrote so much, but mainly because of what he wrote about. It is the difference between people, situations, and subject-matter which generates different kinds of vocabulary, and Shakespeare is acknowledged to be unmatched in the range of his characters, settings, and themes."[26] Crystal lists some of the areas in which Shakespeare wrote, each of them, he thinks, calculated to bring fresh ranges of words to his tally. He concludes, "If you do all of these, and more, inevitably you will end up with a lexical total that makes you stand out from your contemporaries."[27]

In this essay, I present evidence that Shakespeare introduces "fresh" words—that is, words he has not used before—at about the same rate as his contemporaries. He does use more different words overall than they do, not because of any unusual internal variety but only, it would seem, because his canon is larger than theirs. This is one aspect of Shakespeare's exceptionality that can be studied and resolved one way or another, since it takes a numerical form. The statistics of word use should be impervious to the forces impeding comparison. They provide an objective foundation for a consideration of style, if a drastically limited one. Seeing Shakespeare's plays as part of a corpus of dramatic work from the period—comparing him—helps dispel the myth of his exceptional vocabulary.

II

There are different ways of arriving at an overall figure for "Shakespeare's vocabulary," in the sense of the number of different words he used in his surviving works. Poems and collaborative works may or may not be included. Different texts of individual works will make for small variations. Words may be defined as lemmas— as dictionary-type headwords—or as word forms. Working from the combined plays and poems data in Spevack's concordance, Joachim Neuhaus gives round figures of 900,000 words in total, with 30,000 different word forms, which reduce to 20,000 lemmas. If proper names and foreign words are excluded, he says, there are 17,500 lemmas.[28]

In my analysis I count word forms, treating each form of a lexical word as a separate item, so that "light" is a single form, whether used as adjective, noun or verb, while *run*, *ran*, *week*, and *weeks* are all distinct words.[29] (It may be conve-

[26] Crystal, 6.
[27] Crystal, 7.
[28] Neuhaus, "Author Vocabularies Compared with Chronological Dictionaries," 15.
[29] The results that follow are based on counts taken from old-spelling versions of early printed texts of the plays. Variant spellings were consolidated into lists by a computer program. This

nient to think of these as "concordance words," following Slater.)[30] This does not allow precise replication of Hart's work, which involved combining word forms into dictionary-type headwords. But we can make our own separate estimate of the comparative vocabularies of Shakespeare and his fellow dramatists, treating all of them in the same way.

Hart did his counting by hand and was constrained by the sheer labor involved from compiling statistics much beyond those for Shakespeare. While there are still relatively few printed concordances for Shakespeare's contemporaries,[31] it is now possible to assemble sizable collections of electronic texts of plays, and this opens the way to fast and reliable counting of words across the dramatic production of the period. Here, I reconsider Hart's conclusions about the size of Shakespeare's vocabulary relative to his peers by reference to a corpus consisting of a hundred or so plays by others and a core set of twenty-eight plays which are generally accepted as Shakespeare's unaided work (details of the plays are in the Appendix).

The results confirm that Shakespeare has a larger vocabulary than his peers. If we compare Shakespeare's sole-authored plays to five other complete dramatic canons from the period—those of Jonson, Lyly, Marlowe, Middleton, and Peele—we find that Shakespeare used over 20,000 different words, as I define them above, while the range is from Peele at just under 6,000 to Jonson at 18,500. Yet we must bear in mind that more Shakespeare plays survive than those of any other playwright of the era. There is good reason to think that Shakespeare was involved in the writing of 44 plays.[32] He may occasionally

---

means that some word forms in the texts were discarded and left out of the counts because they belong to two or more lists and thus are ambiguous. To arrive at the consolidated counts, proper nouns, such as characters' names, were counted as distinct words. Words in foreign languages were excluded. Plurals and conjugated verbs were treated as distinct words. All the texts were treated alike, so there is no reason to think these factors over such a large corpus would create any bias. For a description of the software used to deal with spelling variants in these texts, see Hugh Craig and R. Whipp, "Old Spellings, New Methods: Automated Procedures for Indeterminate Linguistic Data," *Literary and Linguistic Computing* 25 (2010): 37–52.

[30] Slater, 20, 29.

[31] There are printed concordances for the dramatic works of Kyd, Dekker, Marlowe, and Webster and another planned for Ford, but none available as far as I know for Chapman, Fletcher, Greene, Heywood, Jonson, Lyly, Middleton, or Peele. See V. A. Small, R. P. Corballis, and J. M. Harding, *A Concordance to the Dramatic Works of Thomas Dekker* (Salzburg: Institute für Anglistik und Amerikanistik, 1984); Ch[arles] Crawford, *A Concordance to the Works of Thomas Kyd* (Louvain: Uytpruyst, 1906–10), and *The Marlowe Concordance*, 2 vols. (Louvain: Uystpruyst, 1911–32); Richard Corballis and J. M. Harding, *A Concordance to the Works of John Webster*, 4 vols. (Salzburg: Institut für Englishce Sprache und Literatur, 1978); and "The Complete Works of John Ford," Institute of English Studies, University of London; http://ies.sas.ac.uk/cmps/Projects/Ford/index.htm (accessed 10 January 2011).

[32] This figure takes account of Shakespeare's part in *Edward III* and *Arden of Faversham*, his (conjectured) additions to *The Spanish Tragedy* and *Sir Thomas More*, and the lost plays *Love's*

have been exceeded in sheer output by his peers—Thomas Heywood claimed in the preface to *The English Traveller* to have had "either an entire hand, or at the least the maine finger" in 220 plays, although at the most generous estimate we have records of only 42 of these, and surviving copies of only 25—but Shakespeare's is the largest surviving canon.[33] The next largest is Middleton's, whom the recent Oxford edition associates with 31 plays, 18 as sole author and 13 as collaborator,[34] and next is Jonson with 17 sole-author plays. Probably two factors are at work in this metric: Shakespeare was indeed exceptionally productive, and an unusual proportion of his dramatic work survives because his plays were frequently printed in quarto editions before 1600 and his later plays were collected and printed after his death in the First Folio.

