CHAPTER 1

Introduction

1.1 Chronological Overview

English spelling is in some ways a product of the Early Modern Era. The spelling forms that we use today are the result of a long process of conscious development and change, most of which occurred between the sixteenth and the seventeenth centuries. This portion of history is marked by a number of momentous events in England and the Continent, which had an immediate effect on English culture and language. For example, the end of the fifteenth century marks the beginning of the Tudor era (1485–1603), during which England became one of the most influential powers in Europe, and over the course of which the vernacular enjoyed an ever-growing development (Barber et al., 2009: 186). The year 1531 saw the secession of the religious relations with Rome by King Henry VIII, followed by the publication of multiple translations of the Bible in vernacular English and the outbreak of Protestantism in England (Smith, 1996: 169). The English Reformation marked the beginning of the fall of monasteries as centres of knowledge and development, and a gradual establishment of lay education in the English language (Smoluk, 2012: 119).

The beginning of the sixteenth century also marked the onset of the most fervent period of cultural, artistic, political and economic growth for England. The Renaissance (1300–1600) had an impact on various areas of social activity and collective knowledge in Early Modern England, including a revival of interest in the classical languages. As an immediate consequence of the Renaissance, the English vocabulary was enriched with a number of loanwords from Ancient Greek and Latin (Nevalainen, 1999: 332; Nevalainen & Tieken-Boon van Ostade, 2006: 281; Nurmi, 2012: 59). In addition, partly inspired by the etymologising movement in France, some of the spellings were reshaped according to their etymology, especially in words borrowed from Anglo-Norman or Old French, but also Latin and Ancient Greek, in order to make them resemble their original

root (Scragg, 1974: 52–4). At the time of the Renaissance, English became the object of harsh criticisms from early theorists of the language, who wanted English to 'imitate' the perfection and regularity of classical languages. One of the concerns that fuelled the criticisms advanced by the language theoreticians of the EModE era was the fact that writing was departing considerably from pronunciation. During that time, the Great Vowel Shift was causing a series of radical changes in the way in which words, and particularly vowels within words, were being pronounced (Lass, 1999b: 72–137; Barber et al., 2009: 201–7; see also Stenbrenden, 2016).

The fading of the Early Modern Era, on the other hand, corresponds to a smaller number of historical, political and social events in Europe. The English Civil War (1642–51) was a series of armed conflicts and political clashes between Parliamentarians (Roundheads) and Royalists (Cavaliers) over the manner of England's governance. The first (1642–6) and second (1648–9) conflicts pitted the supporters of King Charles I against the supporters of the Long Parliament, while the third conflict (1649–51) saw fighting between the supporters of King Charles II and the supporters of the Rump Parliament. That conflict ended with the Parliamentarian victory at the Battle of Worcester in 1651, and initiated a period of relative political peace (Görlach, 1999: 463). The changes brought about by the Civil War led to the formation of a bureaucratised and somewhat confessionalised national identity, and a formative political and regulatory power. While fully developed nationalisms had not yet come into existence in England, the late EModE society was certainly on its way to a process of 'nationalisation'. The concept of 'denomination', for example, was gradually replaced by the sense of 'nation' as early as the Tudor era, which essentially marked the birth of nationalism (Greenfield, 1992: 23).

As a consequence of these contextual changes, the Restoration of Charles II to the throne in 1660 marked the beginning of a new era for the English language (Salmon, 1999: 44). The political changes of the second half of the seventeenth century encouraged a renewed public consciousness of language as a bastion of authority. In this context, spelling became a useful nationalising tool, and orthographic debates became embedded in community discourses. In the 1660s, for example, the Royal Society, an organisation that co-ordinated the English scientific endeavours, proposed that the English prose scientists should strip their writing free from ornamentation and emotive tones, and that they should strive to write plainly, concisely and clearly. Events like these bespeak an

intense awareness of language in almost every area of politics, society and culture, which contributed to increasing the prestige of the vernacular and favoured changes on the level of phonology, lexicon and orthography (see Görlach, 1991: 42–59; Nevalainen, 2006: 4–6; Rutkowska, 2013b: 44). Over the course of the eighteenth century, the Enlightenment (1715–89) promoted a range of ideas centred on reason as the primary source of discernment and authority, and encouraged the advances of ideals like liberty, progress, tolerance, fraternity, constitutional government and detachment from the Church and the State. In the age of the Enlightenment, all kinds of manuals appeared for readers with social aspirations, including grammars and pronouncing dictionaries, which strengthened the process of linguistic standardisation, and which would eventually give the vernacular more structural rigour (Lass, 1999a: 8). The appearance of formal guides marked the beginnings of the so-called 'age of prescriptivism', which has been labelled 'the final stage in the standardisation process of the English language' (Tieken-Boon van Ostade, 2009: 3).

