Spelling standardisation in Shakespeare’s first editions: evidence from the Second Quarto and First Folio versions of *Romeo and Juliet*

Rita QUEIROZ DE BARROS
University of Lisbon

ABSTRACT
This paper presents a study on spelling standardisation in Shakespeare’s first editions. Though certainly not central in literature, in which the orthography of Shakespeare’s texts has been considered mainly as an authorial and chronological test or as a tool for textual or phonological reconstruction, this issue deserves attention. An appraisal of the degree of spelling standardisation in Shakespeare’s first editions, which we know incomplete, may (i) contribute to a description of the standardisation of the English spelling system, generally allocated to the Early Modern period but still presenting important lacunae; (ii) provide a better knowledge of the spelling habits and variation patterns in Shakespeare’s first editions, thereby lessening the difficulties involved in the use of digital versions of those texts; (iii) supply a background against which to appraise the alleged manipulation of spelling for stylistic purposes in the Renaissance period, namely the use of visual rhymes and of spelling variants.

Assuming standardisation as a trend towards uniformity, this analysis concentrates on two different Renaissance editions of *Romeo and Juliet* and identifies a significant degree of orthographic regularity in the corpus considered, thus contradicting expectations raised by most references so far.

KEYWORDS: spelling, standardisation, Shakespeare, linguistic variation, Early Modern English

1. Introduction
Though not central in either historical linguistics or Shakespearean studies, which tend to assume orthography merely as a means to register speech, spelling and its standardisation are certainly not new to these research areas.
In fact, as evidence required for language reconstruction, past spelling practices have been treated in literature since the advent of English philology; furthermore, as Gómez Soliño stresses (1985a: 81), the rising of Standard English is a classic issue in the history of the language, and so far the approach to this question has privileged writing and, in particular, spelling (Rissanen 1999: 134, Wright 2000: 2).

Within Shakespearean studies, spelling has been considered not only for purposes of textual reconstruction (e.g.: Hinman 1963), as an authorial and chronological test (as suggested and mentioned by Pollard 1923), and as a tool for phonological reconstruction (e.g.: Cercignani 1981), but also, though less frequently, described for its own sake (e.g.: Partridge 1954, 1964; Blake 2002); and such allusions make it very clear that the spelling of Shakespeare's first editions was far from uniform. As is well known, even the First Folio, which is the product of a fairly careful enterprise by Heming and Condell, was found to be set in print by at least five compositors imposing on the text different spelling systems (Blake 2002: 7). Charlton Hinman has shown, for instance, that what he identifies as compositor A preferred the spellings doe, goe, here, griefe, traytor, young, Ulisses or Troian, while the so-called compositor B preferred the forms do, go, heere, greefe, traitor, yong, Ulysses and Troyan.

In spite of this background, spelling standardisation and its status in Shakespeare's first editions seem to deserve further investigation. In the first place, because descriptions of both the history of English spelling and of the standardisation of the language, of which orthographic regularisation is part and symptom, still present important lacunae.

In fact, despite diffuse treatment of orthography in traditional literature, references on the history of some graphemes (e.g.: Grosse 1937), various and important work on the spelling of particular words (e.g.: The Oxford English Dictionary; Hinman 1948), authors (e.g.: Byrne 1923, Partridge 1964, Samuels 1988 [1983], Diemer 1998), printers (e.g.: Fisher 1996 [1984], Aronoff 1989, Gómez Soliño 1985b, Salmon 1989), texts (e.g.: Partridge 1954; McLaughlin 1963, Blake 1965, Rutkowska 2000) and dialects (e.g.: Fisher 1996 [1977], 1979; McIntosh, Samuels and Benskin eds. 1986), and even three books especially dedicated to the history of English spelling (Vallins 1954, Scragg 1974, Bourcier 1978), approaches to the subject have been too circumscribed and/or flawed by the urge to justify the eccentric relationship between writing and speech in English (e.g.: Craigie...
1927, Vallins 1954), the instrumental value of spelling for phonological reconstruction (e.g.: Jespersen 1909) and the reigning lack of interest in writing since Saussurean times (e.g.: Scragg 1974). The need for a new history of English orthography has even been explicitly stated by Bliss (1975: 511) and, more recently, Görlach (1995: 5).¹