The obvious explanation for Shakespeare's relatively large vocabulary is that he had more opportunity to use different words. Each extra play added more new words to his total. Figure 1 shows how vocabularies relate to the totals of words used in a canon for eleven playwrights each represented by five or more plays. Shakespeare uses about the same number of words in his plays that one would expect, given the size of his canon. The dashed line in Figure 1 is a "line of best fit," showing the overall trend in this set of relationships between total words and different words used. It represents the number of different words we would expect given a canon of a certain size. Jonson's count, the black square, is above the line, indicating that he uses more different words than the norm for this group; Shakespeare's, the black diamond, is below, indicating fewer different words than we would expect. Taken as a whole, therefore, Shakespeare's dialogue is a little denser than that of his peers, in the sense that he uses proportionately fewer different words within a given scene or play. Thus, the data in Figure 1 suggest that Marlowe, Middleton, Jonson, or any of the others would have reached or exceeded Shakespeare's total dramatic vocabulary if they had written as much dialogue as he did, assuming that they used as many or more different words as he did for a given amount of dramatic writing.

However, this simple calculation—the ratio of different words to total words—is not entirely satisfactory as a measure of "vocabulary richness" because larger canons would be expected to be relatively more dense. On the whole, one would expect each new play to add more vocabulary, but at a decreas-

---

Labor's Won and *Cardenio*, as well as *Pericles*, *Two Noble Kinsmen*, and the thirty-six plays in the First Folio.

[33] Thomas Heywood, "To the Reader," in *The English Traveller* (London, 1633), sig. A3r. Twenty-five is the tally of plays associated with Heywood in Alfred Harbage and S. Schoenbaum, *Annals of English Drama 975–1700*, 2nd ed. (Philadelphia: U of Philadelphia P, 1964), leaving aside pageants, "classical legends," and the like.

[34] Gary Taylor and John Lavagnino, eds., *Thomas Middleton and Early Modern Textual Culture* (Oxford: Clarendon Press, 2007).

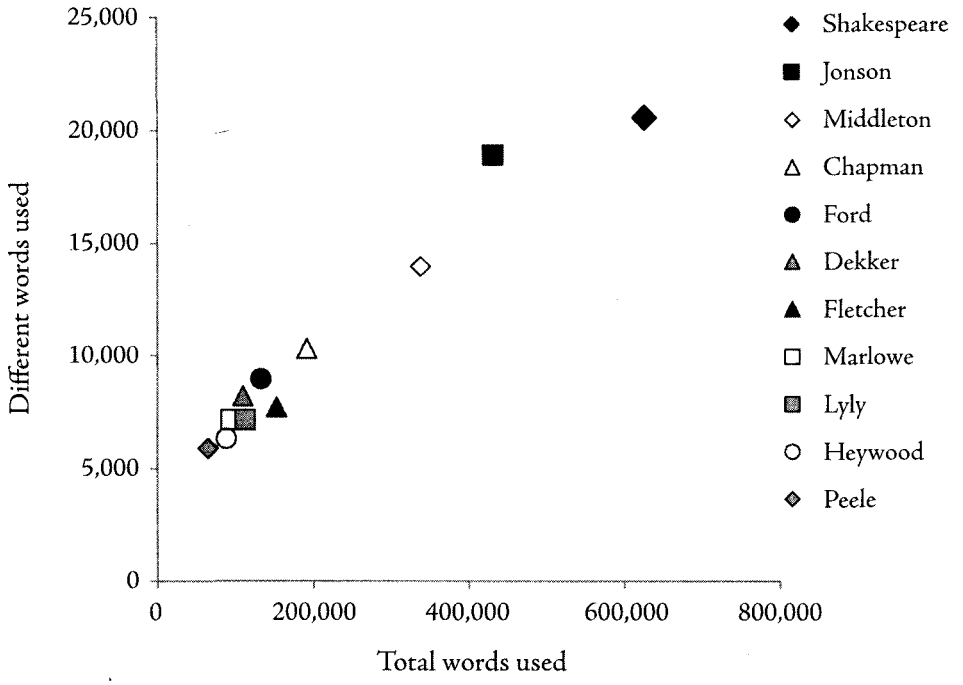

Figure 1: Vocabularies of eleven playwrights.

ing rate. This makes sense if we consider what happens with a very small sample. A single sentence might have all different words. It is much less likely that a second sentence would not repeat any words, and so on. In the same way, we can predict that as successive plays are added to a canon, new words will be added and overall vocabulary size will grow, but at a slower rate, so that the overall ratio of different words to total words will decline.[35]

Shakespeare's large canon and the interactions between sample size and vocabulary size thus make exact comparisons difficult. One way to overcome this and put Shakespeare on the same footing as his peers is to standardize the word length of our samples by using just the first 10,000 words of plays. We can take a set of 121 plays dated from 1580 to 1619, all of them single-author, well-attributed, original compositions, rather than translations, and all written for the stage rather than for the closet. Shakespeare was sixteen at the beginning of this period and had been dead three years at the end of it, so these four decades seem to be a reasonable span to represent the work of the immediately preceding generation, his direct contemporaries, and those who followed immediately afterward. Twenty-four playwrights, in addition to Shakespeare, are represented.

[35] For discussions, see Dugast; Ratkowsky et al.; Ule, "Vocabulary Richness in Shakespeare"; and Youmans, "Measuring Lexical Style and Competence."