In this book, I focus on the development of spelling conventions in printed English within the chronological setting outlined here, beginning with the highest point of the English Renaissance and ending just before the beginning of the Enlightenment. The focus on two centuries of development affords a discussion of some of the key historical events that affected the wholesale standardisation of spelling in English. With regard to chronology, I focus on the interval of time between 1500 and 1700, reflecting something of a majority-based consensus among the scholarly community for defining the beginning and the end of the EModE period (Barber, 1976: 1; Görlach, 1991: 9; Blake, 1992: 18, 26, 28; Nevalainen, 2006: 1; Nurmi, 2012: 48). Needless to say, other proposals have been made over the years for alternative dates to identify, chronologically and linguistically, the Early Modern Era (see Görlach, 1991: 9; Blake, 1992: 14–15; Fisiak, 1994: 48; and Curzan, 2012 for overviews of different proposals). The chronological boundaries identified in this section, however, constitute a sufficient conventional agreement, and a point of departure, within the remits of this work.

1.2 Previous Research

1.2.1 Traditional Literature

Due to its crucial role in the history of English, scholars from different linguistic strands have explored a wide variety of areas in the study of EModE orthography, for example theoretical issues, levels of orthographic

change and standardisation, as well as the linguistic and extra-linguistic contexts that shaped and defined the writing system(s) of English. The diverse range of interests in the study of EModE orthography is not surprising, if one considers the outstanding differences in terms of research into the topic, and the number of linguists with divergent fundamental assumptions and methodologies who share an interest in the possibilities, tendencies and reasons for change in the EModE language. Among the most notable remarks, a number of writers have commented on the process of development as a gradual and somehow logical fixing of English spelling (Vallins, 1954: 79; Strang, 1970: 107; Pyles & Algeo, 1982: 68; Wakelin, 1988: 109; Millward, 1989: 225; Nevalainen & Tieken-Boon van Ostade, 2006: 289–91; Kohnen, 2014: 123–62). Various contributions to aspects of historical English have also weaved a narrative about EModE spelling as a systemic, well-defined entity. Some of the most relevant publications include those by Foster (1953), Dobson (1955), Hinman (1963), Whalley (1969), Barber (1976), McLeod (1979), Devitt (1989), Görlach (1991, 2001), Salmon (1999), Blake (2000), Smith (2005, 2012), Nevalainen (2006), Moessner (2012) and Rutkowska & Rössler (2012). All of these resources provide a compendium of information about developments and structures in EModE orthography and cover areas like spelling, punctuation and capitalisation.

Some of the more traditional scholars from the last century have largely explored author-specific orthographies, with a close focus on specific points within EModE, as well as various orthographic realisations of lexical items and morphological categories, for example inflectional endings and derivational suffixes. Relevant names in this respect include Bambas (1947), Salmon (1986, 1988 [1962], 1989), Osselton (1963, 1984, 1985), Shawcross (1963), Wilson (1963), Partridge (1964), Blake (1965), Graband (1965), Lucas (1973), Fisher (1977), Marckwardt (1977), Weinstock (1978), Brengelman (1980), Gómez-Soliño (1981, 1984, 1986), Little (1984), Johnston (1988), Aronoff (1989), Cram (1989), Sönmez (1993) and Liuzza (1996). Some contributions have also touched upon spellings within the context of EModE phonology, especially Kökeritz (1953), Dobson (1957) and Cercignani (1981), and, only partially, Scott (1967). EModE spelling is also a core area of discussion in general histories of English spelling, namely Vallins (1954), Venezky (1970), Scragg (1974), Bourcier (1978), Carney (1994), Upward & Davidson (2011), Crystal (2012) and Horobin (2013). These books deserve much praise for the difficult task of overviewing key steps and processes of orthographic standardisation in English, and for encompassing knowledge on the matter drawn from a relatively long

tradition of research work in English. The knowledge available from these and all the other titles mentioned so far provides present-day scholars with a set of guidelines and material for research into EModE orthography, but also a vision of spelling standardisation in EModE that is, generally, a gradual, linear and monolithic process from old and haphazard to modern and regular.

