Alpo Honkapohja and Aino Liira

9 Abbreviations and standardisation in the *Polychronicon*: Latin to English and manuscript to print

1 Introduction

Abbreviations were an integral part of the writing systems used in the Middle Ages. They were used both to conserve precious writing materials and to alleviate "the labour of writing Latin" (Hector 1958: 37). Proof of how widespread and so-phisticated the Latin system had become is that the most comprehensive reference work for medieval Latin abbreviations by Adriano Cappelli (1990 [1899]) contains some 14 000 abbreviations. When vernacular languages like English and Anglo-Norman French began to be written down, the system of abbreviations. The system was especially important in a multilingual society, as abbreviations can be language-independent. Towards the end of the Middle Ages the number of abbreviations began to decrease, simultaneously with technological innovations in book production and the emergence of English in a new nationwide function.

The gradual disappearance of the abbreviation and suspension system is linked to the technological developments in book production. As parchment began to be replaced by a cheaper material, paper, and the printing press made it possible to produce multiple copies with ease, the two main needs for using an abbreviation and suspension system lost their importance. The system was eventually abandoned in printed books, although it continued in handwriting used for personal letters and notes, and legal writing (Hector 1958: 28, 38; Kytö, Grund and Walker 2011). Furthermore, a decrease in the use of abbreviations took place at the same time as vernacular English was gaining ground from Anglo-Norman French and even Latin, and a new written standard for English was beginning to emerge. The details of this gradual change, however, remain largely uncharted.

Our aim in this chapter is to help build a foundation for the timeline and reasons of the loss of the abbreviation system. By quantitatively studying changes in the abbreviations and variation across copies of a single work, Ranulph Higden's

Note: We would like to thank Jane Roberts and two anonymous reviewers for their comments. Their feedback led to a much improved article. We are also thankful to Sara Norja for proofreading and Laura Wright for editing the article and making a number of helpful suggestions. All remaining mistakes are our own.

Polychronicon, we hope to provide a descriptive outline for the reduction in their numbers and the changes in the system concurrent with changes in written language. We ask what happened to the system of abbreviation with the rising vernacular written standard on the one hand, and with the new technologies of paper and printing on the other. Our analysis is divided into four sections, the first of which focuses on the proportion of abbreviated words in Latin and English in the data. The second section establishes a picture of the abbreviation types that disappeared first and the kinds of words in which abbreviations survived the longest. In the third section, we examine the effects of technological aspects of book production, such as right-margin justification in double column layouts, on the abbreviation system. Finally, the fourth section compares the reduction of abbreviations and the reduction of spelling variants in general.

Our findings show that the rate of the disappearance was different in Latin and English, and that different abbreviation types disappeared at different rates. While there was a major reduction in the use of abbreviations in large *de luxe* manuscripts like the *Polychronicon*, our data show only minor reduction in spelling variants. Moreover, the density of some abbreviation types actually increases in printed books due to the emergence of a standard set of abbreviations used by early modern printers. Abbreviations were used for line justification, among other purposes, and they seem to have survived the longest in this function. The results thus show that the medieval abbreviation and suspension system underwent both qualitative and quantitative changes. These changes happened early in the standardisation process, but many of the usages survived up to early printed books and later.

2 The emergence of a written standard and the loss of abbreviations: Previous work

The loss of the medieval abbreviation and suspension system in English has, by and large, not been described from the point of view of standardisation and with the quantitative precision of corpus linguistics. Diachronic developments are mainly treated by concise and imprecise statements in palaeographical handbooks. For example, Petti (1977: 22) notes that "[t]he general pattern in English literary manuscripts was one of gradual reduction, so that by the Renaissance, abbreviations were of modest proportions and, in any case, more abundant in drafts than in formal copies, where the practice was hardly more extensive than in printed books". What we do not know is how exactly this change proceeded. This is, however, not to say that there has not been any quantitative work at all. We now have some corpus resources which encode suborthographic phenomena, including abbreviations, in a way that can be studied quantitatively.¹ A handful of studies² have made use of them, proving that studying them can yield interesting results. Two corpus studies overlap with our period. Shute (2017a, 2017b) touches upon abbreviations as a part of her quantitative study of spelling variation in Caxton, noting that they are statistically more likely to occur close to the right margin. Smith (2019) gives a diachronic account of one common abbreviation in the *Linguistic Atlas of Older Scots* (LAOS). Moreover, there have been some interesting discoveries in French scholarship (see, e.g., Hasenohr 2002; Camps 2016; Stutzmann et al. 2018). All of these will be discussed in sections 4.1, 4.2 and 4.3 below. However, with respect to English, an overall quantitative account of the gradual disappearance of many abbreviations between ca. 1350 and 1500 is lacking. We do not know how it lines up with the emergence of the English vernacular as the main language of written communication.

The development of a written standard for English is a process whose broad outlines are known, but the causes and exact mechanisms are incompletely understood and have recently been opened up for new discussion. The broad outlines are that post-conquest England, between 1066 and c. 1350, was a trilingual society in which all educated language users knew Latin and Anglo-Norman French (cf., e.g., Ingham 2012). English may have been the primary spoken

¹ The few corpora that encode abbreviations in a way that enables applying quantitative methods include texts made available by the Medieval Nordic Text Archive (MENOTA), a network of libraries, archives and research departments of Old Icelandic, Old Norwegian and Old Swedish Texts. In English studies, corpora that encode abbreviations include the Edinburgh resources, LAEME (see Laing 2013) and LAOS, as well as the Middle English Grammar Corpus (MEG-C) and Middle English Local Documents (MELD) corpora compiled at the University of Stavanger, and the digital edition *An Electronic Text Edition of Depositions 1560-1760* (Kytö, Grund & Walker 2011). Recently Stutzmann et al. (2018), as a part of the HIMANIS network, have applied Handwritten Text Recognition to a large corpus of French administrative texts from the 14th and 15th centuries; the resulting corpus allows retrieval of both abbreviated and unabbreviated forms through a plain text search. None of these resources has been used exhaustively for the study of abbreviations.