Table 1 shows authorial averages for the numbers of different words used in these segments for playwrights in the set with three or more plays included there. The problem of the unequal size of samples is removed, because each segment in the collection is the same size; we are considering an average, so playwrights with large or small canons are at neither an advantage nor a disadvantage. We might feel that for those authors with a large set, such as Shakespeare, the average is a truer representation of their underlying habits because anomalies will balance themselves out, but there is no reason to expect this average to be higher or lower than those of playwrights with small canons. Shakespeare ranks seventh out of thirteen, with just as many others above him as below him. He used a large number of different words in the course of writing a large number of plays, but a more or less average number of words for a given, standard-sized chunk of dialogue. If we turn to columns three and four, we find that Shakespeare's highest-scoring segment is fourth in the group, and his lowest-scoring is eighth. With twenty-eight segments, he has plenty of opportunity to display an unusual pattern, but he remains comfortably within the envelope of his peers on these measures.

Another way to put Shakespeare on the same footing as his contemporaries is to ask how many new words he and other dramatists with large surviving oeuvres used in successive plays. This is an approach that interested Hart, as we have seen: he noted Shakespeare's unusual capacity to introduce "fresh" words

Table 1: Number of different words used in the first 10,000 words of plays[1]

Author	Segments	No. of different words		
		Lowest	Highest	Average
Webster	3	1,694	1,957	1,827.0
Dekker	5	1,502	2,011	1,772.0
Peele	4	1,649	1,845	1,749.5
Marlowe	5	1,554	1,924	1,740.8
Jonson	12	1,494	1,964	1,727.1
Greene	4	1,512	1,829	1,675.5
Shakespeare	28	1,426	1,930	1,663.5
Lyly	8	1,497	1,688	1,604.5
Chapman	10	1,281	1,783	1,581.3
Heywood	5	1,389	1,769	1,554.6
Middleton	12	1,308	1,655	1,550.8
Fletcher	7	1,394	1,682	1,525.5
Wilson	3	1,333	1,710	1,499.0

[1] Analysis is by author, arranged by average count, largest value first.

with each new play, that is, words he had not used before. Thomas Middleton with eighteen sole-authored plays and Ben Jonson with seventeen are the two writers whose work is closest in number in our corpus to Shakespeare with twenty-eight. Figure 2 shows the raw counts for the new words introduced in their second and following plays, and Figure 3 displays the counts of new words expressed as words per 1,000. Successive Shakespeare plays do not bring with them any extraordinary freight of freshly introduced words. The aberrations in Figure 3, where counts are standardized for size, come more from Jonson and Middleton, especially in their late plays. The general pattern is one of decline, rapid at first and then gentler, but the last Jonson play and the last Middleton play—*The Sad Shepherd* and *A Game at Chess*—reverse the trend by bringing in more "fresh" words proportionately than their predecessors. Each of these plays is a departure for its author, attracting a new range of vocabulary. Shakespeare's line, although it extends far to the right of the other two, follows what seems to be the common underlying profile, a progressively flattening curve.

For the secrets of Shakespeare's undoubted greatness, it seems we must look elsewhere than in a prodigiously rich vocabulary in the particular terms we have been examining—that is, the number of different words he uses and the number of new words in a given work. Jespersen and Crystal were right to be skeptical about the myth about Shakespeare's vocabulary, but they did not make comparisons with Shakespeare's peers and so attributed Shakespeare's large vocabulary to an exceptional range and variety of situations in his drama. The truth is much simpler: Shakespeare has a larger vocabulary because he has a larger canon.

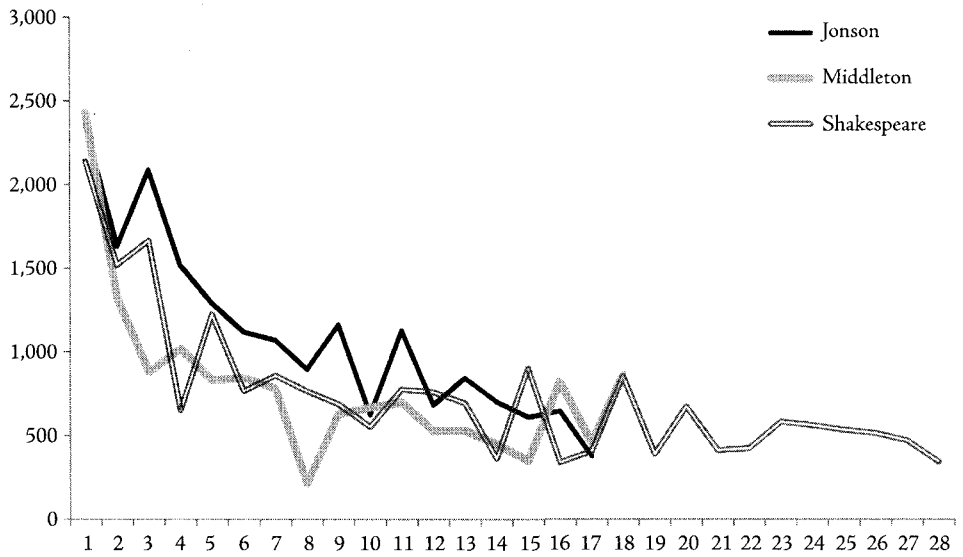

Figure 2: New words in successive plays.