Over the past thirty years, however, researchers have shown increasing awareness of the importance of understanding the process of standardisation in EModE orthography as a complex, dynamic development, and have initiated profound changes in research and empirical attitudes to the topic. Recent approaches to investigating EModE orthography have been subject to a complex interplay between technological innovations, the use of new analytical methods, and the implementation of new theoretical frameworks (Condorelli, 2020a: 2). These innovations are especially owing to recent advances brought about by insights derived from historical sociolinguistics, which have initiated somewhat of a renaissance in English orthography. One of the most immediate products of the influence of historical sociolinguistics is a change in the focus to the correlation between orthographic practices and socio-contextual variables, which has in turn inspired researchers to branch out to a wide range of unexplored, diverse and complementary areas of investigation.

The first studies to investigate orthographic developments in English within a diachronic sociolinguistic framework appeared especially from the late 1990s and the early 2000s (e.g. Hernández-Campoy & Conde-Silvestre, 1999, 2005; Taavitsainen, 2000, 2004a; Conde-Silvestre & Hernández-Campoy, 2004; Rutkowska, 2005). The main focus of these studies was on the diffusion of early standard spelling practices in late fifteenth-century correspondence, the development of standard spelling practices, and the influence of authors' age, gender, style, social status and social networks on orthographic variation. Over the past few decades, book-length contributions with a relatively strong sociolinguistic stance were also published (e.g. Nielsen, 2005; Sebba, 2007; Hickey, 2010; Jaffe et al., 2012; Houston, 2013; Lillis, 2013). These titles have touched upon different aspects of spelling patterns and change, providing useful frameworks of analysis and ideas for new angles of research in English orthography. A more detailed overview of the new perspectives offered by recent research in EModE orthography, whether within the remit of historical sociolinguistics or beyond, is provided in the following section.

1.2.2 *Recent Research Trends*

In general, some of the more recent publications have investigated correlations between orthographic features and various combinations of extra-linguistic variables, including, for example, gender and text type (Sönmez, 2000; Oldireva-Gustafsson, 2002; Sairio, 2009b), gender, register and genre (Markus, 2006; Evans, 2012, 2013; Kaislaniemi et al., 2017; Evans & Tagg, 2020), as well as gender and authorship with a sociolinguistic tone (Evans, 2012; Hernández-Campoy, 2016). Other areas of investigation include text type and register (Gómez-Soliño, 1981, 1984, 1986; Osselton, 1984; Sönmez, 1993; Taavitsainen, 2000), register and level of formality (Tieken-Boon van Ostade, 2003, 2006b), occasional considerations pertaining to spacing, line justification, word division (Howard-Hill, 2006; Agata, 2011; McConchie, 2011; Rutkowska, 2013b; Shute, 2017b), and palaeographic factors (Wolfe, 2009). Some contributions have also covered a variety of both linguistic and extra-linguistic variables, with a somewhat stronger sociolinguistic framework (for references, see Rutkowska & Rössler, 2012). In other work, researchers have paid attention to the standardisation of EModE punctuation and capitalisation (Salmon, 1988 [1962], 1999; Cram, 1989; Rodríguez-Álvarez, 2010; del R. Medina Sánchez, 2015; Salles Bernal, 2016). Orthographic variation has also been explored from the perspective of discourse communities (Taavitsainen, 2004b), social networks (Tieken-Boon van Ostade, 1996, 1998; Sairio, 2008a, 2008b, 2009a, 2009b; Evans, 2015) and communities of practice (Sairio, 2013; Tyrkkö, 2013).