² Wright has studied abbreviations in English/Anglo-Norman/Latin mixed-language documents, with small datasets that a single scholar can handle manually (see, e.g., Wright 2000, 2011 and 2013). Two recent Dutch PhD dissertations (ter Horst 2017 and Stam 2017) have investigated similar multilingual phenomena using corpora, focusing on Latin and Irish. Kestemont (2015) studied scribal profiles, including abbreviations, using a stylometric approach in the letter collection of the Middle Dutch mystical female poet Hadewijch. Other discoveries were made by Rogos (2012), who focused on late Middle English literary manuscripts, noting that word-final characters alternate with graphic sequences rather than substitute them.

language for most of the population, but written English reflected the local dialect and was frequently mixed with Latin and Anglo-Norman in documents (cf., e.g., Wright 2002). The change began in the fifteenth century when English was increasingly used as the language of writing. This led to a loss of much of its variation and a new written standard eventually emerged.

There is, however, no single accepted explanation for what led to the emergence of a new written vernacular standard. A very influential account was written by Samuels (1983 [1963]), who proposed that the development towards standardization for written English can be divided into four Types. The last of these is the so-called "Chancery" standard, according to which a department of Royal administration provided the model for the written standard of English. This idea was developed further by Fisher (1977, 1979, 1996), but his strong claims have been decisively dismantled by Benskin (2004) and Wright (2000), who have shown Fisher's work to be lacking theoretically and selective in its use of data (see also chapters 1 and 2 in this book). Neither Samuels (1983 [1963]) nor Fisher take abbreviations into account, but interestingly, Samuels, in a later article (1983), discusses the frequencies of a few common abbreviation types in two manuscripts of the *Canterbury Tales*, mentioning diachronic developments in reference to his types.³ His findings and views will be discussed in more detail in sections 4.1 and 4.2 below.

While the theory of a Chancery standard has been very influential, and is often still the view found in many textbooks, recent work on standardisation has moved in a more sociolinguistic direction (see, e.g., Deumert and Vandenbussche 2003; Nevalainen and Tieken-Boon van Ostade 2006). Since the present study seeks to situate the loss of abbreviations, on the one hand, in the external context of book production, and on the other hand in linguistic change, it too can be described as sociolinguistic. The linguistic framework is based mainly on ideas proposed by Wright (2013), who calls attention to simplification caused by dialect contact (2013: 71) as a potential explanation for the loss of variety, and examining variation from the point of view of a complex system (2013: 64–66).

The view of variation promoted by Wright is informed in particular by the "linguistics of speech" approach proposed by Kretzschmar (2009, 2015; Kretzschmar & Stenroos 2012). According to this approach, language, when examined through big enough data, behaves like a complex system. A complex system refers to the kind of systems that display self-organised patterns, such as those described for contemporary biology and economics (Kretzschmar and Stenroos 2012: 112). Variation

³ The article is written as a response to critique by Ramsey (1982), who claims that the Hengwrt and Ellesmere manuscripts of the *Canterbury Tales* were copied by different scribes.

in historical linguistic data is a "result of the interplay of historical and cultural forces to which language is always subject", and this interplay is characterised by "the massive interaction of speakers and writers over time, as a complex system from which the regularities of our language emerge" (Kretzschmar and Stenroos 2012: 112). Thus written Middle English can be expected to show variation characteristic of a complex system and standardisation can be seen as gradual reduction of this variation.

From the point of view of the linguistics of speech, the development of a written standard forces unnatural uniformity on the natural variability of language. According to Wright, "[t]here is nothing 'natural' about this process of reduction" (2013: 65). Naturally occurring language data, especially speech data, will show a characteristic complex distribution in which a few variants are very common, but there will be a long tail of many uncommon variants (see, e.g., Kretschmar and Stenroos 2012). When languages change, the relative frequency of variants changes in proportion to one another (Kretzschmar and Stenroos 2012). The process that most characterises the fifteenth century is not selection, but rather a whittling down of variants, or *elimination*. "The actual whittling-down process to one supreme variant used by everybody happened well after 1500, and thus after the period of 'Chancery Standard'" (Wright 2013: 65–66). The type of English written by London Bridge clerks in 1501–1502, which has a reduced number of spelling variants per scribe, but still a large pool of variants and different dominant forms compared to ones that eventually became selected, Wright calls protostandard English (2013: 64). It is this kind of gradual elimination of variation into proto-standard English that we expect to see in our present data.

However, applying the linguistics of speech approach to abbreviations is not entirely straightforward, as its claims are mainly based on spoken data. In a nonstandardised written culture such as Middle English between 1066 and 1350, it can be expected to influence written texts (cf. Kretzschmar and Stenroos 2012), even if writing may be somewhat conservative and there is no one-to-one correspondence between sounds and spellings. Less clear is how directly a linguistics of speech approach can be applied to writing systems and orthographic features. Abbreviations, in particular, are an interesting orthographic feature: On the one hand, they are a device developed to save time and space (Petti 1977: 22) and can thus be expected to be conditioned by the physical properties of the handwritten and printed space (Varila 2016; Shute 2017a; Tyrkkö 2017). On the other hand, they are legitimate spelling variants of their own that were part of the pool of variants available to writers of Middle and Early Modern English (Lass 2004; Driscoll 2009; Rogos 2012).