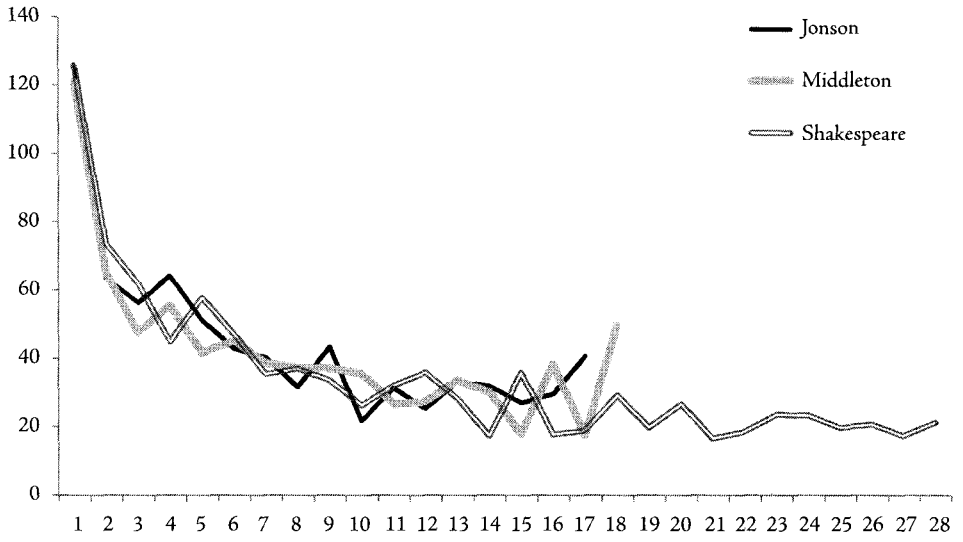

Figure 3: New words in successive plays, per 1,000 words.

When we look closely at this question using the new methods, Shakespeare merges back into his setting, as a working playwright with a pattern closely resembling many others of his time. In this case, the corpus approach allows us to relinquish a mistaken way of understanding what makes Shakespeare exceptional, and frees us to turn to better-founded ones.

Slater's doubts that Shakespeare's plays were necessarily richer in vocabulary than those of his peers have already been mentioned. Taking this a step further, Slater argues that the critical focus should move from exotic vocabulary items and toward the more workaday range. To illustrate, he quotes the third quatrain of Sonnet 13:

> Who lets so fair a house fall to decay,
> Which husbandry in honour might uphold
> Against the stormy gusts of winter's day,
> And barren rage of death's eternal cold?[36]

He points out that although this is "one of [Shakespeare's] more flamboyant images," the words themselves are almost all common ones. Of the fourteen words used in lines 11 and 12, only two, "stormy" and "gusts," are at all unusual, appearing by his count seven and eight times in the canon, respectively. The next most uncommon is "winter's," which occurs twenty-two times in the works as a whole. Slater rightly concludes that the effect of the lines derives not from strik-

[36] Greenblatt, gen. ed., *Norton Shakespeare*, 1950.

ing vocabulary items but "from the content, the imagery, and particularly from the contrast between concrete and abstract, and a certain ambiguity of meaning." Slater suggests that we should look to the opposite tendency as characteristic: Shakespeare typically made "abundant use of the most commonplace words to produce far from commonplace effects," to the extent that "he falls at times into a repetitiveness that makes one grit one's teeth."[37]

We can test this idea with the 121 equal-sized play blocks. Do Shakespeare plays use more of the commoner words than plays by others, as Slater implies? At the other end of the scale, do they use fewer of the rare words than those of his contemporaries? Figure 4 plots each of the segments according to the number of words accounted for by the 500 commonest words overall versus the number of "singletons," words used only once anywhere in these segments. These two measures for this set of segments are obviously related. As the total of words derived from the top 500 increases and we move from left to right across the chart, the number of once-used words in a given segment decreases. To the lower right are segments for which the top 500 words form as much as three-quarters of their total of 10,000 words and which have much fewer than their share of the "singletons," as few as 50. It is easy to understand why this should be so. Plays with varied settings and long speeches of description and persuasion tend to use both a more varied vocabulary and more of the very rare words. Therefore, Crystal is probably right to say that treating a mixture of "people, situations, and subject-matter" brings about a more extensive vocabulary, but Shakespeare does not do any more of this than his contemporaries.[38] Plays with static and domestic settings often have shorter speeches with more dialogic interchange and so have a denser pattern of vocabulary, using commoner words more often and using fewer of the one-off words. John Marston's tragedy *Antonio's Revenge* is the lowest in its use of the top 500 words and fourth highest in the counts of singletons. It is an extravagant revenge tragedy set in Italy, with multiple murders and a climactic masque. John Fletcher's tragicomedy *The Loyal Subject* is second highest in its use of the top 500 and the lowest scoring of all in singletons. Its focus is debate on moral choices within a restricted circle of characters.

Shakespeare's entries shown in Figure 4 show no great variation from the overall pattern. They are encompassed within the scatter of segments by others, bracketed in both dimensions by higher and lower values in the larger cluster of play segments by other authors. Shakespeare's segments have a tendency to be higher in totals of the 500 words than the run of the others—to fall in the right-hand side of the chart—but this would not justify calling them exceptional.

[37] Slater, 23.
[38] Crystal, 6.