The same can be said of the standardisation of English. Despite strong renewed interest on the subject (cf. for instance Gómez Soliño 1984 and Wright ed. 2000), recent research has not yet been able to compensate for the traditional focus on the sources of Standard English, a still ongoing debate, and, in particular, for the apparent acceptance of the emergence of that variety as a linear process taking place in the Early Modern (henceforward EMod) period.²

As a consequence of the lacunae identified in previous paragraphs, the description of the standardisation of English orthography is still very incomplete. In fact, the exact characteristics of the system that turned standard are not known for sure and tend to be confused with the present ones; different opinions on those responsible for its emergence as a model – i.e. printers or linguistic authorities – still remain; the standard spelling’s diffusion along geographical, sociolinguistic and stylistic continua is largely to map; and last but not least, a contradictory chronology is attributed to that phenomenon. In fact, a search for this apparently simple piece of information in literature reveals that, though generally situated in Early Modern English (henceforward EModE), different references situate the standardisation of English orthography at different moments of that period; and some authors even give different dates for its occurrence in the very same text. That is the conclusion we can draw from the table below, which summarizes information collected from various sources:³

---

1 In fact, Scragg (1974), still the most comprehensive reference on the history of English spelling, concentrates on the external history of its subject, which has given rise to severe criticism in literature, namely by Bliss (1975) and Salmon (1976). For a thorough appraisal of the histories of English spelling see Kniezsa (1992).


3 Italics mark references presenting contradictory information on the chronology of spelling standardisation.
Table 1. Chronology of spelling standardisation

<table>
<thead>
<tr>
<th>15th century</th>
<th>16th century</th>
<th>17th century</th>
<th>18th century</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>early</td>
<td>middle</td>
<td>late</td>
</tr>
<tr>
<td>Orcutt (1929: 39)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bühler (1960: 5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pei (1967: 50)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fernández (1993: 87)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scragg (1974: 55)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brangelman (1980: 334)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Górlach (1991: 48)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowles (1997: 102)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salmon (1999: 32)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Potter (1950: 71)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partridge (1964: 2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scragg (1974: 80)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vallins (1954: 7, 16)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strang (1970: 107)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bourcier (1978: 129)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leith (1983: 34)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Freeborn (1992: 196)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As evidenced by the table, the time-span referred to is long. And though the inclusion of the fifteenth century is no longer a valid suggestion but a mere service to exhaustiveness, we are left with three centuries as possible moments for the standardisation of English spelling. This is surely a symptom of the need for further research, which is particularly true as far as descriptions of the spelling praxis are concerned, both in manuscript texts, for more obvious reasons, and in printed ones.4

That is probably why we can find recent research on or connected to the subject, namely by Gómez Soliño (in particular 1981, 1984, 1986), Sönmez (1993), Rodríguez (1999), Hérnandez-Campoy and Conde Silvestre (1999), Conde Silvestre, Hérnandez-Campoy and Pérez Salazar (2000) and Taavitsainen (2000). Gómez Soliño analyses the vertical diffusion of the emerging standard in printed and manuscript texts produced from 1470 to 1540; Sönmez assesses spelling standardisation in late seventeenth century as shown in manuscript and printed texts by the same author; Rodríguez observes the extension of written practices of the Chancery to private correspondence written in the late fifteenth century (part of the Paston Letters); Hérnandez-Campoy, Conde Silvestre and Pérez Salazar develop similar work on some of the Cely, Paston and Stonor Letters, dated from 1424 to 1490, and try to

---

4 In her recent chapter on EModE spelling and punctuation included in The Cambridge History of the English Language, Salmon sustains that “there is (…) no (…) detailed account of the gradual introduction of standard spelling in printed books” (1999: 55).
map the social diffusion of written variants typical of Chancery English; and finally Taavitsainen observes the extension of Chancery and Central Midland spelling practices to medical texts written from 1375 to 1550.