The fifteenth century, which saw the initial stages of a new written standard for English, also saw major developments in written culture. The most famous one is, of course, the printing press, but there were also other developments, including the less sudden but equally important paper revolution, which made manuscript books more affordable to a middle-class audience and promoted functional literacy, such as account-keeping, among merchants (see, e.g., Lyall 1989; Da Rold 2011; Robinson 2014; and Honkapohja 2017: 23–24). It would also be possible to see the loss of abbreviations in terms of shared practices of the scribal community and changes in the copying process. This type of visual pragmatics approach, utilised for the *Polychronicon* by Carroll et al. (2013), correlates the co-occurrence of particular visual elements of the manuscript or printed page with meaningmaking processes (see also Liira 2020: 274, 276). Nevertheless, because the focus of this chapter and book is on standardisation and multilingualism, we will only take material text into account in a somewhat limited manner: script, one- vs. two-column layout and the effect of line breaks are discussed in section 4.3.

As we are examining standardisation within a theoretical framework in which elimination of spelling variation is central, and basing it on studies which lead us to expect reduction in the course of the fifteenth century, it is worth asking whether abbreviations disappear at the same time as other variation. Is their disappearance related to these same processes, or is their gradual abandonment a separate process? This chapter seeks to answer these questions in addition to providing a diachronic outline for the gradual elimination of abbreviations from the late fourteenth century to the early sixteenth century. The *Polychronicon* provides a very good point of comparison for this, as it was consistently popular throughout the period. Our study takes into account both Latin and English, as well as manuscripts and early printed editions.

3 Data

3.1 The Polychronicon

Ranulph Higden's (OSB; d. 1364) *Polychronicon* is a universal chronicle divided into seven books, the first of which presents geographical information about the known world while the other six books narrate the history of the world from the biblical creation to Higden's own time. Higden continued revising the *Polychronicon* throughout his life: three distinct versions have been identified, and the intermediate version ending at 1342–1346 is found in the majority of the copies (Waldron 2004: xiii). The intermediate version of the chronicle was translated from the original Latin into Middle English by John Trevisa (fl. 1342–1402). The translation was requested by Trevisa's patron Sir Thomas

Berkeley, and it was finished in 1387 according to the translator's colophon. While Higden had composed his chronicle for a clerical audience, the readers of the English *Polychronicon* were presumably both aristocrats and clergy (Shepherd 1999: 31; see also Beal 2012: 68). Despite its universal theme, Higden's chronicle is heavily focused on the British and Irish Isles, and the translation reflects a wider interest in both vernacular literature and national history at this time (Matheson 1984: 209; Given-Wilson 2004: 139–140).

The English Polychronicon survives in fourteen manuscripts while the Latin manuscripts number over 120 (Waldron 2004: xiii). The English text was first printed by William Caxton in 1482 and again by Wynkyn de Worde and Peter Treveris, in 1495 and 1527, respectively. The work thus remained popular for over two centuries. It is particularly suitable for the study of standardisation as it covers the period which saw two remarkable changes in book production: the emergence of a new vernacular standard and the introduction of paper and printing. This provides an excellent starting point for a parallel corpus, allowing a comparison of spelling features across passages that are textually close to each other (cf. McIntosh, Samuels, and Benskin 1986: 2.1.3.). Waldron (1991: 67) has noted that "[w]hen the manuscripts of Trevisa's Middle English version of the Polychronicon have been fully transcribed and collated, they will yield (it can be safely said) a good deal of information on scribal attitudes to the language of the text being copied and on movements towards standardization in the written forms of English".⁴ For this reason, we selected copies of the *Polychronicon* for our quantitative study of the development of a supra-regional standard of written English.

3.2 The corpus

We sampled a selection of *Polychronicon* manuscripts for the present study, choosing one Latin manuscript, nine English manuscripts and the three early printed editions. Three aims guided the selection: to have at least two manuscripts from each 25-year period, to maximise the number of different scribes, and to select manuscripts from different parts of the stemmatological tree. Hu (Glasgow, University Library MS Hunter 223) was selected as the Latin manuscript for our corpus, as Waldron (2004: xviii) notes that Trevisa must have used a copy similar to this manuscript as his source text. The oldest extant manuscripts of the English translation, M (Chetham's Library MS Mun.A.6.90)

⁴ Steps towards this full collation have been taken by Waldron in his edition of Book 6 of the *Polychronicon* (2004).

and C (British Library MS Cotton Tiberius D.vii), were likely copied in the Berkeley area and are thus closest to the translator's original copy, now lost, in both date and language. A completely even representation turned out not to be possible, as the third quarter of the fifteenth century is only represented by one manuscript, T.⁵ Moreover, the section selected for our corpus in the two earliest manuscripts, C and M, was copied by the same scribe. To balance the selection we also added a second manuscript from "Scribe Delta" into the second period.

The manuscripts and printed editions used for the present study are briefly described below; for more extensive descriptions, see Waldron (2004: xviii–xxxvii) and Liira (forthcoming 2020).⁶ Figure 9.1 shows how the ones that were selected for the corpus are related to each other according to Waldron's stemmas.



Figure 9.1: Selection of manuscripts based on stemmatological tree (Waldron 2004: xxiii, reproduced with permission).