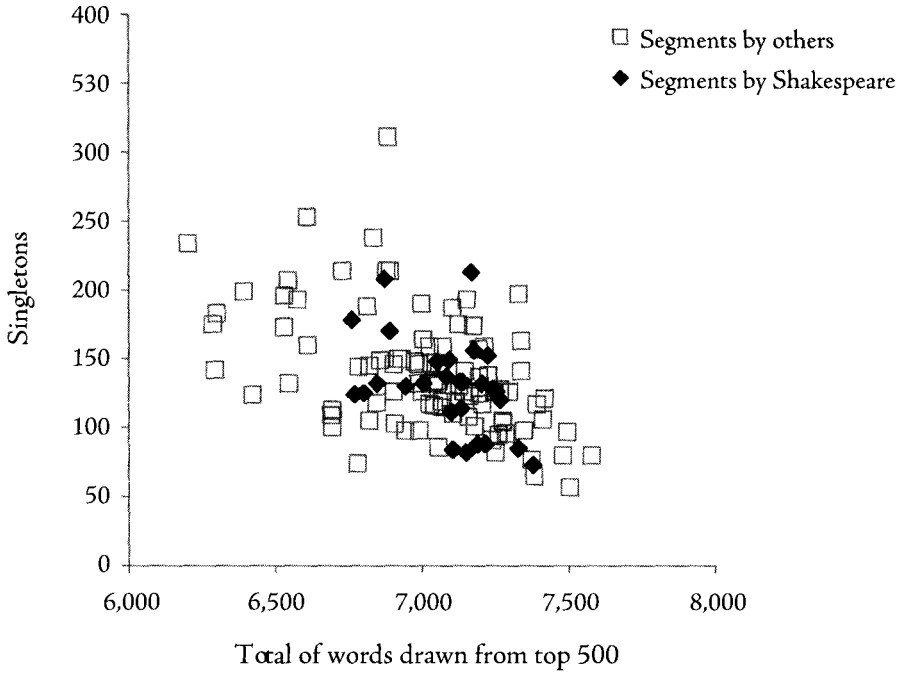

Figure 4: Counts for the top 500 words and for "singletons" for 121 10,000-word play segments.

The diamond to the lower right is *Much Ado about Nothing*, low in singletons and high in totals of the commoner words. The two plays to the top left of the Shakespeare cluster are *Henry V*, lower down, and *Love's Labor's Lost*, higher up.

On this showing, Slater was broadly right—overall, Shakespeare relies a little more than his peers on "commonplace" words—but this is by no means a defining or highly marked feature. In the present set, there are 8 play segments by others among the 121 that are more extreme on this measure than *Much Ado*, Shakespeare's highest-scoring segment. Shakespeare's practice in this regard, like his resort to very rare words, maps comfortably within the habits of his peers.

Osvaldo Rosso and his collaborators (one of whom is the author of this essay) recently suggested that if Shakespeare is exceptional, it is by being unusually close to a standard of word use. They calculate an average rate of use for all the words used in a large collection of plays and poems from the period and then observe how far the various texts included vary from this average. Shakespeare's practice is consistently close to the common standard.[39] We can do something similar with the current set of 121 play segments, which is restricted to one

[39] Osvaldo A. Rosso, Hugh Craig, and Pablo Moscato, "Shakespeare and Other English Renaissance Authors as Characterized by Information Theory Complexity Quantifiers," *Physica A* 388 (2009): 916–26.

mode (drama), more limited in time span than Rosso and company's corpus, and composed of blocks of standard size rather than whole plays. Here, I use a much simpler measure of deviation from the average. I examine only counts from the 100 commonest words, which in this set range from the article *the*, with 301 instances per segment on average, to the verb form *art* with 11. First, I find the difference between the play score and the average score for each word, and I then divide the result by the average score so that all the words are on the same basis. This produces scores for 100 deviations for each play. Finding the average of these in turn then provides a single composite figure for each play's overall deviation from the norm. The results confirm the earlier finding. A Shakespeare play, *As You Like It*, has the lowest average difference from the mean. (George Peele's play *The Battle of Alcazar* has the highest average difference.) Of the ten plays with lowest scores, six are Shakespeare's. There are no Shakespeare plays among the ten highest. Of the thirteen playwrights with three or more plays in the set, Shakespeare has the smallest overall distance from the mean.

What this might mean in terms of style is best explained by the opposite pattern. Characters in *The Battle of Alcazar* insistently modify nouns with *our*, so that this word occurs 91 times, compared to the average of 35; the characters studiously avoid the first-person pronoun *I*—it occurs just 78 times in the segment analyzed, compared to the average of 276 across the whole set. Greene's play is lopsided, the discourse of war largely unrelieved by more individual, less collective concerns. The analysis tells us that Shakespeare plays are generally more mixed, less extreme than this. The Shakespeare play with the highest average deviation is *Much Ado about Nothing*, with spikes in the frequency of *an* and *her*, reflecting a certain kind of allusive dialogue, an outburst of bickering about whether or not Dogberry is "an ass," and an unusual emphasis on women protagonists. Nevertheless, there are twenty-three segments in the set with higher overall deviations. *As You Like It*, *Othello*, *All's Well That Ends Well*, and *Romeo and Juliet* occupy four of the five lowest places in the deviations table. If we look only at their counts for the commonest word, *the*, we find they are 316, 361, 283, and 298, within a fairly narrow band around the average of 301. The overall lowest count for *the* in the set is 186 and an overall highest is 496. High frequencies are associated with passages of exposition and low frequencies with interactive dialogue between characters familiar with each other. The four Shakespeare plays mentioned present a balance between the two.

If we follow Rosso and company to recalculate the word averages without each of the authors' works in turn and use that as the basis for our measure of difference from the mean, Shakespeare still comes out lowest. Repeating the test using only 50 words, which means ending the list at *from*, which occurs on average 29 times in a segment, brought about similar results, with Shakespeare's

average the lowest of the thirteen, whether the averages for the words included his segments or not (Table 2). It is only when the segments are subdivided into genre sets that the pattern breaks down (Table 3).[40] Rosso and his collaborators conclude:

> Our results would suggest that we should think of William Shakespeare in word-use terms as a playwright who was extremely well attuned to the general practice of his day, producing an exceptionally large set of plays which all remain within a broadly defined common linguistic standard. If his dialogue in general strikes audiences and readers as just and satisfying, this impression may derive at least in part from his tendency in vocabulary use to fit with patterns of the drama of his day, which may of course reflect in turn a wider educated usage among English speakers of the time.[41]

Whatever the underlying explanation for this tendency—it is possible, for instance, that some of the effect comes from Shakespeare's influence on his successors—the basic finding that Shakespeare's practice brings him close to an overall mean is undeniable.