A few researchers have compared orthographic practice across several editions of the same book, mapping out the spelling systems of these texts and their diachronic developments (Sönmez, 1993; Horobin, 2001; Queiroz de Barros, 2007; Rutkowska, 2005, 2013a, 2013b, 2013c, 2015, 2016; Caon, 2010). Other researchers have focused on dialectal variants of EModE spelling, with reference to Scots (Devitt, 1989; Kniezsa, 1997), while yet other authors have concentrated on regional variation (Nevalainen & Raumolin-Brunberg, 1989; Wełna, 2011). Some researchers have explored the orthographic developments of lexical items and morphological categories, like inflectional endings and derivational suffixes (Berg & Aronoff, 2017). Others have discussed the competition between phonological and etymological principles in orthographies, resulting in different levels of phonography and morphography in particular languages (e.g. Sönmez, 2000; Conde-Silvestre & Hernández-Campoy, 2004; Taavitsainen, 2004b; Hernández-Campoy &

Conde-Silvestre, 2005; Sairio, 2009b; Wójcik, 2016). In yet other research work, specific diatopic and non-standard orthographic instances of variation and standardisation have been examined (Sönmez, 2000; Auer, 2005; Nevalainen, 2012; Rutkowska, 2015). Likewise, abbreviations and contractions have received some attention for at least a couple of decades (Shevlin, 1999; Tieken-Boon van Ostade, 2003, 2006b; Markus, 2006; Rutkowska, 2013b). Most recently, the dynamic and continuous production of transcribed corpora have encouraged a number of researchers to explore technological approaches to investigating EModE spelling on a larger-scale basis (Schneider, 2002; Baron et al., 2009a, 2011a, 2011b; Craig & Whipp, 2010; Archer et al., 2015; Basu, 2016; Shute, 2017a; Condorelli, 2020b). The study of English orthography using computational approaches has opened up the door to a whole new range of possibilities, and has made it possible to address questions that were difficult to answer in the past. In this book, I aim to provide a contribution to the widely diverse, ever-growing dialogue in the field, introducing a new quantitative model for analysing diachronic spelling developments, and investigating questions about standardisation in a new way.

My work focuses, specifically, on spelling, rather than orthography more broadly. Spelling involves the graphical realisations of all spoken items, whereas orthography is limited to a more or less binding norm (Rutkowska & Rössler, 2012: 214). In general, the term *orthography* is used as a hyponym of *spelling* and refers to one or more writing systems that comprise not only the spelling of particular lexical items and morphemes, but also capitalisation, word division and punctuation, as well as functional and decorative marks. Orthography depends on the practices of a community of writers at a given moment in time, and has to be established and accepted by the community. Recent explorations of spelling developments in EModE (e.g. Rutkowska, 2013b; Basu, 2016) have shown, with a number of examples, that spelling forms the basic unit for the process of standardisation in English and represents the most important element of analysis for core developments in a writing system. Among the spelling units, graphemes in particular behave in a rather interesting way: their development seems to occur autonomously and within their own well-defined time frame. The grapheme is intended as 'the minimal functional distinctive unit of any writing system' (Henderson, 1984: 15), 'a purely distinctive visual unit, part of an autonomous semiotic system' (Liuzza, 1996: 28), and, whether alone or in combination with other graphemes, it represents the lowest denominator of spelling.

The following are some of the most notable examples of variant developments (mostly alternations) at the level of individual graphemes or groups of graphemes that have been discussed within the remits of EModE spelling: the alternation between <u> and <v>;[1] the alternation between <i> and <y> (sometimes also <e>);[2] the regularisation of word-final <e>;[3] the alternation between <i> and <j> (and <g>);[4] the standardisation of word-final <ie> to <y> (sometimes also alternating with <ye>);[5] the introduction of <ea>, <oe> and <oa>;[6] the variation between <ll> and <l> (e.g. <all> with <al>, <ell> with <el>), and <el> and <le>;[7] the standardisation of <(e)th> to <(e)s>;[8] the replacement of <þ> with <th> or <y>;[9] the alternation between <s> and <z>, <ss> and <zz>;[10] the alternation between <o(r)> and <ou(r)>, <re> and <er>;[11] the alternation between <ck> and <k>, <ick> and <ic>;[12] the replacement of <ch> with <tch>;[13] consonant doubling after short vowels (and the replacement of <tt> with <t>);[14] the alternation between

[1] Classen (1919), Vallins (1954), Strang (1970), Scragg (1974), Brengelman (1980), Salmon (1986, 1999), Wakelin (1988), Millward (1989), Görlach (1991), Freeborn (1992), Barber (1997), Burchfield (2002), Smith (2005), van Gelderen (2006), Nevalainen (2006), Craig & Whipp (2010), Lehto et al. (2010), Weiner (2012), Horobin (2013), Basu (2016), Rutkowska (2016).