Latin manuscript

s. xiv **Hu** Glasgow MS Hunter 223 325 x 215 mm, single-column layout Single hand, Anglicana Formata

⁵ There is some fluctuation in the dating of MS B: Waldron (2004) dates it to xv^{med} but Mooney, Horobin & Stubbs (2011) to 1450–1475, following Dutschke (1989: 683), which would make it contemporary with MS T. Our findings, interestingly, suggest an earlier dating or perhaps simply conservative scribes.

⁶ The manuscript sigla used in this article are Waldron's; the sigla for the printed editions are ours.

English manuscripts

s. xiv/xv C London, British Library MS Cotton Tiberius D. vii (vol. 1) Current size (inlaid) 380 x 280 mm, single-column layout Single hand in vol. 1, Anglicana Formata: Hand 1, "The Polychronicon Scribe" (1–169r)

M Manchester, Chetham's Library MS Mun.A.6.90 350 x 265 mm, single-column layout Single hand, "The Polychronicon Scribe", Anglicana Formata

s. xvⁱⁿ

H London, British Library MS Harley 1900 350 x 240 mm, single-column layout Single hand, Anglicana Formata

 $s. xv^1$

A London, British Library MS Additional 24194 420 x 290 mm, two-column layout Single hand, "Delta", Anglicana Formata

F Tokyo, Senshu University Library MS 1 (olim Oslo/London Schøyen Collection MS 194)

420 x 285 mm, two-column layout

Single hand, "Trevisa-Gower Scribe", Anglicana Formata with some Textualis forms

J Cambridge, St John's College MS 204 (H.1) 395 x 300 mm, two-column layout Single hand, "Delta", Anglicana Formata

s. xv^{med}

B San Marino, Huntington Library MS HM 28561

380 x 275 mm, two-column layout

Four hands, two in the Polychronicon: Hand 1 (ff. 1–78r), Anglicana Formata and Secretary; Hand 2 (ff. 123v–319v), Secretary; the two alternating ff. 78–123r.

G Glasgow, University Library MS Hunter 367

260 x 255 mm, two-column layout

"Possibly three different hands, but could be one scribe" (Mooney, Horobin, and Stubbs 2011), Anglicana with some Secretary/continental forms

s. xv³

T Princeton, University Library MS Taylor 6 460 x 310 mm, two-column layout Single hand, the "Hooked-g Scribe", Bastard Secretary

English printed eds

Cax Caxton, William (Westminster, 1482), STC 13438⁷ 2°, single-column layout, Type 4:95B⁸

Wor De Worde, Wynkyn (Westminster, 1495), STC 13439 2°, two-column layout, Type 4:96G⁹

Tre Treveris, Peter (Southwark, 1527), STC 13440 2°, two-column layout

For our corpus, we selected a passage in Book 1, within chapters 1–7. This passage was chosen to maximise the amount of Latin, as it contains a list of Higden's references, copied in Latin even in the English manuscripts. In addition to the list of authorities, Latin occurs, for instance, in chapter titles and headings, and in the frequent source references in the running text. See Table 9.1 for the word count in English and in Latin in each transcribed and tagged corpus sample.

3.3 Encoding

We use a corpus-based approach with the intention of describing how abbreviations are reduced in the present corpus of 13 samples taken from the *Polychronicon*. To enable the quantitative analysis, a number of important divisions were annotated in the data: language, headings, line breaks, word divisions as well as both abbreviations and their expansions. The system of encoding was based on TEI P5 XML. However, to facilitate the encoding, the manuscript and printed witnesses were transcribed in MS Word, adding preliminary mark-up, which was converted to TEI XML using the scripts in the OxGarage web service. This was processed by running the automatically converted XML through a number of XSLT scripts which converted the automatically created tags into more semantically justified ones (cf. also Cummings 2009: 309–312).

⁷ Designated K by Waldron (2004).

⁸ BMC 11: 127.

⁹ BMC 11: 195.

MS/ed	Date	Word count	English	Latin	Percent of Latin
Hu	s. xiv	2112		2112	100%
С	s. xiv/xv	2828	2474	354	12.52%
М	s. xiv/xv	2861	2488	373	13.03%
Н	s. xv ⁱⁿ	2871	2497	374	13.03%
A	s. xv ¹	2744	2391	353	12.86%
F	s. xv ¹	2851	2471	380	13.33%
J	s. xv ¹	2812	2447	365	12.98%
В	s. xv ^{med}	2842	2489	353	12.42%
G	s. xv ^{med}	2730	2417	313	11.47%
Т	s. xv ³	2804	2440	364	12.98%
Cax	1482	2854	2491	363	12.72%
Wor	1495	2865	2482	383	13.37%
Tre	1527	2832	2458	374	13.2%
Total		36006	29545	4349	12.08%

Table 9.1: Corpus and word count.

TEI XML is well-suited for this type of study as it provides tags for all the features we wanted to study. As the aim is to take multilingualism into account, sections in Latin were tagged as <foreign lang="Lat">. The <foreign> tag "identifies a word or phrase as belonging to some language other than that of the surrounding text" (TEI Guidelines 3.3.2.1). Line breaks were tagged as <lb/>lb/>, marking "the beginning of a new (typographic) line in some edition or version of a text" (TEI Guidelines 3.10.3). Headings and marginal comments were indicated by <seg>-tags, which indicate any kind of segment (16.3), and the specific type of division was specified by attributes.