Despite their undeniable importance for the study of spelling standardisation, the references just mentioned are not enough. Most of them limit themselves to only a few variables; and all of them concentrate on either the beginning or the end of the EMod period and never consider its central decades, which may have played a crucial role in the standardisation of English spelling. This is one of the reasons to consider Shakespeare’s first editions in the particular perspective of spelling standardisation – those texts are certainly an important sample of the printed production of those times.

But lacunae in the description of the history of English orthography and standardisation are not the single motives to engage in a study on spelling regularisation. A second and no less central reason is the importance acquired by past spelling practices with the advent of electronic textual reproduction and analysis. In fact, access to original spelling editions, which are certainly to prefer, has become widespread; but so have automatic searches on such corpora, which are, almost fatally, based on graphic forms. Historical orthography has therefore ceased to be the interest of the specialist alone and turned into a tool required by all those who no longer dispense with electronic aids for their analysis of textual material: they have to be aware of the patterns and variation tendencies they can encounter. An appraisal of the degree of spelling standardisation in Shakespeare’s first editions will therefore lessen the difficulties involved in the use of the “more original” versions of such fundamental texts.

Last but not least, such a study may supply a background against which to assess the alleged manipulation of spelling for stylistic purposes in the Renaissance period. This is a tendency mentioned in some references, namely the use of visual rhymes (Wrenn 1943: 34ff) and the resource to spelling variants as a means

---

5 Gómez Soliño, Conde Silvestre et al. and Taavitsainen consider the graphic representation(s) of a limited set of words (ca. twenty), chosen because of their use in the LALME (Language Atlas of Late Medieval English); Rodríguez considers a similar sample of functional and lexical words, to which she adds some morphemes and grapheme clusters; the most detailed analysis is Sönmez’s, which contemplates bound morphemes, graphemes and some other graphical devices (i.e. apostrophe, hyphen, capitalization, macron, paragraph, blank space, abbreviation, etc.).
to enact the metamorphic style then in vogue (Adamson 1999: 555). But both claims seem at least problematic: the concept of visual rhymes appears to involve that of a fixed orthography, which is far from certain in EModE; and the deliberate resource to graphic variants for stylistic reasons implies the existence of standard spelling patterns, which are still to identify.

Since they were devoted to stating the motives to consider spelling standardisation in Shakespeare’s texts, previous paragraphs were not clear as to what is meant by standardisation. In this article, it is understood as the process resulting in the implementation of a linguistic standard, conceived as a “written variety varying minimally in form and maximally in function, whose norms are codified in grammars and dictionaries” (Kytö and Romaine 2000: 189).

In the particular plan of spelling, and since there were already prescriptive instruments in EModE – not only spelling books but also a major reference like Mulcaster’s Elementarie (1582) –, the study of spelling standardisation in Shakespeare’s texts seems to require the contemplation of the following aspects of their orthography:

(i) its degree of uniformity and/or variation;
(ii) its conformity to potential coeval models;
(iii) its similarity to the present spelling of English; and finally
(iv) the possible inclusion of future regionalisms that still permeate some sixteenth century texts.6

Considering all these issues at once would be too demanding. Points (i) and (iii) seem most urgent, as their results can be of service to the use of digital versions of Shakespeare’s first editions. So, and given the existence of spread information and a concise description of the most important differences between the present spelling and that of Shakespeare’s time by Blake (2002: 30-33), this article focuses on the degree of spelling uniformity and/or variation in those texts.

This aim will be approached via an electronically supported quantitative study. Its precise goal is to determine the relative weight of words with variant and invariant spelling in Shakespeare’s first editions. Though a simple way of assessing spelling standardisation, it is also an expectedly effective one (Görlach 1999: 4).