[2] Classen (1919), Vallins (1954), Blake (1965), Scragg (1974), Brengelman (1980), Wakelin (1988), Aronoff (1989), Freeborn (1992), Barber (1997), Salmon (1999), Culpeper & Archer (2009), Craig & Whipp (2010), Lehto et al. (2010), Nevalainen (2012), Weiner (2012), Evans (2013), Horobin (2013), Basu (2016), Rutkowska (2016), Evans & Tagg (2020).

[3] Classen (1919), Vallins (1954), Brengelman (1980), Wakelin (1988), Aronoff (1989), Görlach (1991), Freeborn (1992), Sönmez (1993), Salmon (1999), van Gelderen (2006), Nevalainen (2006, 2012), Caon (2010), Craig & Whipp (2010), Lehto et al. (2010), Evans (2013), Horobin (2013), Rutkowska (2013b).

[4] Classen (1919), Vallins (1954), Scragg (1974), Brengelman (1980), Millward (1989), Görlach (1991), Freeborn (1992), Barber (1997), Salmon (1999), Smith (2005), Craig & Whipp (2010), Lehto et al. (2010), Weiner (2012), Horobin (2013), Basu (2016), Rutkowska (2016).

[5] Classen (1919), Vallins (1954), Strang (1970), Scragg (1974), Brengelman (1980), Osselton (1984), Görlach (1991), Salmon (1999), Craig & Whipp (2010), Lehto et al. (2010), Horobin (2013).

[6] Blake (1965), Scragg (1974), Millward (1989), Görlach (1991), Salmon (1999), Howard-Hill (2006), Horobin (2013), Rutkowska (2013b), Tyrkkö (2013), Basu (2016).

[7] Vallins (1954), Osselton (1984), Sönmez (1993), Salmon (1999), Rutkowska (2005, 2013b), van Gelderen (2006), Howard-Hill (2006), Sairio (2009b), Kaislaniemi et al. (2017).

[8] Bambas (1947), Strang (1970), Freeborn (1992), Barber (1997), Blake (2002), Barber et al. (2009), Raumolin-Brunberg (2011), Horobin (2013).

[9] Classen (1919), Strang (1970), Millward (1989), Freeborn (1992), Sönmez (1993), Burchfield (2002), Weiner (2012), Tyrkkö (2013).

[10] Vallins (1954), Wakelin (1988), Görlach (1991), Salmon (1999), Nevalainen (2006), Evans (2012, 2013), Horobin (2013).

[11] Vallins (1954), Scragg (1974), Osselton (1984), Sönmez (1993), Salmon (1999), Culpeper & Archer (2009), Tyrkkö (2013).

[12] Vallins (1954), Brengelman (1980), Osselton (1984), Sönmez (1993), Burchfield (2002), Nevalainen (2006), Horobin (2013).

[13] Blake (1965), Brengelman (1980), Osselton (1984), Sönmez (1993), Salmon (1999), Rutkowska (2013b).

[14] Vallins (1954), Scragg (1974), Salmon (1999), Howard-Hill (2006), Culpeper & Archer (2009), Rutkowska (2013b).

word-final <'d> and <ed>, <t> and <'t>;[15] the alternation between <a(n)> and <au(n)>;[16] the restoration of etymological <b>, <c>, <d>, <l> and <p>;[17] the replacement of <owh> with <ough> and the alternation between <(o)u> and <(o)w>;[18] the alternation between <sch> and <(s)sh> (and other combinations);[19] the replacement of <f> with <ph>;[20] the alternation between <c> and <t>, especially in *-tion;*[21] the alternation between <vv> and <w>;[22] the rationalisation of <er> and <ar>;[23] the variation between word-initial <g>, <gh> and <gu>;[24] the alternation between <g(g)> and <dg>, <dg(e)> and <g(e)>;[25] the alternation between <d> and <th>;[26] vowel doubling (ee, e . . . e and oo, o . . . o);[27] the alternation between word-final <f> and <ff>;[28] the alternation between <ei> and <ie>;[29] the alternation between <wh> and <w>;[30] the alternation between <gh> and <ght>;[31] the alternation between <c> and <s>;[32] the introduction of <æ> and <œ> for <e>;[33] and the replacement of <ȝ> with <gh>, <y> or <s>.[34] This book focuses on a few of the examples mentioned, including both individual graphemes and groups of individual graphemes, and traces their history of standardisation in EModE as a complex, dynamic development. The following section discusses in greater detail the scope and the outline of the book, starting with some methodological and theoretical remarks that contextualise the goals pursued.