Each word was tagged inside <w> tags. If a word does not contain an abbreviation, the encoding is simple: <w>word</w>. Our definition of words is based on editorially identified word divisions. According to the TEI P5 Guidelines, <w> represents "a grammatical (not necessarily orthographic) word" (TEI Guidelines 17.1). The words were tagged according to what we considered to be separate words in transcription. The transcription is thus semi-diplomatic: it mainly follows the manuscript, but makes occasional editorial normalisations of word division. Words such as *shalbe* were annotated as two

separate words <w>shal</w><w>be</w>, and words such as *not withstandynge* as one word <w>not withstandynge</w>, regardless of where the scribes and typesetters left a space. Word token counts in the analysis are based on the numbers of <w> tags identified in the transcription.

The most important part of the tagging is the need to annotate both the abbreviations and their expansions. The TEI P5 Guidelines provide mechanisms which allow precisely this. The system for encoding used in this study is represented in detail in Honkapohja (2013) and discussed in Honkapohja (2018: 246–248),¹⁰ but we will present it here using one example, *P*^ofore 'therefore'. In its XML format the word looks like this:

```
<w>
<choice>
<abbr>Þ<am>hook</am>fore</abbr>
<expan>Þ<ex>er</ex>fore</expan>
</choice>
</w>
```

The possibility of including both the abbreviated and expanded form is enabled by <choice> tags, which are used for every abbreviated word. The tag "groups a number of alternative encodings for the same point in a text" (3.4). Within the <choice> structure, the two alternative encodings are marked with <abbr> (abbreviation), which "contains an abbreviation of any sort" and <expan> (expansion), which "contains the expansion of an abbreviation" (3.5.5). There are also specific tags for encoding abbreviations and expansions within these alternative structures. The <am> (abbreviation marker) tag is used inside the <abbr> tag, marking "a sequence of letters or signs present in an abbreviation" (TEI Guidelines 11.3.1.2).¹¹ Similarly, inside the <expan> tags, an <ex> (editorial expansion) is used to annotate "a sequence of letters added by an editor or transcriber when expanding an abbreviation" (TEI Guidelines 11.3.1.2). The tagging of abbreviations, Latin sections, line breaks and individual words enabled us to carry out the quantitative analyses presented in the following section.

¹⁰ See also Cummings (2009), Driscoll (2009) and Stutzmann (2014) for the application of similar systems of encoding.

¹¹ Because the aim of the paper was a corpus study and not displaying the abbreviations, we used names for all of them (e.g. hook).

4 Results and analysis

The natural starting point for studying the disappearance of abbreviations and suspensions is to provide descriptive statistics on how much they were used. While the higher number of abbreviations in Latin compared to the vernacular is often noted (Hector 1958: 36–37; Hasenohr 2002: 82–83; Laing 2013: 3.4.5.1), there are not many quantitative studies which would give us exact numbers; these would be useful for giving us some indication of what to expect. There are, however, a couple of recent studies. Stutzmann et al. (2018) report ca. 60% for Latin and 30% for French in a corpus comprising registers and formularies connected to the French royal chancery in the fourteenth and fifteenth centuries. Honkapohja (2018: 250–251) finds a mean abbreviation density of 34.8% for Latin recipes and a mean density of 9.2% in Middle English in medical texts copied ca. 1450–1490. These figures give some indication of what to expect with roughly contemporaneous data. However, it has to be noted that they represent different genres: administrative texts and medical treatises and recipes. Neither of the studies are structured diachronically.

4.1 Abbreviation density

Firstly, we calculated the abbreviation densities for the manuscripts and printed editions. These were counted by dividing the number of abbreviated words (<abbr> tags) by the word count (<w> tags). Figures 9.2, 9.3 and 9.4 display the abbreviation densities in chronological order. They reveal major differences between Latin and English frequencies, and a different rate of disappearance.

The most immediately striking aspect of the results displayed in Figure 9.2, which shows the combined abbreviation density of Latin and English, is how much more frequent abbreviation is in Latin than it is in Middle English. This is precisely what manuscript scholars are well aware of, but which has only rarely been quantified. In the *Polychronicon* manuscripts, more than half of the words, 55.49%, are abbreviated in the Latin Hu, whereas the two most heavily abbreviated English manuscripts, H and F, only abbreviate between 20 and 30% of the words: 26.82% and 23.36%, respectively. The figures for both Latin and Middle English are slightly lower than those reported by Honkapohja (2018) or Stutzmann et al. (2018). The likely explanations are genre and grade, as the *Polychronicon* manuscripts are *de luxe* productions, whereas Honkapohja examined medical texts and Stutzmann et al. administrative documents. Abbreviations were used more in less important and more utilitarian and workmanlike texts and copies (Roberts 2005: 9–12; Kopaczyk 2011: 95).

282 — Alpo Honkapohja and Aino Liira



Figure 9.2: Abbreviation density, English and Latin combined.



Figure 9.3: Abbreviation density, Latin.

The *Polychronicon* corpus is, however, able to provide us with a diachronic view to these developments. The other English manuscripts copied before the mid-fifteenth century (C, M, A, J) contain between 10 and 20% abbreviated words. The overall density drops close to 10% and below in the three manuscripts copied in the mid- to late fifteenth century (B, G, T), and below 5% in early printed books (Cax, Wor, Tre). There is, however, some variation between copies contemporary to each other, as the s. xv^3 manuscript T contains more abbreviations, 7.35%, than the s. xv^{med} manuscript G, in which the density is 5.49%. Even more surprisingly, the lowest number of abbreviations is used by Caxton, whereas their number is



Figure 9.4: Abbreviation density, English.

higher in the two later printed books, De Worde and Treveris. These differences show that while there is a general downward trend, the reduction of abbreviations varied based on the preferences and practices of scribes and typesetters.