---

6 This tendency was empirically attested by Gómez Soliño in research mentioned above (e.g. 1981).
2. Corpus

The corpus considered in this study is a sample of Shakespeare’s first editions, since it would be impossible, and indeed unnecessary, to attend to all of them. It is composed of two Renaissance editions of Romeo and Juliet, namely (i) a copy of the Second Quarto, dated 1599, and (ii) a copy of the First Folio, published in 1623. Both these texts were collected in digital format from the site of the Internet Shakespeare Editions, prepared and maintained by the University of Victoria, Canada. They were transferred into word processor documents by means of a simple copy and paste procedure. Since the Internet Shakespeare Editions present each scene of each version separately, such sections were selected one by one in the original files, and then pasted into two documents, one per version. The final documents were saved as simple text files, the format required by the analytical software adopted, namely Mike Scott’s Wordsmith Tools.

The choice of this sample from the extensive list of Shakespeare’s writings was not random. It seemed advisable to avoid a poem, since spelling may be too constrained by stylistic factors in poetry; and within plays, Romeo and Juliet seems to present some advantages. In fact, it includes various styles – rhymed verse, blank verse and prose; it presents a medium length (990 lines); it is available, in the source mentioned, in various EMod editions, thereby providing evidence also on the chronological progress of spelling standardisation; and, finally, it was printed in both one (Second Quarto) and two (First Folio) columns, thus allowing for a control of the alleged increase of spelling variation in double column

---

7 This possibility is explicitly mentioned by Puttenham in *The Arte of English Poesie* (Bk. III, Ch. 1, apud Blake 1996: 230-232):

> A Word as he lieth in course of language is many wayes figured and thereby not a little altered in sound, which consequently alters the tune and harmonie of a meeter as to the eare. And this alteration is sometimes by adding sometimes by rabbating of a sillable or letter to or from a word either in the beginning, middle and ending ioynig or vnioynig of sillables and letters suppressing or confounding their severall sounedes, or by misplacing of a letter, or by cleare exchaunge of one letter for another, or by wrong ranging of the accent. [...] These many wayes may our maker alter his worde, and sometimes it is done for pleasure to give a better sound, sometimes vpon necessitie, and to make vp the rime. But our maker must take heed that he be not to bold specially in exchange of one letter for another, for vnlesse vsuall speach and custome allow it, it is a fault and no figure.
texts due to an also double need of line-justification (Pollard 1923: 5-6).

The two texts considered amount to a total of 50,270 words - 25,234 from the Second Quarto and 25,036 from the First Folio.

3. Methodology
This study intends to assess spelling standardisation in Renaissance editions of Shakespeare's works by determining the relative weight of graphic words with variant and invariant spelling in a sample of those texts.

But the choice of the graphic word as the basic unit for analysis is not self-evident: English is known to use a mixed spelling system, in which graphic units represent either phonemes, morphemes or words. For instance, the letter <a> regularly stands for the diphthong [ei] (e.g. lady) or the vowel [æ] (e.g. lad); but the sequence <ed> at the end of verbal forms represents the preterite morpheme, variously pronounced as [d] as in mowed, [t] as in packed or [id] as in started; and knight is a graphic form not obviously divisible into shorter units and that should be considered as a whole representing a lexical item. In Aronoff’s words (1989: 96),

A competent modern English speller must be aware of individual sounds, individual words (inasmuch as a majority of common words have peculiar spellings), individual morphemes (the spelling of most affixes is morphologically determined), classes of sounds (mostly in the use of silent e and consonant doubling to mark differences in vowel length and stress) and classes of words (especially those which are related by morphophonological rules).

This type of writing system was, according to the same author, already in existence by the time of Wynkyn de Worde, the famous second generation printer who died ca. 1534. A study of EModE spelling contemplating simply words is therefore incomplete from the start.

