

God-Language and Scepticism in Early Modern England

An Exploratory Study Using Corpus Linguistics Analysis as a Form of Distant Reading

William M. Hamlin (Washington State University, USA)

Abstract This essay incorporates new trends of corpus linguistics research within an investigation of sceptical thought, particularly in the period of roughly 1580 to 1620. The analysis uses several standard texts, such as the Shakespeare First Folio (1623), the Florio translation of Montaigne (1613 ed.), the collected plays of Marlowe, Bacon's *Advancement of Learning* (1605), etc., juxtaposed against larger reference corpora, such as a fully-keyed collection of several thousand English books printed between 1580 and 1620. Such juxtapositions enable contrastive analysis using various techniques of collocation, proximity, and syntactic pattern examination.

My purpose in this essay is to describe and reflect upon a preliminary exploration of the value of corpus analytic techniques for the study of ideological change during the sixteenth and seventeenth centuries. Specifically, I have made an investigative foray into the evolution of god-language in English printed texts from this period, focusing especially on the spread of epistemological and religious scepticism in multiple discursive realms.¹ By 'god-language' I mean language which registers the ways in which human beings – including fictional characters – imagine, describe, and discuss their conceptions of divinity, as well as the locutions upon which they rely as they address, worship, petition, doubt, or condemn divine beings (along with the affiliated concepts, assumptions, and institutions underwritten by presumed divine authority). Early modern scepticism, meanwhile, is a topic that has been extensively studied by many scholars over the past fifty years, so before I began my inquiry I already had a strong sense of what

1 When I speak of ideological change, I do not wish to limit ideology to forms of false consciousness. Rather, I follow Raymond Williams, Terry Eagleton, David Hawkes, and others in broadening the term so as to include systems of belief characteristic of a particular class, culture, or social group, as well as systems of illusory beliefs. I am nonetheless particularly intrigued by assumptions, claims, and convictions which are non-falsifiable. See Williams (1977, pp. 55-71); Williams (1983, pp. 153-157); Eagleton (1991, pp. 1-31), and Hawkes (2003, *passim*, esp. pp. 1-14, 15-37). See also Stephen Greenblatt's comments on the frequent insensitivity of modern discussions of ideology (2001, pp. 45-46).

I was likely to learn.² But because I am a literary historian rather than a computational linguist, I came to this project with methodological biases and levels of technical ignorance which have predisposed me to investigate topics and interpret results in ways that may seem peculiar, even counterintuitive, to scholars with greater expertise in digital humanities methods and computer-assisted discourse analysis.³

I must also preface this essay with several caveats. First of all, the digital corpus with which I have primarily worked is the pared-down version of EEBO-TCP housed within the CQPweb processor at Lancaster University.⁴ This corpus is currently comprised of 44,422 fully-keyed texts, just under two-thirds of the total number of documents eventually slated for complete digital transcription by EEBO-TCP. Nonetheless, the corpus is broadly representative of EEBO's full coverage; it contains increasingly large numbers of texts within the unfolding decades of the sixteenth and seventeenth centuries, and altogether it comprises a total of more than 1.2 billion words.⁵ I should note, however, that the method of 'tokenization' used in its CQPweb conversion has allowed for typographic entities other than alphabetic units (e.g., punctuation marks and numerical digits) to count as 'words' – a practice that presents drawbacks as well as undeniable advantages.⁶ Finally, while this corpus has been curated so as to provide a significant degree of word-level annotation, it cannot be queried in some of the ways that other curated corpora currently enable (e.g., the Shakespeare and Early Modern Drama sites in WordHoard). It also displays a small number of residual transcription errors. Nonetheless, the CQPweb version of EEBO constitutes a massive digital database of English printed

2 See, e.g., Popkin (2003); Schmitt (1972); Larmore (1998); and Hamlin (2005). Williams (1983) still stands as a classic model of the discussion of denotative evolution and cultural relevance for specific English words, e.g. «ideology».

3 Two recent essays from which I have learned a good deal are each co-authored by Jonathan Hope and Michael Witmore (2014, 2010). I take the term «distant reading» from its various iterations in the work of Franco Moretti – most recently in his monograph of that title (2013). See Moretti (2005, pp. 1-2).

4 I am much indebted to Professors Paul Rayson, Andrew Hardie, and Alistair Baron, who introduced me to CQPweb during my visit to the University Centre for Computer Corpus Research on Language (UCREL) at Lancaster University during May and June of 2013. CQP is an acronym for Corpus Query Processor; EEBO-TCP is the common abbreviation for *Early English Books Online - Text Creation Partnership*. At the time I gathered my evidence for this study, EEBO-TCP offered a database of 40,061 keyed full texts, but that number has now risen considerably. CQPweb's Version 3 of EEBO-TCP, meanwhile, offers a database of 44,422 fully keyed texts. But that figure, too, will rise when Version 4 is loaded to the processor.

5 To be precise: 1,202,214,511 words. Decades are calculated in CQPweb not from 1 to 0 but from 0 to 9; for example, the 1580s run from 1580 to 1589 rather than from 1581 to 1590.

6 For a description of the development of the CQPweb version of EEBO, see Pumfrey, Rayson, Mariani (2012, esp. pp. 401-406).

documents ranging from 1473 to 1700, and its query processor allows for an array of search and analysis techniques which can generate vast batches of quantitative and statistical information with impressive speed.

After spending an initial few days conducting miscellaneous searches and becoming familiar with CQPweb, I decided that my first substantial task should be the creation of a stable lexicon of 'god-terms' and 'doubt-terms' that I could then study in depth through scrutiny of such forms of data as the following: (1) diachronic distribution of words and phrases across the full EEBO time-span, (2) dispersion of terms within multiple texts (and groups of texts), (3) keyword frequency comparisons within juxtaposed corpora, and (4) collocation and proximity analyses (again attending to diachronic change among specific collocates). I thus prepared a preliminary list of 175 relevant words which I then reduced through a process of elimination to a final list of 104; this reduction was based upon scores of keyword searches, extensive examination of specific attestations, and frequent recourse to the Oxford Historical Thesaurus. The word «lord», for instance, cannot serve usefully as a god-term in early modern English because it is too often employed as a form of address (e.g., «good my lord»); similarly, «pray» and «curse», despite their apparent promise as god-verbs, must be disqualified for similar reasons, «curse» being too secular in general usage and «pray» figuring too often in locutions such as «pray tell» and «I pray you, madam». Indeed, many words which function simultaneously as nouns and verbs (e.g., «curse», «grace», «sacrifice», «minister», «host», «spirit», «elect») prove ultimately unsuitable for a god-lexicon, their very flexibility rendering their denotative range too broad for a study such as mine. And many more specific religious terms – «rosary», «atone», «conventicle», «ordain», «absolution», «manna», «eucharist», and so on – appear with such comparatively low frequencies that their usage tends to be less valuable from a statistical perspective than that of other, more common, words.⁷

Here, then, are the god-terms and doubt-terms I have chosen for my lexicon (see Tables 1 and 2). Table 1 presents these terms using spelling variations and statistical data derived from EEBO-TCP in its current iteration; Table 2 presents the same terms using analytic techniques and spelling regularisation software available in Version 3 of EEBO on CQPweb. I wish to stress that this is a provisional list, and that I expect to modify it as I move forward. It is also profoundly Christian in orientation – and predictably so, given the verbal corpus from which it is derived. At the same time, it is capable of reflecting shifts in attitude towards other religious

7 God-phrases, meanwhile (e.g. «holy ghost», «grace of god», «eternal soul», etc.), are far too numerous to limit to a stable lexicon, but I have nonetheless attempted to keep track of roughly forty such phrases through multiple distribution and collocation searches. With a few major exceptions («christ[s]», «mahomet[s]», «satan[s]», etc.), I have excluded proper names from my god-lexicon.

outlooks. The words «faith(s)», «religion(s)», and «doctrine(s)», for instance, have the potential to be deployed in non-pejorative ways regarding non-Christian forms of devotion or spirituality, and varying frequencies in such terms as «heathen(s)», «pagan(s)», «infidel(s)», and «heresy(ies)» can also expose attitudinal shifts. For the words «god» and «heaven» I have included separate statistical computations for singular, plural, and conjoined forms, but it is crucial to remember that in each case the plural spelling also comprehends large numbers of the singular possessive (i.e., «god's»), since the typographic use of the possessive apostrophe was uncommon until the later decades of the seventeenth century.⁸ I have presented all initial letters in lower-case format, and indeed all my searches have been non-case sensitive. Finally, the lexicon moves gradually towards more secular terms at the end, so that ideological developments of the sort that interest me have a better chance of making themselves visible.