Together with the evidence that Shakespeare is in fact no different from his contemporaries in the number of different words he uses, this result helps redirect our attention toward the effects Shakespeare created within a normal range of vocabulary. When Shakespeare is placed on a computational-stylistic map alongside his contemporaries—that is, when we look at a "corpus Shakespeare"—he follows rules about vocabulary density and about the introduction of new words in new plays, rather than breaking them. If anything, his linguistic profile is exceptional in being unusually close to the norm of his time. His language is an extraordinary achievement with the regular resources of the English of his day rather than a linguistic aberration.

[40] To test further the persistence of Shakespeare's pattern of deviations from the mean, I made a new set of segments with only comedies included. There are fifty-seven of these, twelve of them by Shakespeare (see the Appendix for details). Six authors in this smaller set have three or more segments. When 100 words are used for calculations and all fifty-seven segments are included in the averages, Shakespeare once again has the lowest score of the six authors. If we recalculate the word means without the plays for each author in turn, however, Shakespeare has the highest score of the six. When 50 words are used, he is the second lowest when the mean from the full set of fifty-seven is used and the second highest when means excluding the relevant authors in turn are employed (see Table 3). It is, then, possible to find a combination of play group and method for which the rule does not apply. There is a definite and strong tendency for Shakespeare to use common words at a rate closely following the pattern of his peers, but it is a tendency rather than an unvarying pattern.

[41] Rosso et al., 925.

Table 2: Average deviation from the mean for thirteen authorial groups of play segments[1]

Author	Segments	All-segments mean		Nonauthorial mean	
		100 words	50 words	100 words	50 words
Chapman	10	0.288	0.243	0.295	0.248
Dekker	5	0.308	0.257	0.312	0.259
Fletcher	7	0.333	0.270	0.342	0.276
Greene	4	0.358	0.282	0.365	0.287
Heywood	5	0.303	0.297	0.307	0.302
Jonson	12	0.302	0.267	0.312	0.280
Lyly	8	0.319	0.283	0.329	0.293
Marlowe	5	0.326	0.301	0.332	0.307
Middleton	13	0.287	0.240	0.302	0.252
Peele	4	0.299	0.253	0.299	0.254
Shakespeare	28	0.265	0.224	0.279	0.235
Webster	3	0.282	0.276	0.288	0.282
Wilson	3	0.294	0.267	0.298	0.272

[1] Results of means calculated with all-segments and nonauthorial means are shown in Tables 2 and 3; data were obtained using 50 or 100 words.

Table 3: Average deviation from the mean for six authorial groups of comedy segments

Author	Segments	All-segments mean		Nonauthorial mean	
		100 words	50 words	100 words	50 words
Chapman	8	0.253	0.228	0.262	0.235
Dekker	4	0.281	0.225	0.251	0.222
Jonson	10	0.272	0.242	0.281	0.234
Lyly	3	0.344	0.342	0.279	0.246
Middleton	10	0.255	0.208	0.282	0.236
Shakespeare	12	0.250	0.211	0.284	0.238

Appendix

Author	Title	Genre[1]	Date[1]	Copy text	Copy text date	Group[2]
Beaumont	*Knight of the Burning Pestle*	Burlesque romance	1607	STC 1674	1613	1
Chapman	*All Fools*	Comedy	1604	STC 4963	1605	1
Chapman	*Blind Beggar of Alexandria*	Comedy	1596	STC 4965	1598	1
Chapman	*Bussy d'Ambois*	Foreign history	1604	STC 4966	1607	1
Chapman	*Gentleman Usher*	Comedy	1602	STC 4978	1606	1
Chapman	*Humorous Day's Mirth*	Comedy	1597	STC 4987	1599	1
Chapman	*May Day*	Comedy	1602	STC 4980	1611	1
Chapman	*Monsieur d'Olive*	Comedy	1604	STC 4984	1606	1
Chapman	*Revenge of Bussy d'Ambois*	Tragedy	1610	STC 4989	1613	1
Chapman	*Sir Giles Goose-Cap*	Comedy	1602	STC 12050	1606	1
Chapman	*Widow's Tears*	Comedy	1605	STC 4994	1612	1
Chettle	*Hoffman*	Tragedy	1602	STC 5125	1631	1
Day	*Isle of Gulls*	Comedy	1606	STC 6412	1606	1
Dekker	*2 Honest Whore*	Comedy	1605	STC 6506	1630	1
Dekker	*If This Be Not Good, the Devil Is in It*	Comedy	1611	STC 6507	1612	1
Dekker	*Old Fortunatus*	Comedy	1599	STC 6517	1600	1
Dekker	*Shoemaker's Holiday*	Comedy	1599	STC 6523	1600	1
Dekker	*Whore of Babylon*	Allegorical history	1606	STC 6532	1607	1
Fletcher	*Bonduca*	Tragedy	1613	Wing B1581	1647	1
Fletcher	*Faithful Shepherdess*	Pastoral	1608	STC 11068	1610	1
Fletcher	*Loyal Subject*	Tragicomedy	1618	Wing B1581	1647	1
Fletcher	*Mad Lover*	Tragicomedy	1617	Wing B1581	1647	1
Fletcher	*Monsieur Thomas*	Comedy	1615	Wing B1581	1647	1
Fletcher	*Valentinian*	Tragedy	1614	Wing B1581	1647	1
Fletcher	*Woman's Prize*	Comedy	1611	Wing B1581	1647	1
Ford	*Broken Heart*	Tragedy	1629	STC 11156	1633	3
Ford	*Fancies Chaste and Noble*	Comedy	1635	STC 11159	1638	3