Looking at abbreviation densities for only the Latin sections reveals higher density, less variation and a slower decrease. These are illustrated by Figure 9.3. More than half of Latin words are abbreviated in all but two manuscripts (A and J) copied before 1425. Moreover, abbreviation density remains at least 25% even in the later manuscripts. MS B from the mid-fifteenth century still contains 53.42% of abbreviated words and G and T contain 27.48% and 32.14%, respectively. This shows that the scribes copying both languages retained the abbreviation and suspension system for Latin much longer than they did for the vernacular. A likely reason for the large amount of abbreviation in Latin is that in the present data it is used in paratextual elements as well as bibliographical information. For example, Stam (2017: 79, paraphrasing Tristram 1997) lists the economical nature of the Latin abbreviation system as one of the reasons for using Latin. Consequently, the present figures are likely to show the importance and practicality of the highly developed Latin abbreviation system in citations and headings.

What is particularly striking is that the Latin manuscript (Hu) no longer stands out as the most heavily abbreviated witness. The abbreviation density of Hu is actually lower than in three out of the four earliest English manuscripts (C, M, H). The density is lower regardless of whether the point of comparison is the entire text (Hu) or solely the sections and headings that are in Latin in all manuscripts (Hu_bibl). The abbreviation density for Hu_bibl is even slightly lower (54.47%) than the entire sample (Hu: 55.49%). This indicates that the Latin abbreviation and suspension system remained in use at least until 1425, and that, unlike in English, the decrease in density remained moderate even later in the century.

In English, the abbreviation density shows a sharper contrast between the practices of individual scribes and drops below 5% after the mid-fifteenth century. Figure 9.4 displays the abbreviation density of all the Middle English in the present data. It shows a gradual reduction that happens earlier than in Latin, but also major differences in the abbreviation profiles of individual scribes. The work of the two heavily abbreviating scribes (H, 21.97%, and F, 18.74%), in particular, stands out as they abbreviate twice as many words as the other scribes active before 1425 (C, M, A, J). Another striking feature is that the slightly later H contains several more abbreviations than the two earliest, C and M, which according to Waldron (2004: xxxix) are close to Trevisa's archetype. It has to be noted, though, that C and M both represent the work of one scribe, "The Polychronicon Scribe", as do A and J, copied by "Scribe Delta" (see section 3.2 above). It would thus appear that individual scribes largely vary in their use of abbreviations and that the abbreviation densities of scribes are fairly similar across different manuscripts they have copied. In order to understand what constitutes variation in the work of various scribes, it is necessary to look closer at which words different scribes abbreviate and which they expand.

The higher abbreviation density of the heavily abbreviating scribes, H and F, can be explained partly by their propensity to use abbreviations for frequent function words and partly by their frequent use of two general signs of abbreviation, the macron (for a nasal $fr\bar{o}$ 'from', $sec\bar{o}d$ 'second') and the hook (for <er>, <re> or <e>: ger^9 'year', man^9 'manner', $p^9 fore$ 'therefore'). Table 9.2 shows the ten most frequent abbreviated words in four early manuscripts, which makes it possible to compare the practices of H and F with the less enthusiastically-abbreviating M and A. Abbreviations for small grammatical words such as & 'and' and p^t 'that' are very high in all of the earlier manuscripts, and constitute the top two in three of the manuscripts (M, H, F). However, the scribes of H and F tend to abbreviate two additional function words, 'from' (H: 41 tokens, F: 14 tokens) and 'in' (H: 14 tokens, F: 9 tokens), which contributes to their higher abbreviation density. Manuscript A already has a much lower percentage of the ampersand (36 tokens) in comparison with the other three pre-1425 manuscripts (142, 142, 168).

Lexical words also get abbreviated and some types of abbreviation are applicable to many different words. The heavily abbreviating H and F scribes also have a number of abbreviations which both use frequently, but other scribes normally spell out, such as abbreviating 'Christ' with a superscript (H: 9 tokens, F: 12 tokens, interestingly, the Hu scribe uses Nomina Sacra based abbreviations for all of these, but the practice is completely different in the English tradition). Two abbreviations in particular, the macron and the hook, are used productively

Chetham (M)		Harley (H)		Tokyo (F)		Add 24194 (A)	
& 'and'	142	& 'and'	142	& 'and'	168	& 'and'	36
þ ^t 'that'	þ ^t 'that' 38 þ ^t 'that' 52		þ ^t 'that'	23	þ ^e 'the'	18	
vnd ^r 'under'	11	frō 'from'	41	frā 'from'	14	secoūde 'second'	10
her ⁹ 'here'	6	īʻin'	14	all ⁹ 'all'	12	oþ ⁹ 'other'	7
diu ⁹ sce 'diverse'	5	3er ⁹ 'year'	10	c ⁱ st 'christ'	12	ptie ^s 'parties'	6
frā 'from'	5	vnd ⁹ 'under'	10	ī 'in'	9	þ ⁹ fore 'therefore'	6
man ⁹ 'manner'	5	c ⁱ st 'christ'	9	secōd 'second'	9	acoūted 'accounted'	5
oþ ⁹ 'other'	5	oþ ⁹ 'other'	8	man ⁹ 'manner'	6	seuē 'seven'	5
₽tyes 'parties'	5	bygȳnȳg 'beginning'	7	oþ ⁹ 'other'	6	seuēþ 'seventh'	5
þ ⁹ fore 'therefore'	5	pties 'parties'	6	pties 'parties'	6	acoūteþ 'accounts'	4