A few preliminary comments are in order. First, it is abundantly clear, though scarcely surprising, that early modern English is saturated with religious language. The word «god», for instance, surfaces at least once in 78% of all printed texts dating from 1473 to 1700, and in those texts it displays an average frequency of 121 appearances.⁹ Words such as «faith(s)», «soul(s)», «sin(s)», «church(es)», and «heaven(s)» are likewise extremely common, all of them manifesting themselves in more than half the period's documents. «Christ(s)», «jesu(s)», «saint(s)», «christian(s)», «priest(s)», «prayer(s)», «worship(s)», «scripture(s)», «gospel(s)», «eternal(ly)», «angel(s)», «devil(s)», and «hell(s)» each appear in more than a third of all printed texts, «righteous(ly)», «almighty(ies)», «salvation(s)», «saviour(s)», «providence(s)», «rejoice(s)», «jew(s)», and «heathen(s)» in more than a quarter, and the adjective «holy» is used with far greater frequency than the adjective «natural» (both in raw numbers and in rates of dispersion). It will of course be news to no one that sixteenth- and seventeenth-century English is much more oriented towards religious concerns than the English of today, but numbers such as those in Tables 1 and 2 demonstrate this fact with decisive clarity.

Attending to the diachronic distribution of god-terms across the nearly 230-year time-span covered by EEBO yields results that, once again, will be largely unsurprising to scholars familiar with the period. I should

8 For most other nouns I have used a combined singular/plural search; a search for «soul» and «souls», for example, is represented as «soul(s)».

9 My figure of 78% was derived from an EEBO-TCP search on 23 January 2014. At that point there were 40,061 fully-searchable texts within the database, though as noted earlier this figure has now risen. Version 3 of EEBO on CQPweb, by contrast, currently provides access to 44,422 fully-searchable texts, and «god» appears in 77.35% of them. Searching for «god(s)» across all relevant decades, I find the highest frequency in the 1530s (95.35%) and the lowest in the 1680s (71.96%). Version 3 of EEBO also employs a spelling regularization program called «VARD», designed by Professor Alistair Baron at the University of Lancaster.

reiterate, however, that the CQPweb version of EEBO currently includes fewer than two-thirds of the roughly 70,000 texts eventually planned for complete digitisation; when the full body of digitised texts is incorporated within the CQPweb processor, the resulting numbers will differ from those I offer here. Nonetheless, I believe that my current figures are broadly indicative of major trends within the two centuries under examination. Terms such as «jesu(s)», «scripture(s)», «gospel(s)», «sabbath(s)», «rejoice(s)», «amen(s)», «jew(s)», «blaspheme(s)», «sinner(s)», «heresy(ies)», and «purgatory(ies)», for instance, display a sharp rise in incidence during the 1520s, 1530s, and 1540s, a fact that makes perfect sense given the growth of English Reformation polemic during precisely these years. Similarly, such terms as «church(es)», «prayer(s)», «conscience(s)», «spiritual(ly)», «baptise(s)», «baptism(s)», «jew(s)», «heathen(s)», «blasphemy(ies)», «unholy(ily)», and «unnatural(ly)» occur with distinctly increased frequency during the 1640s and 1650s, very likely in concert with social anxieties surrounding the English Civil War, the rise of presbyterianism, the abolishment of the episcopacy, and the establishment of the Commonwealth.¹⁰ The terms «heaven(s)», «rejoice(s)», «hell(s)», «devil(s)», «antichrist(s)», «heathen(s)», «blaspheme(s)», and «blasphemy(ies)» exhibit declining levels of usage during the latter half of the seventeenth century, while the terms «reason(s)», «scepticism(s)», «sceptical(ly)», «uncertain(ly)», «perhaps», and «possible(ly)» show gradual rates of increase during the same period. The 1590s and 1600s, meanwhile, present something of a puzzle insofar as they reveal a sharp decline in the use of certain terms along with a sharp increase in the use of others. The words «god(s)», «christ(s)», «jesu(s)», «faith(s)», «almighty(ies)», «holy(ily)», «scripture(s)», «gospel(s)», «salvation(s)», «amen(s)», and «sabbath(s)», for example, all show strikingly diminished rates of usage during this twenty-year period (and particularly during the 1590s), while the words «hell(s)», «despair(s)», «damn(s)», «infidel(s)», and «doubtful(ly)» all display a distinct rise in incidence. How may we account for this? Are such results related to *fin de siècle* English monarchical anxiety? Are they based on flawed data – too few documents, for instance, or skewed textual representation within the two decade sub-corpus? At the moment I cannot answer these questions, but my strong suspicion is that the statistical figures for the 1590s and 1600s, even though they will change when we have more thorough information, are nonetheless indicative of slight ideological shifts during the period. Intriguingly, moreover, these are precisely the decades when Marlowe, Shakespeare, Jonson, Middleton, and other English playwrights wrote a substantial portion of their plays.

10 The Thomason Tracts are included among the keyed full texts from the 1640s, and thus the total number of published texts from that decade greatly exceeds that of previous decades.

But let us return to the evidence. One useful way to move forward from broad-based keyword distribution analysis is to juxtapose the presence of specific keywords in multiple sub-corpora. As an example of such investigation I have examined the frequency of two dozen keywords across sixteen separate texts published between 1583 and 1623, among them Bacon's *Advancement of Learning* (1605), Burton's *Anatomy of Melancholy* (1621), Foxe's *Acts and Monuments* (1583), Hakluyt's *Principal Navigations* (1599), Montaigne's *Essays* (1613), Raleigh's *History of the World* (1617), Shakespeare's *First Folio* (1623), and Spenser's *Faerie Queene* (1596). Fourteen of these texts are individual printed volumes such as those just listed, but two are amalgamations: one of them a compilation of the various first editions of Marlowe's seven plays, and the other a massive corpus of 4716 separate works published between 1580 and 1619 which currently comprise the section of fully-transcribed digital texts in CQPweb's Version 3 of EEBO.¹¹ This latter corpus serves in essence as a control text against which the other fifteen corpora may be juxtaposed. Thus, for example, the keyword «god(s)» appears at a rate of 4093.56 instances per million words ('ipmw') in the forty-year corpus of texts dating from 1580 to 1619 (hereafter '1580-1619'). In Hooker's *Laws of Ecclesiastical Polity* (1604), however, it appears at the much higher rate of 6222.23 ipmw (52% more frequently), while in Montaigne's *Essays* (1613) it displays a rate of only 1038.82 ipmw (74% less frequently).¹² Similarly, the keyword «hell(s)» has an ipmw frequency of 163.01 in the broad 1580-1619 corpus, but in Marlowe's plays its frequency is a stunning 727.85 (446% higher), while in Bacon's *Advancement* (1605) its frequency is only 10.54 (93% lower). Hell tended to occupy Marlowe's thoughts far more than it did Bacon's.

11 In addition to the titles and corpora just listed, I have also included the following volumes among my sixteen texts: Holinshed's *Chronicles* (1587), Hooker's *Laws of Ecclesiastical Polity* (1604), Jonson's *Workes* (1616), Plutarch's *Philosophie* (1603), Seneca's *Tenne Tragedies* (1581), and Sidney's *Arcadia* (1593). As for my compilation of Marlovian drama, I have drawn on the 1590 edition of *Tamburlaine I and II* (STC 17425), the 1594 edition of *Dido, Queen of Carthage* (STC 17441), the 1594 edition of *Edward the Second* (STC 17437), the 1594 edition of *The Massacre at Paris* (STC 17423), the 1604 edition («A-Text») of *Doctor Faustus* (STC 17429), and the 1633 edition of *The Jew of Malta* (STC 17412). Readers will note that I rely both on first editions (e.g., Shakespeare's *First Folio*, 1623) and on second or third editions (e.g., Montaigne's *Essayes*, 1613), depending on the texts currently available in Version 3 of EEBO-TCP on CQPweb. I use John Florio's spelling, *Essayes*, for Montaigne's book; I know from close scrutiny that the 1603 and 1613 editions are almost identical in substance, the second merely correcting some of the more egregious typographic errors of the first, as well as introducing various new errors. I had hoped to include one of the early editions of the Authorized ("King James") Version of the Bible, but none of the printings from 1611 to 1619 are currently held within EEBO Version 3 on CQPweb.