Ford	*Lady's Trial*	Comedy	1638	STC 11161	1639	3
Ford	*Lover's Melancholy*	Tragicomedy	1628	STC 11163	1629	3
Ford	*Love's Sacrifice*	Tragedy	1632	STC 11164	1633	3
Ford	*Perkin Warbeck*	History	1633	STC 11157	1634	3
Ford	*'Tis Pity She's a Whore*	Tragedy	1632	STC 11165	1633	3
Greene	*Alphonsus*	Heroic romance	1587	STC 12233	1599	1
Greene	*Friar Bacon and Friar Bungay*	Comedy	1589	STC 12267	1594	1
Greene	*James IV*	History	1590	STC 12308	1598	1
Greene	*Orlando Furioso*	Romantic comedy	1591	STC 12265	1594	1
Haughton	*Devil and His Dame*	Comedy	1600	Wing G1580	1662	1
Haughton	*Englishmen for My Money*	Comedy	1598	STC 12931	1616	1
Heywood	*Four Prentices of London*	Heroic romance	1600	STC 13321	1615	1
Heywood	*If You Know Not Me, You Know Nobody*	History	1604	STC 13328	1605	1
Heywood	*Rape of Lucrece*	Tragedy	1607	STC 13363	1638	1
Heywood	*Wise Woman*	Comedy	1604	STC 13370	1638	1
Heywood	*Woman Killed with Kindness*	Tragedy	1603	STC 13371	1607	1
Jonson	*Alchemist*	Comedy	1610	STC 14755	1612	1
Jonson	*Bartholomew Fair*	Comedy	1614	STC 14753.5	1631	1
Jonson	*Case Is Altered*	Comedy	1597	STC 14757	1609	1
Jonson	*Catiline His Conspiracy*	Tragedy	1611	STC 14759	1611	1
Jonson	*Cynthia's Revels*	Comedy	1601	STC 14773	1601	1
Jonson	*The Devil Is an Ass*	Comedy	1616	STC 14754	1640	1
Jonson	*Epicœne*	Comedy	1609	STC 14751	1616	1
Jonson	*Every Man in His Humor*	Comedy	1598	STC 14766	1601	1
Jonson	*Every Man out of His Humor*	Comedy	1599	STC 14767	1600	1
Jonson	*Magnetic Lady*	Comedy	1632	STC 14754	1640	2
Jonson	*New Inn*	Comedy	1629	STC 14780	1631	2
Jonson	*Poetaster*	Comedy	1601	STC 14781	1602	1
Jonson	*Sad Shepherd*	Comic pastoral	1637	STC 14754	1640	2
Jonson	*Sejanus His Fall*	Tragedy	1603	STC 14782	1605	1
Jonson	*Staple of News*	Comedy	1626	STC 14753.5	1631	2

Appendix (*continued*)

Author	Title	Genre[1]	Date[1]	Copy text	Copy text date	Group[2]
Jonson	*Tale of a Tub*	Comedy	1633	STC 14754	1640	2
Jonson	*Volpone*	Comedy	1606	STC 14783	1607	1
Kyd	*Spanish Tragedy*	Tragedy	1587	STC 15086	1592	1
Lodge	*Wounds of Civil War*	Classical history	1588	STC 16678	1594	1
Lyly	*Campaspe*	Classical legend (comedy)	1584	STC 17048	1584	1
Lyly	*Endymion*	Classical legend (comedy)	1588	STC 17050	1632	1
Lyly	*Gallathea*	Classical legend (comedy)	1585	STC 17080	1592	1
Lyly	*Love's Metamorphosis*	Pastoral	1590	STC 17082	1601	1
Lyly	*Midas*	Comedy	1589	STC 17083	1592	1
Lyly	*Mother Bombie*	Comedy	1589	STC 17084	1594	1
Lyly	*Sappho and Phao*	Classical legend (comedy)	1584	STC 17086	1584	1
Lyly	*Woman in the Moon*	Comedy	1593	STC 17090	1597	1
Marlowe	*Doctor Faustus* (A-text)	Tragedy	1592	STC 17429	1604	1
Marlowe	*Edward II*	History	1592	STC 17437	1594	1
Marlowe	*Jew of Malta*	Tragedy	1589	STC 17412	1633	1
Marlowe	*1 Tamburlaine the Great*	Heroic romance	1587	STC 17425	1590	1
Marlowe	*2 Tamburlaine the Great*	Heroic romance	1588	STC 17425	1590	1
Marston	*Antonio's Revenge*	Tragedy	1600	STC 17474	1601	1
Marston	*Sophonisba*	Tragedy	1605	STC 17488	1606	1
Middleton	*Chaste Maid in Cheapside*	Comedy	1613	STC 17877	1630	1
Middleton	*Game at Chess*	Political satire	1624	Trinity College MS	1624	2
Middleton	*Hengist, King of Kent*	Tragedy	1620	Ed. R. C. Bald	1938	2
Middleton	*Mad World, My Masters*	Comedy	1605.2	STC 17888	1608	1
Middleton	*Michaelmas Term*	Comedy	1604	STC 17890	1607	1
Middleton	*More Dissemblers Besides Women*	Comedy	1614	Wing M1989	1657	1
Middleton	*Nice Valor*	Comedy	1622	Wing B1581	1647	2
Middleton	*No Wit, No Help Like a Woman's*	Comedy	1611.1	Wing M1985	1657	1