However, alternative choices would be unwise: considering phonemes would imply phonological reconstructions requiring decisions that would largely surpass the scope of this research; and taking into account morphemes would involve the segmentation of complex lexical items, which would be unnecessarily time-consuming. The graphic word seemed therefore to be the best
variable to attend to and so it is the object of the study described below.\textsuperscript{8}

The analysis of the two versions considered, which were always treated as autonomous texts, proceeded along the five following steps:

(i) Generation of a list of all the graphic forms in each text. This was produced automatically, namely by the WordList program included in Wordsmith Tools.

(ii) Edition of the word list mentioned above, from which had to be excluded several items initially identified by Wordsmith Tools. Such exclusions are composed of the following items:

(a) Forms appearing only once in the text and with no identified variant, since they were invalid evidence for a study on spelling uniformity – e.g. accident.

(b) Proper names, which were excluded because of their well-known spelling variability in EModE, as seems to have happened with Shakespeare's name itself – e.g. Romeo.

(c) Abbreviations. These correspond in most cases to characters' names, but include other sporadic elements – e.g. lul, for Juliet; coz, for cousin.

(d) Interjections, given their onomatopoeic roots – e.g. ah, o.

(e) A few foreign words used in the text – e.g. passado.

(f) Obvious misprints, i.e. graphic forms pointing to a pronunciation which we know for sure did not exist at that time – e.g. couragi.

(g) Items graphically represented as one word in the corpus but corresponding to separate Present English (henceforward PresE) words. The major reason for their exclusion is that it was impossible to control the use of multiword variants of the same element. Possible examples are yfaith and almaner.

(h) Incomplete words appearing in different lines but assembled by a hyphen in the original. The reason for this exclusion was the fact that Wordsmith Tools interpret them as separate words – e.g.: daugh- and ters from daugh-/ters.

(i) Problematic graphic forms known to represent different PresE words according to respectable references. The forms in

\textsuperscript{8} It should be stressed that the choice of the word as the basic unit for an analysis of English orthography is supported by past research – not only by random studies on historical orthography (e.g. Aronoff 1989, Goméz Soliño 1997, Diemer 1998, Rodríguez 1999, Conde Silvestre et al. 2000), but also by one of the most important descriptions of the PresE writing system, i.e. Carney (1994).
question are: I, because it could represent either the personal pronoun or yes; a, which represented both the indefinite article and the weak form of the personal pronoun he; and finally to and too and of and off, still undistinguished.

This long list of categories of excluded graphic forms obviously led to an important reduction of the material considered. The impact of such a reduction can be appreciated in the following table, which presents the number of words and graphic forms originally contained by each text and those that were kept for analysis.

<table>
<thead>
<tr>
<th>Edition</th>
<th>Words Original</th>
<th>Words Kept</th>
<th>Graphic Forms Original</th>
<th>Graphic Forms Kept</th>
</tr>
</thead>
<tbody>
<tr>
<td>Second Quarto</td>
<td>25,234</td>
<td>20,481</td>
<td>4,082</td>
<td>1,916</td>
</tr>
<tr>
<td>First Folio</td>
<td>25,036</td>
<td>19,222</td>
<td>3,973</td>
<td>1,845</td>
</tr>
</tbody>
</table>

Table 2. Number of words and graphic forms identified and kept for analysis

(iii) Preparation of a database, in Microsoft Excel, including the graphic forms from originals kept for research, their absolute frequency and their PresE equivalent. The PresE equivalent was added because it could be used as a tool to automatically assemble and count the different graphic forms of the same word by means of Excel Sort and Subtotals functions. Supplying such an equivalent was however not always an easy task, since some words were difficult to interpret. Whenever in doubt, the Oxford edition of the play was consulted.

(iv) Identification of the number of graphic variants per word with the help of Excel Sort and Subtotals functions.

(v) Retrieval of quantitative information from the database by means of Excel’s Automatic Filters.