12 For the twenty-year period of 1590-1609, however, the ipmw figure is 3571.42, a drop of about 12.75%.

Analysis of this sort is consistently intriguing and occasionally surprising – but also extremely time-consuming despite the speed of the CQPweb processor. One might easily have predicted that the ipmw rates for such keywords as «christ(s)», «faith(s)», «scripture(s)», «holy(ily)», «sin(s)», and «damn(s)» would far exceed 1580-1619 averages in a work like Foxe's *Acts and Monuments* (1583), and that the same keywords – with the exception of «damn(s)» – would fall far short of those averages in Shakespeare's First Folio. Similarly, one is unlikely to be struck by the fact that works such as Hakluyt's *Principal Navigations* and Montaigne's *Essays* display much higher-than-average frequency rates for the keyword «custom(s)». But that «custom(s)» appears still more often in Hooker's *Ecclesiastical Polity* than in Montaigne's *Essays* is perhaps less predictable, as is the fact that Hooker's use of «reason(s)» in both verb and noun forms far exceeds Montaigne's, as well as Bacon's in the *Advancement*. Sidney's *Arcadia* (1593), meanwhile, exhibits strikingly lower frequency rates for multiple god-terms than do either Spenser's *Faerie Queene* or Shakespeare's First Folio – but strikingly higher rates for such words as «doubt(s)», «reason(s)», and «nature(s)». May we therefore regard it as a more secular text? How, if at all, does this fact challenge us to recalibrate our sense of the general secularity of Shakespeare?

A useful variant of this form of analysis lies in the direct comparison of keyword frequency rates in pairs of corpora. In essence, CQPweb allows us to study the relative frequencies of specific words as they appear in juxtaposed bodies of text, and it ranks these words according to the magnitude of the disparity between their rates of usage. The specific statistical measure of disparity chosen by the CQPweb designers for this purpose is referred to as a «log-likelihood»: it contrasts observed frequencies with expected frequencies, using a logarithmic formula which takes into account the size (in total numbers of words) of the juxtaposed texts.¹³ Thus, for instance, a comparative keyword analysis of Shakespeare's First Folio and Montaigne's *Essays* reveals that the pronoun «you» is the word which exhibits the most striking degree of disparity between the two corpora. Specifically, this word appears only 765 times in Florio's Montaigne, whereas it makes 13,577 appearances in the Folio – a rate of almost nine times greater relative frequency.¹⁴ This translates to a log-likelihood figure

13 The log-likelihood calculator used within CQPweb is presented and explained at the following site: <http://ucrel.lancs.ac.uk/llwizard>. A log-likelihood of just 3.84, for instance, represents a disparity between observed and expected keyword frequencies that ranks in the 95th percentile for statistical improbability; a log-likelihood of 15.13 represents a statistical improbability ranking in the 99.99th percentile. Most of the log-likelihoods I record in the following pages are much higher than 15.13: many of them in the hundreds, and some in the thousands. They thus represent instances of extraordinary lexical disparity between the textual bodies being contrasted.

14 To be precise, 8.764 times greater relative frequency. The 1613 edition of Montaigne's

of 6617.87, which essentially indicates the colossal improbability that such a common word as «you» would nonetheless appear at such vastly differing rates as it does in Shakespeare and Montaigne. As students of early modern literature we immediately recognize that the reason for this difference is fundamentally a matter of genre: a playwright is more likely to deploy pronouns with great frequency than is an essayist or moral philosopher writing in sustained expository prose. And indeed we find that various other first- and second-person pronouns follow «you» in displaying high log-likelihood figures in the Folio: «your», «I», «thou», «thy», «my», «thee», «me», «thine», and «yours». Interestingly, however, third-person plural pronouns such as «their», «they», «them», and «themselves» appear with greater relative frequency in Montaigne than in Shakespeare. Without examining individual attestations in systematic detail I can only speculate as to why this is true, but my guess is that the demand in playwriting for third-person pronouns is comparatively lower than in expository prose because the very nature of dramatic composition foregrounds presence over absence, mandating a greater need for second- than third-person speech forms.

My concern at the moment, however, lies with god-language, and I therefore wish to point to a few of the more intriguing results in my specific searches. Since none of the early editions of the King James Bible (1611, 1612, 1614, etc.) are included as fully-transcribed texts in CQPweb's Version 3 of EEBO, I have relied instead on the 1561 edition of the Geneva Bible as an example of an early modern text dense with god-terms.¹⁵ Thus, for instance, when we juxtapose keywords in the Geneva Bible against those in Foxe's *Acts and Monuments* (1583), we find that «lord», «god», «altar», «tabernacle», «covenant», «rejoice», «righteous», «angel», and «sabbath» all display log-likelihoods revealing huge levels of disparity between observed and expected rates of frequency. Moving in the other direction, however, we see that «bishop», «church», «pope», «christ», «christians», «christian», «popes», «churches», «faith», «holy», «doctrine», «scripture», «priests», and «heresies» appear with much greater frequency in Foxe than in Geneva, which in turn confirms our sense of Foxe's book as one deeply embedded in the doctrinal controversies of its day. Jumping forward two decades to the 1604 edition of Hooker's *Ecclesiastical Polity*, we discover that far fewer obvious god-terms predominate in Geneva, while far more such terms appear with high log-likelihood values in Hooker. Some of these are identical to those in Foxe, such as «scripture», «church», «churches», and «christian», but others are new, among

Essays is comprised of 514,047 words, while the 1623 First Folio is slightly more than twice that size, at 1,041,042 words. Bear in mind that 'words' include digits, punctuation marks, and other non-alphabetic signs.

¹⁵ I.e., STC 2095. The first edition of the Geneva Bible, published a year earlier in 1560, is identified as STC 2093.

them «divine», «apostle», «apostles», «saint», «papists», «savior», and «reformation». A further comparison between Foxe and Hooker would clearly be useful, as would comparisons between such histories as those of Holinshed and Raleigh, or such fictions as those of Sidney and Spenser. But in the interests of displaying the nature of largely religious versus largely secular language in the decades under consideration, I turn now to more striking juxtapositions: those between Montaigne and Geneva, Geneva and Shakespeare, and Shakespeare and Montaigne – the last of these with additional reference to the small corpus of Marlovian plays as well as to the much larger corpus of 2395 texts from the twenty-year period of 1590-1609.¹⁶

In the first of these juxtapositions we find, not surprisingly, that the Geneva Bible exhibits massively greater frequencies in the use of such words as «god», «christ», «jesus», «holy», «church», «priest», «sin», «prophet», «temple», «pray», «blessed», «heaven», and «faith». Montaigne's *Essays*, meanwhile, display a similarly massive lopsidedness with respect to «reason», «fortune», «opinion», «natural», «virtue», «nature», «custom», «senses», «divers», «imagination», «body», and «experience». Indeed the only god-term in Montaigne that exceeds its equivalent in Geneva by a log-likelihood factor of 200.00 or higher is «divine». And when we turn to a comparison of the *Essays* with the twenty-year period of 1590-1609 (the period during which Florio's translation was prepared, mainly from about 1595 to 1603), a similar picture emerges, despite the significant drop in god-language during the 1590s. «Christ», «god», «church», «faith», «sin», «holy», «jesus», «pope», «doctrine», «prayer», «heaven», and other similar terms are still hugely predominant in the broader corpus, while terms such as «judgement», «self», «fortune», «philosophy», «natural», «imagination», «reason», «custom», «discourse», «health», «science», «fantasy», «women», and «uncertainty» predominate in Montaigne. There are in fact no god-terms from the *Essays* among the top 600 most disparate words in the juxtaposed corpora. Does this warrant a claim that the *Essays* bear greater lexical resemblance to Geneva than to 1590-1609? Perhaps. My supposition, however, is that it suggests that both Geneva and 1590-1609 differ vastly from the *Essays* in overall lexical character – and in ways that may underline common features in their deployment of god-language. A comparison of Geneva and 1590-1609 reveals that only 32 of the top 600 most disparate words are god-terms (5.33%), but that the log-likelihood figures throughout this list are extremely high, none of them falling below 200.00. Clearly this is a juxtaposition that requires further investigation.