Table 9.2: The ten most frequently abbreviated English words in the early manuscripts.

and can thus contribute to a higher abbreviation density. The H and F scribes, for example, use the hook in words like $3er^{\circ}$ 'year' (H: 10 tokens) and all° 'all' (F: 12 tokens), and the macron in words like $byg\bar{y}n\bar{y}g$ 'beginning' (H: 7 tokens). However, these abbreviations do not affect the overall density as much as the frequent function words. Scribe Delta (A and J), who is the least abbreviating scribe before 1425, uses the macron very frequently, abbreviating numerals such as *secoūde* 'second' (10), *seuē* 'seven' (5) and *seuēp* 'seventh' (5) and verbs *acoūted* 'accounted' (5), *acoūtep* 'accounts' (4). Nevertheless, his lower frequency of using the ampersand (36) and omission of p^t 'that' brings the overall density down. Consequently, because of their frequency, function words constitute a major part of an individual scribe's abbreviation repertoire. By the mid-fifteenth century, the scribes copying the English *Polychronicon* were usually expanding them.

The habit of the later scribes of spelling out function words becomes striking when one moves to the manuscripts copied in the latter half of the fifteenth century. The majority of abbreviations in Table 9.3, which shows abbreviations from the three manuscripts dated to the mid- to late fifteenth century, are lexical words. There is one exception, since the scribe of T uses the ampersand

San Marino (B)		Hunter 367 (G)		Taylor (T)	
age ^s 'ages' ¹²	34	eu ⁹ ych 'every'	4	& 'and'	57
oþ ⁹ 'other'	9	t ^a nsmig ^a cioun a 'transmigration'	4	comoū 'common'	2
pties 'parties'	6	abrah ^a m 'Abraham'	3	sūme 'some'	2
sōme 'some'	6	man ⁹ 'manner'	3	Abrah ^a m 'Abraham'	1
Man ⁹ 'manner'	5	mē 'men'	3	accoūt 'account'	1
Ryu ⁹ 'river'	5	descripcioū 'description'	2	Cesar ⁹ 'Caesar'	1
diu ⁹ se 'diverse'	4	incarnacioū 'incarnation'	2	comoū 'common'	1
eu ⁹ ech 'every'	4	Isrłl 'Israel'	2	correccioū 'correction'	1
secūde 'second'	4	Aporcioū 'proportion'	2	cūnyng 'cunning'	1
t ^a nsmigracion 'transmigration'	4	sūme 'some'	2	depted 'departed'	1

 Table 9.3: The ten most frequently abbreviated words in the late manuscripts.

frequently (57 tokens). It is this frequency that causes the slight increase in abbreviation density from the mid-fifteenth century to the third quarter (see Figure 9.4 above). If one does not take the ampersand into consideration, there is a steady reduction in abbreviation counts. The frequencies of words making it to the list are also generally lower. The only other abbreviated word with more than ten tokens for this period is the practice of the B scribe to abbreviate 'ages' with a superscript <s> (34 tokens). The frequencies of abbreviated lexical words, in contrast, do not experience a rapid drop but rather a slow and steady reduction as the century progresses. Thus, the major drop in the abbreviated forms of function words. Surprisingly, they do experience a small-scale renaissance in the early printed books (see Table 9.4).

¹² Strictly speaking the form age^s is not an abbreviation as it does not shorten the word. It does, however, serve the function of saving space.

Caxton (Cax)		De Worde (Wor)		Treveris (Tre)	
& 'and'	10	& 'and'	34	y ^e 'the'	19
acouted 'accounted'	1	y ^e 'the'	13	& 'and'	13
acoūten 'account'	1	y ^t 'that'	6	y ^t 'that'	4
circūsicon 'circumcision'	1	acoūte 'account'	2	acouted 'accounted'	2
venemo9 'venomous'	1	acouted 'accounted'	2	foūdē '[was] found'	2
		foūden '[was] found'	1	mē 'men'	2
		Gouernyge 'governing'	1	Whā 'when'	2
		Normās 'Normans'	1	wodres 'wonders'	2
		Puynce 'province'	1	wrytē 'write'	2
		quatyte 'quantity'	1	acountē 'account'	1

Table 9.4: The most frequently abbreviated English words in the early printed books.

Table 9.4 shows two developments. On the one hand, there is a further reduction in the types and density for abbreviations of lexical words; on the other, a small "standard" set of abbreviations has emerged. This small set of popular types used by the printers comprises the ampersand (&) and the superscript abbreviations y^e 'the' and y^t 'that'. We will refer to it as the Standard Printer Set of Abbreviations (henceforth SPSA). The higher abbreviation density of De Worde and Treveris compared to Caxton is largely explained by the absence of two items from SPSA, as Caxton only uses the ampersand. Other types of abbreviations are still used occasionally. All of the typesetters sometimes abbreviate a nasal with the macron. The macron, too, is used by De Worde and Treveris slightly more often. Moreover, all still make occasional use of Latin abbreviations for Romance loan words, such as *venemo9* 'venomous' by Caxton or *puynce* 'province' by De Worde.¹³ This indicates that they did have these types as a part of their printing sets, even if they were not used very frequently.