4. Conclusions

Information resulting from the analytical steps described above is summarized in Table 3 below. It presents the distribution of words in the two editions of Romeo and Juliet considered in the study per number of spelling variants:
Table 3. Distribution of words per number of spelling variants

<table>
<thead>
<tr>
<th>Number of spelling variants</th>
<th>Second Quarto (1599)</th>
<th>First Folio (1623)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>fi⁹</td>
<td>fri (%)¹⁰</td>
</tr>
<tr>
<td></td>
<td>1237</td>
<td>79.1</td>
</tr>
<tr>
<td></td>
<td>306</td>
<td>19.6</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>Total</td>
<td>1564</td>
<td>100</td>
</tr>
</tbody>
</table>

These data allow for some comments and conclusions.
In the first place, results on the Quarto (1599) and Folio (1623) editions are extremely similar. Despite a lexical item with a total of 5 spelling variants only in the 1599 edition, numbers are quite alike in all cells. This implies that:
(i) The first quarter of the seventeenth century has played only a minor role in spelling standardisation.
(ii) Unlike suggestions made in literature, two column pages, as those of the First Folio, are not a necessary cause for spelling variation within the same text, despite their stronger demand for line-justification. This does not mean that text layout is without consequences for spelling - the two column First Folio makes a much more frequent use of the apostrophe than the Second Quarto version, printed in a single column; but the degree of spelling variation within the same text does not seem to be affected by that feature of text lay-out.¹¹

A second note to make is that the number of spelling variants identified for the same word is quite modest. It reaches 5 in the Second Quarto and 4 in the First Folio; but the number of words with such graphic variability is almost negligible. These data can therefore appease researchers who fear to ignore too many graphic

⁹ Absolute frequency.
¹⁰ Relative frequency.
¹¹ Statistics produced by Wordsmith Tools identify 377 apostrophes in the First Folio against only 34 in the Second Quarto.
representations of a lexical item when searching a digital and original spelling edition of Shakespeare’s plays.

Last but not least, the numbers presented above show that the percentage of lexical items with an invariant spelling in the corpus considered amount to ca. 80% - 79,1% in the Second Quarto and 83,2% in the First Folio. Even though a full appreciation of this number would require a comparison with parallel studies on previous and later texts of the same type, it is a significant conclusion. In fact, expectations raised in literature are quite different – to state just a few examples, Partridge considers the existence, in Renaissance times, of “perhaps, a sensus communis or common denominator of correct usage [i.e. writing] (…. ) but very small and obvious” (1954: 36); and Brengelman (1980: 345) sustains that at the end of the sixteenth century most English words admitted at least two graphic variants. The empirical data just discussed denounce, on the contrary, that spelling standardisation was already significant by the end of the sixteenth century.

This conclusion is not only an important advancement as far as the history of spelling standardisation is concerned, but also a reference for analysts of digital editions of Shakespeare’s texts and an empirical confirmation that English spelling already supported, at this moment of its history, the stylistic manipulation mentioned before.

The high percentage of words with invariant spelling is also a characteristic that Romeo and Juliet’s editions share with the almost contemporary Authorised Version of the Bible, which was considered in another study developed along similar lines (Queiroz de Barros 2003). It is therefore a proof that such degree of spelling standardisation in the early seventeenth century is not confined to a special text as the Bible and, in particular, its Authorised Version.

It should finally be stressed that the data collected allow and indeed require further investigation, namely on the interference of style over spelling variation and especially on the nature of the spelling patterns permeating the corpus considered.

References


Hinman, Charlton 1948. “‘Nether’ and ‘Neither’ in the Seventeenth Century.” Modern Language Notes: 333-335.


Internet Shakespeare Editions, University of Victoria, 2005 (http://ise.uvic.ca)


Author’s address:
Faculdade de Letras · Universidade de Lisboa · Departamento de Estudos Anglísticos · Alameda da Universidade · 1600-214 Lisboa
ritaqb@netcabo.pt