¹⁶ The 40-year corpus of 1580-1619 would be more revealing and valuable here, but at present it is too large for keyword juxtaposition analysis on CQPweb; it is comprised of more than 100 million words.

With Shakespeare's First Folio the case is somewhat different. Comparing Geneva and the Folio we find that while many of the former's god-terms are still dominant (e.g., «god», «christ», «jesus», «church», «temple», «holy», «sin», «doctrine», and so on), a few have dropped from prominence (e.g., «blessed», «heaven», «faith»), and others have become dominant in Shakespeare, among them «hell», «damned», and «devil». With plays such as *Othello* and *Hamlet* in mind it is not difficult to imagine why this is the case, but the fact is nonetheless striking. When we compare the Folio to 1590-1609 more broadly, however, we note that «hell», «damned», and «devil» have vanished from high predominance in the Folio, while «heaven», «jew», and «fiend» have taken their places. «God», «christ», «church», «religion», «doctrine», «sins», «jesus», «holy», «salvation», and «spiritual», meanwhile, all retain enormous log-likelihood dominance within the 1590-1609 corpus. If the 1606 ban on religious language in stage-plays had any significant impact, then perhaps we see evidentiary traces of that impact in indications such as these from the Folio.¹⁷ But my guess is that the ratio between published plays and the full textual corpus – particularly with respect to absolute numbers of words – is so low that any inferences regarding the consequences of the 1606 ban are fraught with statistical difficulty. The steep decline in god-language during the 1590s, moreover, presents further vexing complications for any study of Shakespeare's god-lexicon relative to the period during which he composed the majority of his plays.

With Shakespeare and Montaigne, as I have noted, the word «you» surfaces as the most striking lexical anomaly between the two corpora – at least with regard to disparities among single word frequencies. In terms of god-language, however, the corpora exhibit comparatively little differentiation. Only nine god-terms occur among the top 600 most disparate words between the Folio and the *Essays* (1.5%), six of these in Shakespeare («heaven», «hell», «devil», «faith», «blessed», «worship»), and the remaining three in Montaigne («religion», «divine», «divinity»).¹⁸ Even here

17 The «Act to Restrain Abuses of Players» (27 May 1606) stipulates that «if at any time or times, after the end of this present Session of Parliament, any person or persons do or shall in any stage play, interlude, show, May game, or pageant, jestingly or profanely speak or use the holy name of God, or of Christ Jesus, or of the Holy Ghost, or of the Trinity, which are not to be spoken but with fear and reverence, such person or persons shall forfeit for every such offence by him or them committed, ten pounds»: quoted in Pollard (ed. 2004, p. 328).

18 By contrast, a juxtaposition of the First Folio with the seven-play corpus of Marlovian drama reveals eighteen god-terms among the top 373 most disparate words (the list stops at 373 since the log-likelihood figure has dropped by that point to 15.13, indicating that any further words would exhibit something lower than a statistical improbability in the 99.99th percentile). All eighteen of these terms appear with hugely greater frequency in Marlowe than in Shakespeare. In order of rank within the full list of 373, these terms are as follows: «jew», «christians», «lucifer», «soul», «jews», «christ», «pope», «friars», «hell», «heaven», «religion», «christian», «nun», «friar», «eternal», «damned», «divine», and «devils».

the evidence is qualified, since in Shakespearean plays «faith» and «worship» appear quite often in non-religious contexts, «faith» as an interjection and «worship» as a form of address. Doubt-terms, meanwhile, along with words commonly found in epistemological discussion, occur with massively greater frequency in Montaigne than in Shakespeare – words, that is, such as «reason», «natural», «knowledge», «opinion», «custom», «certain», «experience», «senses», «nature», «appearance», «belief», «memory», «ignorance», and «perceive». What we have, then, are a pair of corpora which exhibit striking similarities in their common absence of god-language (i.e., compared to the broad norms of 1590-1609), but striking dissimilarities in their deployment of the language of epistemological discourse. And since Montaigne's book is one that thousands of readers have felt is deeply sceptical about multiple forms of authority, knowledge, and custom, may we begin to decipher its lexical signature, so to speak, in such verbal constellations as those I have isolated through sequential juxtapositions against other contemporary corpora? As for Shakespeare, how should we interpret the comparative absence of epistemological vocabulary? Does it tell us anything significant about Shakespearean scepticism – or, indeed, about Shakespeare more broadly? My immediate reaction to this question is largely negative: no, we have far too little information at this point to make any such inference. Yet at the same time these statistics regarding lexical configuration obviously mean *something*. The relevant questions would seem to be these: what do such figures mean, how do we determine their meanings, are these meanings in fact important, and what new ways of thinking about the intersections of ideological history and computational analysis might be prompted by such forms of investigation?

Leaving Shakespeare and Montaigne behind for the moment, I would like to turn to a separate set of corpus juxtapositions. In this instance I have generated keyword frequency lists for sequential pairs of decades from the 1500s to the 1690s, again ranking words in terms of disparity based on log-likelihood calculations. Thus in Table 3 I present a series of nineteen specific juxtapositions: the 1500s against the 1510s, the 1510s against the 1520s, and so on. In each case I have examined the top 300 most disparate terms within relevant pairs of decades, isolating god-language in the process. Since the CQPweb spelling regularisation software does not work with this form of query, I have to some extent abandoned the constraints of my initial god-lexicon (i.e., the master list of terms provided in Tables 1 and 2), admitting both variant spellings and other relevant terms (e.g., «catholyque», «presbyters») not included in the main lists. We therefore see, for instance, that in the 1530s «cryste», «israel», «crysten», and «sauyour» all exhibit log-likelihood dominance, whereas in the 1540s we find that «christ», «israell», «christian», and «saviour» are highly predominant. But once we filter out orthographic anomalies of this sort, we are positioned to observe significant lexical disparities between juxtaposed decades. Perhaps most

prominent among these are the comparatively low levels of god-language in the 1520s, 1550s, 1570s, and 1590s. Another way of expressing this, of course, is that the surrounding decades – i.e., the 1510s, 1530s, 1540s, 1560s, 1580s, and 1600s – all exhibit much higher-than-average frequencies with respect to the broad god-lexicon. Either way, however, the results are perplexing. Why would it be the case, for instance, that among the 300 most disparate words between the 1590s and 1600s, 37 are god-terms (12.33%), and *all* of them predominate in the latter decade? Why does religious discourse exhibit such a sharp decline during the final years of Elizabeth's reign? Again, I have no ready answer to this question, but I suspect that the best way to begin addressing it is systematically to examine the 1046 printed texts from the 1590s against the 1349 printed texts from the 1600s.¹⁹ In other words, something akin to 'close reading' may be the best strategy for solving a particular conundrum brought to our attention by 'distant reading'.

Elsewhere in Table 3 it is mildly gratifying to witness the emergence of various familiar god-terms during successive decades: «ungodly», «heretykes», and «heresie» in the 1530s, «papists» in the 1550s, «righteousness» in the 1560s, «equivocation», «romanists», «popish», and «exorcistes» in the 1600s, and «presbyterian», «episcopacy», «church-government», and «synods» in the 1640s. A database including more than the top 300 most disparate terms would no doubt yield further satisfactions. But my suspicion is that until the CQPweb processor allows for more sophisticated searching – and especially for multi-variable constellations of lexical traits – we will not be able to move very far towards the isolation of statistically-convincing patterns of ideological change across sequential decades. At present I think that CQPweb is a tremendous resource for examining such matters as connotative variation over time: why, for instance, does «heathen(s)» exhibit declining levels of usage during the later decades of the seventeenth century, while «pagan(s)» and «infidel(s)» both show slight rates of increase? Similarly, I suspect that keyword juxtaposition analysis might be valuable in displaying detailed lexical differences between Protestant and Catholic styles of discourse, perhaps in poetic works such as Spenser's *Faerie Queene* (1596) or Southwell's *Saint Peter's Complaint* (1595), but still more likely in devotional polemics of the sort that Chloe Preedy has examined in her recent book on Marlowe (2012). With regard to epistemological and religious scepticism, however, I am at present unsure how corpus investigative techniques of the kind I have here employed can help in further sharpening our sense of the gradual erosion of god-language and related forms of religious discourse during the early modern era, particularly from the late sixteenth century forward.