1.3 Scope, Rationale and Structure of the Book

1.3.1 Scope

In recent years, statistical methods, visualisation techniques and corpus-driven approaches have enriched the scope of English spelling, allowing researchers to query large datasets and capture and analyse big data, sidestepping traditional approaches based on qualitative perspectives,

[15] Osselton (1984), Millward (1989), Barber (1997), Culpeper & Archer (2009).
[16] Vallins (1954), Blake (1965), Salmon (1999), Evans (2013).
[17] Wakelin (1988), van Gelderen (2006), Baugh & Cable (2013), Rutkowska (2013b).
[18] Blake (1965), Scragg (1974), Sönmez (1993).
[19] Strang (1970), Evans (2012, 2013), Rutkowska (2013b).
[20] Osselton (1984), Horobin (2013), Rutkowska (2013b).
[21] Vallins (1954), Weiner (2012), Rutkowska (2013b).
[22] Görlach (1991), Smith (2005), Craig & Whipp (2010).
[23] Blake (1965), Wakelin (1988), Rutkowska (2013b). [24] Scragg (1974), Salmon (1999).
[25] Scragg (1974), Salmon (1999). [26] Sönmez (1993), Rutkowska (2013b).
[27] Salmon (1999), Weiner (2012). [28] Burchfield (2002). [29] Kaislaniemi et al. (2017).
[30] Evans (2013). [31] Evans (2013). [32] Sönmez (1993). [33] Horobin (2013).
[34] Millward (1989).

and endeavour to identify previously undetected patterns. A common denominator to all computational methods is that they often encourage a theory-neutral or data-driven perspective on primary data and emphasise the need to find tendencies, patterns and trends, which may (or may not) show something new, or confirm an already advanced hypothesis about a complex dataset. While it is arguable whether quantitative methods alone can make much of a contribution to our knowledge of language, texts or cultures, they certainly offer intriguing and sometimes unexpected insights into large-scale patterns that we would normally not be able to see with our naked eye (Tyrkkö, 2017: 100). In this book, quantitative insights are often integrated with a careful, qualitative take, in order to investigate question-driven issues, while also gaining data-driven perspectives about the mechanisms of EModE spelling standardisation.

In previous work, it has been considered 'commonplace' to attribute the gradual standardisation of English spelling in the EModE period to the impact of printing technology (Tyrkkö, 2013: 151). However, the relative influence of a number of agents on the wholesale standardisation of English spelling remains, to date, anything but a settled topic (see Scragg, 1974: 52–87; Salmon, 1999: 32; Nevalainen & Tieken-Boon van Ostade, 2006: 290; Percy, 2012: 1008). Among the previous commentators on the issue, Brengelman (1980: 343) sees the theoreticians' contribution as a way to promote patterns of consistency which 'were actually followed (if tardily) by authors and printers'. For Howard-Hill (2006: 16), on the other hand, Early Modern printers were largely not affected by authors or theoreticians, and used the authors' original spelling as a starting point to introduce their own in-house set of rules. Even among present-day researchers, the issue of which factors were important for standardising English spelling as we know it today remains unresolved. Nevalainen (2012: 156), for example, suggests that more parties were involved and worked together as agents in the process of fixing of English spelling, including theoreticians, schoolmasters, authors, readers and printers. For Basu (2016), modern spelling features arose mainly because of the influence of printers. For Rutkowska (2016: 187), the theoreticians/printers debate is likely to remain an unresolved issue for a long time. For Berg & Aronoff (2017), the standardisation of English spelling occurred predominantly as a relatively spontaneous, self-organising internal process. Clearly, more work is needed in order to increase our understanding of the factors responsible for spelling standardisation in EModE. In my opinion, the statements made by previous scholars on the matter are conditioned