Middleton	*Phoenix*	Comedy	1603	STC 17892	1607	1
Middleton	*Puritan*	Comedy	1606.1	STC 21531	1607	1
Middleton	*Revenger's Tragedy*	Tragedy	1606.2	STC 24150	1608	1
Middleton	*Second Maiden's Tragedy*	Tragedy	1611.2	Malone Society	1909	1
Middleton	*Trick to Catch the Old One*	Comedy	1605.1	STC 17896	1608	1
Middleton	*Widow*	Comedy	1615	Wing J1015	1652	1
Middleton	*Witch*	Tragicomedy	1616	Malone Society	1949	1
Middleton	*Women Beware Women*	Tragedy	1621	Wing M1989	1657	2
Middleton	*Yorkshire Tragedy*	Tragedy	1605.3	STC 22340	1608	2
Middleton	*Your Five Gallants*	Comedy	1607	STC 17907	1608	1
Munday	*John a Kent and John a Cumber*	Pseudohistory	1589	Malone Society	1923	1
Nashe	*Summer's Last Will and Testament*	Comedy	1592	STC 18376	1600	1
Peele	*Arraignment of Paris*	Classical legend (pastoral)	1581	STC 19530	1584	1
Peele	*Battle of Alcazar*	Foreign history	1589	STC 19531	1594	1
Peele	*Edward I*	History	1591	STC 19535	1593	1
Peele	*King David and Fair Bethsabe*	Biblical history	1587	STC 19540	1594	1
Porter	*1 Two Angry Women of Abington*	Comedy	1588	STC 20121.5	1599	1
Rowley	*All's Lost by Lust*	Tragedy	1619	STC 21425	1633	1
Shakespeare	*All's Well That Ends Well*	Comedy	1604	STC 22273	1623	1
Shakespeare	*Antony and Cleopatra*	Tragedy	1606	STC 22273	1623	1
Shakespeare	*As You Like It*	Comedy	1599.2	STC 22273	1623	1
Shakespeare	*Comedy of Errors*	Comedy	1593	STC 22273	1623	1
Shakespeare	*Coriolanus*	Tragedy	1608	STC 22273	1623	1
Shakespeare	*Cymbeline*	Tragicomedy	1609.2	STC 22273	1623	1
Shakespeare	*Hamlet*	Tragedy	1600	STC 22276	1604	1
Shakespeare	*1 Henry IV*	History	1596	STC 22280	1598	1
Shakespeare	*2 Henry IV*	History	1597.2	STC 22288	1600	1
Shakespeare	*Henry V*	History	1598.2	STC 22273	1623	1
Shakespeare	*Julius Caesar*	Tragedy	1599.1	STC 22273	1623	1
Shakespeare	*King John*	History	1596	STC 22273	1623	1
Shakespeare	*King Lear* (quarto)	Tragedy	1604	STC 22292	1608	1

Appendix (*continued*)

Author	Title	Genre[1]	Date[1]	Copy text	Copy text date	Group[2]
Shakespeare	*Love's Labor's Lost*	Comedy	1594	STC 22294	1598	1
Shakespeare	*Merchant of Venice*	Comedy	1596	STC 22296	1600	1
Shakespeare	*Merry Wives of Windsor*	Comedy	1597.1	STC 22299	1602	1
Shakespeare	*Midsummer Night's Dream*	Comedy	1595.1	STC 22302	1600	1
Shakespeare	*Much Ado about Nothing*	Comedy	1598.1	STC 22304	1600	1
Shakespeare	*Othello*	Tragedy	1603	STC 22305	1622	1
Shakespeare	*Richard II*	History	1595.3	STC 22307	1597	1
Shakespeare	*Richard III*	History	1592	STC 22314	1597	1
Shakespeare	*Romeo and Juliet*	Tragedy	1595.2	STC 22323	1599	1
Shakespeare	*Taming of the Shrew*	Comedy	1592	STC 22273	1623	1
Shakespeare	*Tempest*	Comedy	1610	STC 22273	1623	1
Shakespeare	*Troilus and Cressida*	Tragedy	1602	STC 22331	1609	1
Shakespeare	*Twelfth Night*	Comedy	1601	STC 22273	1623	1
Shakespeare	*Two Gentlemen of Verona*	Comedy	1590	STC 22273	1623	1
Shakespeare	*Winter's Tale*	Tragicomedy	1609.1	STC 22273	1623	1
Tourneur	*Atheist's Tragedy*	Tragedy	1609	STC 24146	1611	1
Webster	*Devil's Law Case*	Tragicomedy	1617	STC 25173	1623	1
Webster	*Duchess of Malfi*	Tragedy	1614	STC 25176	1623	1
Webster	*White Devil*	Tragedy	1612	STC 25178	1612	1
Wilson	*Cobbler's Prophecy*	Comedy	1590	STC 25781	1594	1
Wilson	*Three Ladies of London*	Moral	1581	STC 25784	1584	1
Wilson	*Three Lords and Three Ladies of London*	Moral	1588	STC 25783	1590	1

[1] For plays by Middleton, publication dates follow those of Gary Taylor and John Lavagnino, gen. eds., *Thomas Middleton: The Collected Works*; the order within a year is indicated by a decimal point. Publication dates for plays by Shakespeare follow Greenblatt, gen. ed., *The Norton Shakespeare*; the order within a year is indicated by a decimal point. Publication dates for all other plays, and genres for all plays in the appendix, follow Harbage and Schoenbaum.

[2] Groups are defined as follows: group 1, 121 plays of more than 10,000 words dating from 1580 to 1619; group 2, 11 later or shorter plays to complete the Jonson and Middleton canons; and group 3, Ford plays.