¹⁹ These numbers represent the current totals available in EEBO Version 3 on CQPweb; they will rise when subsequent versions are installed.

I turn now to one final tactic of investigation which is somewhat more complex than the forms I have hitherto discussed: rates of lexical co-occurrence derived from proximity and collocation analysis. Using fourteen high-frequency terms drawn from my current god-lexicon, I have generated individual collocation tables for each term as it appears in ten consecutive decades from the 1550s through the 1640s. Each of these 140 tables, in turn, lists a minimum of 150 collocates, and some of them as many as 500. The collocates are ranked, once again, in terms of log-likelihood value, so that the most common (and thus the most statistically improbable) collocations appear at the top of each list. To give just one example, the highest-frequency collocate for «holy» in all ten decades is the word «ghost», which registers astronomical log-likelihood figures ranging from 13,629.64 in the 1550s to 113,827.29 in the 1640s.²⁰ To put this another way, «ghost» co-occurs with «holy» at a rate far exceeding that of any other word during each of the decades under consideration. The «holy»/«ghost» conjunction may thus be said to exhibit extraordinary collocational strength between 1550 and 1649. As for the parameters I have chosen for collocational analysis, they are simply the default settings currently provided within CQPweb: namely, a collocate must appear no further than three words to the left or right of a specific search term, and it must occur a minimum of five times in the relevant subcorpus to be considered valid for inclusion in the table.

Many other examples of consistent collocational strength might be offered. With «christ(s)», for instance, the words «jesus», «blood», and «saviour» are collocates which display tremendous prominence and consistency across the ten-decade span; with «soul(s)», such collocates include «body» and «salvation»; and so on with the following keywords: «scripture(s)» («holy», «canonical»); «christian(s)» («true», «religion»); «heaven(s)» («earth», «kingdom»); «nature(s)» («law»); and «god(s)» («almighty», «word»). Four of my fourteen god-terms are verbs («sin», «believe», «doubt», and «reason»), and in each case I have conducted my search using the following set of six conjugational forms: first-, second-, and third-person present tense (e.g., «sin», «sins», «sindest», «sindesth»), simple past tense («sinned»), and gerund («sinning»). Thus, in the case of «sin(s)», collocates of high prominence and consistency include «remission» and «original»; with «believe(s)», such collocates include the pronouns «I» and «we»; and with «doubt(s)», «I» and «no» top the list for consistency from 1550 to 1649.

More interesting than collocational stability, however, is collocational variation over time. And it is within this sphere that we can once again

²⁰ For «holy» as a keyword I have relied (as I have throughout this study) on the EEBO-TCP variant list, which in this case includes «holier», «holiest», and «holily» as well as the base term «holy».

begin to gather evidence for ideological change, although extreme caution must be used in the effort. Returning to the keyword «holy», for instance, we note that prominent collocates (besides «ghost» and «scriptures») include «martyrs», «communion», «sacrament», and «trinity». Only with the last of these terms, however, do we observe significantly increased rates of frequency during the ten-decade span under consideration. Similarly, we can note striking increases in «reformed» as a collocate for «church(es)» and in «instinct» as a collocate for «nature(s)». Still notable, but less impressive rates of increased collocational strength may be found with the following terms: «people» and «blessing» (as collocates for «god[s]»); «love» (as a collocate for «christ[s]»); «church» and «liberty» (as collocates for «christian[s]»); and «because» (as a collocate for the verb «reason»). The word «hell», meanwhile, displays a sharp rise in frequency as a collocate for «heaven(s)», which leads me to wonder whether the expression «heaven and hell» grew in popularity during the century from 1550 to 1649.²¹ «Truth» and «sinful», on the other hand, show only slight rates of increased collocational strength with respect to the keywords «doubt(s)» and «nature(s)». Finally, «experience» gains prominence as a collocate for the verb «reason» - but only in the 1630s and 1640s. Prior to that point the word fluctuates considerably in its levels of collocational strength, achieving significant gains in the 1590s but decreasing precipitously in the 1610s and 1620s.

Decreased rates of co-occurrence also appear with «christ» (as a collocate of «church[es]»), «kingdom» (as a collocate of «god[s]»), «against» (as a collocate of «reason(s)»), «passion» (as a collocate of «christ[s]»), «frail» and «against» (as collocates of «nature[s]»), and «deadly» (as a collocate of «sin[s]»). There is no question, in other words, that the phrase «the seven deadly sins» underwent marked decline (at least in its printed deployment) during the ten decades from 1550 to 1649. «Gospel», «christ», and «jesus», meanwhile, all exhibit slight decreases in collocational strength with respect to the verb «believe(s)». And in the 1610s and later, similarly slight decreases may be observed in the collocates «damned» (with respect to «soul[s]») and «religion» (with respect to «faith[s]»). Quite intriguingly, the collocates «up» and «down» both show gradual diminishment in collocational strength with regard to the keyword «heaven(s)» - a fact that may suggest a comparably slight diminishment in the degree to which «heaven» was imagined and described in distinctly spatial terms over the century under review. Finally, the collocate «vengeance» declines markedly in its rates of co-occurrence with the keyword

21 A quick 'string of words' search in CQPweb's Version 3 of EEBO suggests that this is indeed the case; i.e., the ipmw rate for «heaven and hell» is twice as high during the 1640s as it is during the 1550s. The rate fluctuates in the intervening decades, but is particularly low during the 1560s and 1570s, and particularly high during the 1600s, 1610s, and 1620s.

«god(s)». But lest we conclude from this fact that during the ten decades from 1550 to 1649 the notion of divine retribution – particularly within a Christian context – diminished significantly in its frequency of expression, we should note that the collocate «wrath» displays no comparable decrease in frequency with respect to «god(s)» as a search term. And this, I think, serves well as an illustration of why we must proceed with great caution in drawing inferences about potential ideological change from evidence of diachronic collocational variation such as that presented here. Synonyms and near-synonyms, for example, must be queried intensively so as to ensure that rates of variation point not merely to inevitable linguistic shifts over time, but to genuine alterations in structures of thought and belief.

In the end there can be no substitute for the close reading of individual texts. Distant reading through corpus analytic techniques can provide fascinating data-sets and reveal patterns of language use that might never be detected even through the trained intuitive discernment of the most accomplished and erudite scholars. Further forms of computer-assisted calculation, moreover, can sharpen the quality of the initial linguistic evidence and thus enrich and prolong the period during which an investigator might examine vast swaths of data and statistical assessment. One of my own goals in this regard would be to chart multiple vectors of diachronic collocational change by programming a computer not merely to assemble lists of collocates but to track log-likelihood variations over sequential decades and then to display results in visual form, presumably in graphs which juxtapose particularly intriguing examples. This, I believe, could bring us a step closer to discerning subtle patterns of ideological change through large-scale tactics of corpus linguistics analysis. Still, the questions raised by distant reading will always send us back to close reading – although with a more well-rounded perspective and with newly-conceived strategies of potential investigation. Indeed, we will rely upon close reading as long as we practice and value ‘close writing’: that is, the meticulous word-by-word composition of verbal artifacts which attend to the concerns of specific audiences and which display emotional sensitivity, reasoned argument, lexical innovation, irony, metaphor, humor, and all the other elements of advanced cognitive expression that may ultimately evade the interpretive reach of distant reading even in its unknown future forms.

Table 1. God-Words and Doubt-Words (EEBO-TCP Analysis [1473–1700]).

This is a list of the words I have currently selected for examination. The words are presented in lower-case spelling; all of my searches are non-case sensitive. The Word-Class column uses the following abbreviations: «v» for verb, «n» for noun, «pn» for proper noun, «int» for interjection, «adj» for adjective, and «adv» for adverb. The Variants column indicates the total number of spelling and verb-tense variants identified by EEBO-TCP.²² Dispersion has been calculated by dividing the number of records in a specific search by the total number of full-text records available in EEBO-TCP at the time I undertook that search (40,061). Density has been calculated by dividing the number of matches (or 'hits') by the number of records; this yields the average number of appearances of a given word in each text where it appears. I do not have a means to indicate the frequency of a word in terms of 'instances per million words' (as in CQPweb), since I do not know how many million words were included in the 40,061 fully-keyed texts available in EEBO-TCP at the time I compiled this table. But the dispersion and density figures should be roughly equivalent in this table and in that derived from CQPweb data (Table 2).

Word	Word-Class	Variants	Matches	Records	Dispersion	Density
faith(s)	n; int	9(5+4)	660,107	20,532	51.25%	32.15
religion(s)	n	20(13+7)	413,765	20,834	52.00%	19.86
church(es)	n	10(10+7)	1,388,972	23,868	59.57%	58.19
doctrine(s)	n	11(6+5)	329,604	15,249	38.06%	21.61
divinity(ies)	n	13(11+2)	50,532	9142	22.82%	5.52
divine(s)	n; v; adj; adv	42	317,190	19,141	47.77%	16.57
sacred(ly)	adj; adv	5(3+2)	92,228	14,294	35.68%	6.45
spiritual(ly)	adj; adv	27(20+7)	237,860	14,058	35.09%	16.91
righteous(ly)	adj; adv	27(21+6)	85,971	10,036	25.05%	8.56
almighty(ies)	n; adj	19(16+3)	67,373	13,034	32.53%	5.16
holy	adj; adv	25	676,527	20,597	51.41%	32.84
unholy	adj; adv	11	4606	2098	5.23%	2.19
pious(ly)	adj; adv	5(3+2)	38,125	8882	22.17%	4.29
impious(ly)	adj; adv	9(4+5)	17,928	5434	13.56%	3.29
piety(ies)	n	8(6+2)	55,765	10,188	25.43%	5.47
impiety(ies)	n	8(6+2)	22,385	6180	15.42%	3.62
idol(s)	n	16(7+9)	56,171	8131	20.29%	6.90
sabbath(s)	n	9(3+6)	58,912	5516	13.76%	10.68
god	n; pn	3	3,794,456	31,248	78.00%	121.43

22 Thus, for nouns, I have selected all spelling variants for singular and plural forms. For verbs I have chosen to include variants for the first, second, and third person in the present tense (e.g., «sin», «sins», «sinnest», «sinneth») along with the simple past tense («sinned») and the gerund form («sinning»). For adjectives I have chosen to include adverbial forms (e.g. «sceptically» as well as «sceptical»).

gods	n; pn	8	624,625	20,167	50.34%	30.97
god(s)	n; pn	11(3+8)	4,419,081	31,884	79.58%	138.59
christ(s)	n; pn	12(9+3)	1,725,023	18,578	46.37%	92.85
jesu(s)	n; pn	9(5+4)	375,362	14,758	36.83%	25.43
saviour(s)	n	18(12+6)	191,440	13,294	33.18%	14.40
holyghost(s)	n; pn	40(12+[9+7])	124,470	9822	24.51%	12.67
trinity(ies)	n	9(1+8)	33,723	5830	14.55%	5.78
christian(s)	n; pn	14(9+5)	502,912	19,528	48.74%	25.75
saint(s)	n	23(10+13)	367,967	17,533	43.76%	20.98
soul(s)	n	18(9+9)	736,640	22,346	55.77%	32.96
conscience(s)	n	19(10+9)	279,063	19,605	48.93%	14.23
prayer(s)	n	11(5+6)	321,547	18,329	45.75%	17.54
providence(s)	n	8(6+2)	90,826	13,204	32.95%	6.87
salvation(s)	n	27(23+4)	193,113	12,308	30.72%	15.69
scripture(s)	n	22(10+12)	462,939	15,399	38.43%	30.06
bible(s)	n	5(3+2)	30,020	7219	18.02%	4.15
gospel(s)	n; v	9(5+4)	292,121	14,274	35.63%	20.46
pastor(s)	n	7(2+5)	65,350	6676	16.66%	9.78
priest(s)	n	18(7+11)	283,209	15,118	37.73%	18.73
altar(s)	n	8(4+4)	77,573	8935	22.30%	8.68
sacrament(s)	n	10(4+6)	167,686	9080	22.66%	18.46
baptism(s)	n	10(8+2)	113,558	7226	18.03%	15.71
baptise(s)	v	17	28,812	4183	10.44%	6.88
bless(es)	v; n	27	340,524	22,519	56.21%	15.12
worship(s)	v; n	80	256,065	15,365	38.35%	16.66
rejoice(s)	v	47	92,547	12,837	32.04%	7.20
amen(s)	int; adv	2(1+1)	44,126	8973	22.39%	4.91
eternal(ly)	adj; adv	14(9+5)	195,710	15,108	37.71%	12.95
everlasting(ly)	adj; adv	22(13+9)	94,153	11,431	28.53%	8.23
heaven	n; pn	9	426,093	20,468	51.09%	20.81
heavens	n; pn	14	65,171	10,585	26.42%	6.15
heaven(s)	n; pn	23(9+14)	491,264	21,232	52.99%	23.13
angel(s)	n	18(7+11)	205,618	15,139	37.78%	13.58
hell(s)	n; pn	5	144,458	14,410	35.97%	10.02
purgatory(ies)	n; pn	8(6+2)	23,907	3722	9.29%	6.42
devil(s)	n	43(22+21)	244,890	16,196	40.42%	15.12

satan(s)	pn	4(2+2)	101,414	8946	22.33%	11.33
lucifer(s)	pn	2(2+0)	5637	2161	5.39%	2.60
antichrist(s)	n; pn	10(5+5)	48,815	5089	12.70%	9.59
sin(s)	v; n	34	1,044,021	20,350	50.79%	51.30
sinner(s)	n	6(2+4)	119,253	9659	24.10%	12.34
blaspheme(s)	v	25	17,299	5486	13.69%	3.15
despair(s)	v; n	51	44,647	9676	24.13%	4.61
damn(s)	v; int	15	42,614	8630	21.54%	4.93
blasphemy(ies)	n	17(11+6)	30,338	6997	17.46%	4.33
heresy(ies)	n	13(7+6)	73,785	8210	20.49%	8.98
pagan(s)	n; adj	6(3+3)	26,256	4851	12.10%	5.41
heathen(s)	n; adj	5(3+2)	75,130	10,090	25.18%	7.44
infidel(s)	n; adj	10(4+6)	24,007	5573	13.91%	4.30
jew(s)	n; pn	14(5+9)	231,156	13,088	32.67%	17.66
mahometan(s)	n; pn; adj	8(4+4)	9187	1920	4.79%	4.78
mahomet	pn	7	17,345	2492	6.22%	6.96
atheist(s)	n; adj	8(3+5)	15,538	4656	11.62%	3.33
atheism(s)	n	4(3+1)	10,050	3455	8.62%	2.90
sceptic(s)	n; adj	17(8+9)	1406	728	1.81%	1.93
scepticism(s)	n	8(4+4)	714	351	0.87%	2.03
sceptical(ly)	adj; adv	3	527	357	0.89%	1.47
pyrrhonian(s)	n; adj	10(5+5)	124	79	0.19%	1.56
doubt(s)	v; n	57	228,865	20,994	52.28%	10.92
doubtful(ly)	adj; adv	19(10+9)	30,275	8327	20.78%	3.63
appear(s)	v	77	441,016	25,826	64.46%	17.07
appearance(s)	n	18(11+7)	50,506	10,598	26.45%	4.76
perceive(s)	v	74	132,854	14,826	37.00%	8.96
perception(s)	n	6(4+2)	5080	820	2.04%	6.19
experience(s)	n; v	17	111,669	15,911	39.71%	7.01
experiment(s)	n; v	7	25,925	4455	11.12%	5.81
believe(s)	v	96	529,492	22,631	56.49%	23.39
belief(s)	n	17(12+5)	49,394	9113	22.74%	5.42
unbelief(s)	n	23(18+5)	19,371	3644	9.09%	5.31
knowledge(s)	n	29	272,262	19,032	47.50%	14.30
science(s)	n	7(4+3)	38,232	7201	17.97%	5.30
opinion(s)	n	21(12+9)	283,323	18,245	45.54%	15.52

dogma(s)	n	2(1+1)	877	466	1.16%	1.88
dogmatism(s)	n	2(2+0)	4	2	0.00%	2.00
dogmatic(al)	adj; adv	7	2181	960	2.39%	2.27
custom(s)	n	15(7+8)	146,319	15,302	38.19%	9.56
customary(ily)	adj; adv	5(3+2)	5403	2307	5.75%	2.34
perhaps	adv	8	97,969	12,620	31.50%	7.76
possible(ly)	adj; adv	10(3+7)	136,620	17,141	42.78%	7.97
reason(s)	v; n	34	773,554	25,879	64.59%	29.89
certain(ly)	adj; adv	45	492,805	23,051	57.53%	21.37
uncertain(ly)	adj; adv	44(25+19)	34,375	8856	21.10%	3.88
nature(s)	n	5(2+3)	613,447	22,957	57.30%	26.72
natural(ly)	adj; adv; n	15(8+7)	283,213	17,897	44.67%	15.82
unnatural(ly)	adj; adv	14(8+6)	16,086	6044	15.08%	2.66

Table 2. God-Words and Doubt-Words (CQPweb Analysis).

This is a basic CQPweb examination of the god-words and doubt-words listed in Table 1. Version 3 of EEBO on CQPweb currently contains a total of 44,422 fully-keyed texts, amounting to 1.2 billion words. Frequency is calculated according to instances per million words (i.e., within the corpus of 44,422 texts). Dispersion is the number of texts in which matches occur divided by the total number of fully-keyed texts. Density is the number of matches divided by the number of texts (this figure can be compared to the density figure derived from EEBO-TCP so as to gauge the relative compatibility of my results). My searches here incorporate spelling variants derived not from EEBO-TCP but from the VARD spelling regularization software developed by Dr. Alistair Baron at Lancaster University: (<http://ucrel.lancs.ac.uk/vard/>).

Word	Matches	Texts	Frequency	Dispersion	Density
faith(s)	765,161	23,322	636.46	52.50%	32.80
religion(s)	417,846	23,011	347.56	51.80%	18.15
church(es)	1,378,480	25,452	1146.62	57.29%	54.15
doctrine(s)	345,481	17,200	287.37	38.71%	20.08
divinity(ies)	50,037	9992	41.62	22.49%	5.00
divine(s)	327,483	21,156	272.40	47.62%	15.47
sacred(ly)	94,988	15,706	79.01	35.35%	6.04
spiritual(ly)	254,679	15,841	211.84	35.66%	16.07
righteous(ly)	91,411	11,199	76.04	25.21%	8.16
almighty(ies)	71,620	14,363	59.57	32.33%	4.98
holy	710,705	22,720	591.16	51.14%	31.28

unholy	3896	1787	3.24	4.02%	2.18
pious(ly)	39,058	9776	32.49	22.00%	3.99
impious(ly)	18,163	5809	15.11	13.07%	3.12
piety(ies)	58,917	11,365	49.01	25.58%	5.18
impiety(ies)	23,599	6822	19.63	15.35%	3.45
idol(s)	53,513	8777	44.51	19.75%	6.09
sabbath(s)	66,003	6447	54.90	14.51%	10.23
god	4,174,430	34,362	3472.28	77.35%	121.48
gods	640,962	21,958	533.15	49.43%	29.19
god(s)	4,815,392	35,069	4005.43	78.94%	137.31
christ(s)	2,017,605	21,116	1678.24	47.53%	95.54
jesu(s)	400,008	16,554	332.73	37.26%	24.16
saviour(s)	204,465	14,871	170.07	33.47%	13.74
holy ghost(s)	127,862	10,995	106.36	24.75%	11.62
trinity(ies)	33,493	6358	27.86	14.31%	5.26
christian(s)	523,211	21,808	435.21	49.09%	23.99
saint(s)	381,857	19,500	317.63	43.89%	19.58
soul(s)	798,232	24,854	663.97	55.94%	32.11
conscience(s)	295,099	21,826	245.46	49.13%	13.52
prayer(s)	345,667	20,415	287.53	45.95%	16.93
providence(s)	97,236	14,649	80.88	32.97%	6.63
salvation(s)	213,900	13,984	177.92	31.47%	15.29
scripture(s)	492,901	17,396	409.99	39.16%	28.33
bible(s)	31,333	7961	26.06	17.92%	3.93
gospel(s)	297,482	15,983	247.45	35.97%	18.61
pastor(s)	66,300	7343	55.15	16.53%	9.02
priest(s)	269,469	16,522	224.14	37.19%	16.30
altar(s)	74,599	9624	62.05	21.66%	7.75
sacrament(s)	164,356	9969	136.71	22.44%	16.48
baptism(s)	116,211	7887	96.66	17.75%	14.73
baptise(s)	80,305	7226	66.80	16.26%	11.11
bles(s)es	411,712	25,506	342.46	57.41%	16.14
worship(s)	263,488	17,014	219.17	38.30%	15.48
rejoice(s)	98,664	14,310	82.07	32.21%	6.89
amen(s)	47,515	10,079	39.52	22.68%	4.71
eternal(ly)	211,672	16,901	176.07	38.04%	12.52

everlasting(ly)	102,753	12,947	85.47	29.14%	7.93
heaven	503,504	23,384	418.81	52.64%	21.53
heavens	49,058	8480	40.81	19.08%	5.78
heaven(s)	552,562	23,970	459.62	53.95%	23.05
angel(s)	216,767	16,619	180.31	37.41%	13.04
hell(s)	152,342	15,762	126.72	35.48%	9.66
purgatory(ies)	24,703	4069	20.55	9.15%	6.07
devil(s)	249,749	17,907	207.74	40.31%	13.94
satan(s)	111,697	10,254	92.91	23.08%	10.89
lucifer(s)	5999	2422	4.99	5.45%	2.47
antichrist(s)	51,499	5553	42.84	12.50%	9.27
sin(s)	1,136,736	22,538	945.54	50.73%	50.43
sinner(s)	133,298	10,926	110.88	24.59%	12.20
blaspheme(s)	17,954	6039	14.93	13.59%	2.97
despair(s)	44,569	10,336	37.07	23.26%	4.31
damn(s)	44,855	9444	37.31	21.25%	4.74
blasphemy(ies)	31,169	7799	25.93	17.55%	3.99
heresy(ies)	71,587	8949	59.55	20.14%	7.99
pagan(s)	25,582	5197	21.28	11.69%	4.92
heathen(s)	80,020	11,248	66.56	25.32%	7.11
infidel(s)	23,609	6034	19.64	13.58%	3.91
jew(s)	221,198	13,882	183.99	31.25%	15.93
mahometan(s)	2954	963	2.46	2.16%	3.06
mahomet	3676	930	3.06	2.09%	3.95
atheist(s)	15,772	4960	13.12	11.16%	3.17
atheism(s)	10,676	3849	8.88	8.66%	2.77
sceptic(s)	542	328	0.45	0.73%	1.65
scepticism(s)	606	312	0.50	0.70%	1.94
sceptical(ly)	467	319	0.39	0.71%	1.46
pyrrhonian(s)	105	72	0.09	0.16%	1.45
doubt(s)	231,710	23,083	192.74	51.96%	10.03
doubtful(ly)	29,839	8905	24.82	20.04%	3.35
appear(s)	444,838	28,597	370.02	64.37%	15.55
appearance(s)	51,958	11,642	43.22	26.20%	4.46
perceive(s)	128,414	16,285	106.81	36.65%	7.88
perception(s)	5053	791	4.20	1.78%	6.38

experience(s)	119,185	17,821	99.14	40.11%	6.68
experiment(s)	26,747	4785	22.25	10.77%	5.58
believe(s)	540,076	24,924	449.23	56.10%	21.66
belief(s)	48,807	9638	40.60	21.69%	5.06
unbelief(s)	15,057	3128	12.52	7.04%	4.81
knowledge(s)	283,926	21,220	236.17	47.76%	13.38
science(s)	38,663	7700	32.16	17.33%	5.02
opinion(s)	283,917	20,113	236.16	45.27%	14.11
dogma(s)	1985	894	1.65	2.01%	2.22
dogmatism(s)	5	3	0.00	0.00%	1.66
dogmatic(al)	1602	790	1.33	1.77%	2.02
custom(s)	144,185	16,611	119.93	37.39%	8.68
customary(ily)	5369	2485	4.47	5.59%	2.16
perhaps	99,958	13,917	83.14	31.32%	7.18
possible(ly)	139,568	18,929	116.09	42.61%	7.37
reason(s)	803,264	28,695	668.15	64.59%	27.99
certain(ly)	489,537	25,675	407.20	57.79%	19.06
uncertain(ly)	35,312	9631	29.37	21.68%	3.66
nature(s)	648,449	25,680	539.38	57.80%	25.25
natural(ly)	296,675	19,925	246.77	44.85%	14.88
unnatural(ly)	16,508	6442	13.73	14.50%	2.56

Table 3. God-Word Juxtapositions in Sequential Decades (CQPweb EEOB Version 3).

This table offers preliminary evidence for the development of god-language over the sixteenth and seventeenth centuries in English printed texts. Using log-likelihood statistics for the top 300 most disparate words between pairs of adjacent decades, I provide ranked lists of the predominant god-terms within each juxtaposed pair. Spellings are not regularized (unlike those in Table 2), and thus some of the contrastive terms merely reflect changes in orthography.

Decade	Predominant God-Terms
1500s	god, sin, holy, confessor, hell, confession, heaven, crysten, synner, church, dampned, faith, cryste, penance, blyssed, lord, salvation, paradyse, devil, confess, ihesu
1510s	feythe, monastery, providence, heuyn, mynysters, monasterye
1510s	soul, ihesu, ghostly, charity, grace, holy, blessed, monastery, souls, smote, dyuyne, providence, prayer, devout, mynysters, prayers
1520s	christ, faith, christen, scripture, believe
1520s	pope
1530s	god, israel, jesus, christ, moses, soul, jerusalem, priests, sin, heaven, ungodly, holy, blessed, disciples
1530s	cryste, israel, saint, crysten, iesvs, heuyn, friar, church, purgatory, heretykes, iesv, sayntes, lord, spyryte, byleued, heresies, altar, catholyque, byshop, sauyour, soul, heithen
1540s	gods, christ, godly, spirit, israell, phariseis, scriptures, doctrine, baptism, heathen, salvation, fayeth, christians, christian, saviour, heavenly
1540s	lord, jesus, god, moses, ghospell, gospel, spirit, phariseis, israel, israell, disciples, faith, christ, host, jews, preestes, sin
1550s	sacrament, catholic, heuin, saluioir, papists, nature, synnis, mass
1550s	christen, saluioir, heuin, christin
1560s	pope, church, lord, bishop, bishops, luther, gods, israel, righteousness, popes, psalme, ecclesiastical, communion, bishoppes, priest, churches, sacrifice, canons, priests
1560s	isral, lord, sacrament, mass, luther, bishop, priest, bishoppes, church, israel, sacrifice, christ, chrystes, catholic, priests, pope, bisshops, holy, ghospell, blessed, ordanit, communion, synod, psalme
1570s	[no god-terms appear in this sample]
1570s	[no god-terms appear in this sample]
1580s	church, moses, jesuits, christis, israel, luther, god, soul, iesus, catholic
1580s	god, christ, church, holy, gospel, doctrine, jesus, bishop, faith, scriptures, apostles, heretikes, scripture, sacrament, epistle, gods, salvation, lord, beleue, catholic, godly, pope, luther, moses
1590s	[no god-terms appear in this sample]
1590s	[no god-terms appear in this sample]
1600s	pope, jesuits, popes, catholic, protestants, church, faith, religion, parsons, doctrine, scripture, gods, papists, christis, cardinal, christ, christian, priests, catholikes, religious, christians, luther, iesuit, holy, equivocation, sacrament, catholics, sin, baptism, mass, romanists, canon, bishop, popish, catholickes, presbyters, sacrifice
1600s	papists, protestants, jesuits, religion, exorcistes, popish, sacrament, scriptures, catholic, sheol

1610s	god, israel, lord, iesvs, sin, presbyters, moors, shrive, spiritual, apostle, bishops, diocesan, spirit, blessed, heaven
1610s	moyses, shrive, lord
1620s	iehovah, protestants, saint, arian
1620s	pope, israel, iehovah, popes, saint, priests, bishoppes, parsons, jesuits
1630s	sabbath, spirit, christ, heavens, catholiks, nature
1630s	saint, iesus
1640s	covenant, churches, god, ministers, lords, reformation, believers, presbyters, church, presbytery, religion, congregations, congregation, gospel, baptism, presbyterian, apostles, sectaries, episcopacy, church-government, israel, synods, anabaptists, saints, baptised, protestant, judas, jesus, worship, presbyteriall, excommunication
1640s	church, churches, religion, reformation, prelates, ecclesiastical, bishops, presbyteriall, sectaries, papists, god, synods, popish, psalme, conscience, brownists, saint, protestant, congregations, presbyters, presbytery, gods, popery, church-government
1650s	quakers, nature, sin, unbelief, faith, natural
1650s	christ, baptism, christs, baptised, gods, covenant, grace, believers, jews, sin, faith, apostle, gentiles, gospel, baptize, sacrament, baptizing, saints, psalme, scripture, godly
1660s	experiment, bishop, religion, popes, liturgy, ecclesiastical
1660s	god, gods, israel, godly, sins, saint, psalm
1670s	quaker, cardinal, quakers, divine, righteousness, jesuits
1670s	christ, sin, righteousness, quakers, god, grace, faith, scriptures, gods, quaker, christs, covenant, scripture, holiness, saints, knowledge, nature
1680s	church, bishops, protestant, popish, bishop, popery, sacrament, protestants, churches, religion, communion, papists, reformation
1680s	church, gods, sacrament, papists, popery, protestant, christs, communion, transubstantiation, worship, popish, pope, protestants, eucharist, saint, popes, catholic
1690s	righteousness, quakers, god, nature, covenant, trinity, synod

Works cited

- Eagleton, Terry (1991). *Ideology*. London: Verso.
- Greenblatt, Stephen (2001). *Hamlet in Purgatory*. Princeton: Princeton University Press.
- Hamlin, William M. (2005). *Tragedy and Scepticism in Shakespeare's England*. Basingstoke: Palgrave Macmillan.
- Hawkes, David (2013). *Ideology*. 2nd ed. London: Routledge.
- Hope, Jonathan; Witmore, Michael (2010). «The Hundredth Psalm to the Tune of 'Green Sleeves': Digital Approaches to Shakespeare's Language of Genre». *Shakespeare Quarterly*, 61 (3), pp. 357-390.
- Hope, Jonathan; Witmore, Michael (2014). «The Language of *Macbeth*». In: Thompson, Ann (ed.), *Macbeth: The State of the Play*. London: Bloomsbury, pp. 183-208.
- Larmore, Charles (1998). «Scepticism». In: Garber, D.; Ayers M. (eds.), *The Cambridge History of Seventeenth-Century Philosophy*. 2 vols. Cambridge: Cambridge University Press, pp. 1145-1192.
- Moretti, Franco (2005). *Graphs, Maps, Trees: Abstract Models for Literary History*. London: Verso
- Moretti, Franco (2013). *Distant Reading*. London: Verso.
- Pollard, Tanya (2004). *Shakespeare's Theater: A Sourcebook*. Malden, (MA): Blackwell Publ.
- Popkin, Richard H. (2003). *The History of Scepticism from Savonarola to Bayle*. New York: Oxford University Press.
- Pumfrey, Stephen; Rayson, Paul; Mariani, John (2012). «Experiments in 17th century English: manual versus automatic conceptual history». In: *Literary and Linguistic Computing*, 27 (4), pp. 395-408.
- Schmitt, Charles B. (1972). *Cicero Scepticus: A Study of the Influence of the *Academica* in the Renaissance*. The Hague: Martinus Nijhoff.
- Williams, Raymond (1977). *Marxism and Literature*. Oxford: Oxford University Press.
- Williams, Raymond (1983). *Keywords: A Vocabulary of Culture and Society*. Oxford: Oxford University Press.