Interestingly, two of the abbreviations which later became part of the SPSA, b^t 'that' and w^t 'with' (which is not used by the scribes in the present *Polychronicon* data), are discussed by Samuels (1983), who observes their

¹³ De Worde also uses the 9 '-us' abbreviation for a number of Latin words, including personal names: *Methodi9* 'Methodius', *Marian9* 'Marianus'.

frequencies in connection with Types II and III of London English and diachronic change between 1400 and 1420. His data consist of samples from one scribe. known as scribe B (cf. Doyle and Parkes 1978), perhaps to be identified as Adam Pinkhurst (cf. Mooney 2006), who copied the two earliest Canterbury Tales manuscripts.¹⁴ Samuels examines the proportion of b^t and *that* used by the scribe in the stints he contributed to three manuscripts.¹⁵ He notes far fewer uses of the abbreviated form in the latest manuscript, and explains this with changes in the scribal habits of scribe B (Samuels 1983: 51). The scribe was adapting to changes of spelling fashions in London, which had "only recently undergone a complete metamorphosis [...] from type II to type III" (1983: 53). This change happened in 1400–1420 and "was crucial for the development of Standard English, for it was from the competing and changing fashions in spellings at this time that the new written standard was to evolve" (1983: 53). One of the changes Samuels mentions was an overall move away from thorn, which is also reflected in the abbreviated form, as "*b* was obsolescent in London by this period and being replaced by *th*" (1983:59).¹⁶

Even though the change happens slightly later, our data is generally in line with Samuels' observations. The proportion of p^t does decrease sharply after the early fifteenth century, so it appears that there was a shift in fashion of the practices of scribes contributing to these types of *de luxe* books. The copyists of M, C, H and F (see Table 9.9 below) use it as the major form; others as a minor form, probably when the constraints of space require it. Nevertheless, the abbreviated forms did not disappear completely, and they survived into printed books, which shows that they were still part of the repertoire. In addition to their appropriation by early printers, abbreviations like y^e 'the' and y^t 'that' were used in correspondence centuries later. Our results also show that it is necessary to look at how the frequencies of different abbreviation types developed diachronically, which will be the focus of the next section.

¹⁴ For criticism of Mooney's identification, see Roberts (2011) and Warner (2015, 2018). For a different account of spellings, see Thaisen (2011), who discusses his short and long variants in terms of space.

¹⁵ These are the "Hengwrt Chaucer" Peniarth MS 392D, the "Ellesmere Chaucer" Huntington Library, in San Marino, California (EL 26 C 9), and a copy of Gower in Cambridge, Trinity College, MS R.3.2.

¹⁶ Thaisen (2011: 84) attributes this variation purely to constraints of time and space. Samuels (1983: 58–59) also notes a shift in the proportions of & and *and* in favour of the expanded variant.

4.2 Abbreviation types

Developments in frequencies of abbreviation types have been the subject of a few interesting discoveries, but these have mainly focused on a few individual types. The view in palaeographical handbooks is that abbreviating in the vernacular never was as common as in Latin. Both Hector (1958) and Petti (1977: 22) note that the Latin system reached its most elaborate form by the twelfth century and then began to decrease. According to Hector (1958: 28): "After about 1200 no new abbreviations were introduced into the writing of Latin, and during the later Middle Ages some of those that had formerly been in regular use were gradually discarded". He, however, does note their presence in Latin documents after 1500 (1958: 38). The English system was based on the Latin one from the beginning. For example, Hector (1958: 37) says that it was applied to native proper names in Latin documents and "when archives came to be written in English language there was thus already established a tradition". Petti (1977: 22) notes a general pattern of reduction, but also survival to the Renaissance in formulaic uses.¹⁷

Nonetheless, there are some results which point to exceptions to this overall trend. One of these is identified by Samuels (1983), who, in contrast to his proposed drop in frequencies for p^t , notes that the frequencies of some abbreviations actually increase. These include a shift from expanded <er> to abbreviated \circ 'er', "from h to crossed h [...] and d to tailed d" (58) as well as from *with* to w^t . Thus, according to him, the move from Type II to III by 1420 also involved an increase of a few common types of abbreviation. These observations are supported by recent work on Older Scots and continental French.

The Older Scots results are presented by Smith (2019: 202–203, 208), who studied the use of a single abbreviation, the final loop f for '-is' plural in legal documents from 1380 to 1500 in the *Linguistic Atlas of Older Scots* (LAOS). Smith suggests that "as vernacular writing became more common, scribes began to use more 'shorthand' features" (2019: 208). Moreover, she notes that scribes from densely populated Lowland repositories of Scotland use it more than ones in more peripheral areas.

Similar developments in which some abbreviations increase in frequency have been noted by French scholars, who have elaborated them into a hypothesis. French scholarship on abbreviations consists of observational work by

¹⁷ Parkes (2008 [1969]: preface), which would be the most directly relevant source for the Anglicana and Secretary hands in the present study, does not discuss abbreviations due to "limitations of space".

Bozzolo et al. (1990), Careri et al. (2001), Hasenohr (2002) and Careri et al (2011). Their observations are studied quantitatively using a small corpus of manuscripts of *La Chanson d'Otinel* by Camps (2016). The French account of Latin abbreviations is essentially in line with Hector (1958) and Petti (1977). The number of abbreviations in Latin multiplied up to the twelfth-century renaissance, which saw an expansion of written culture outside monasteries, and even a new scholastic way of reading and writing, which differed from the slow meditative reading practised in monasteries (Hasenohr 2002: 81–82). In contrast, vernacular romance manuscripts copied in the twelfth century lack abbreviations almost completely (2002: 81–82).¹⁸ But whereas the density and type count of Latin abbreviations gradually decreased after the twelfth century, the vernacular abbreviation system grew increasingly independent of it in a process that can be labelled *francisation* 'Frenchisisation' (Camps 2016: cclix).

It seems likely that a similar process took place in English, which explains the developments noted by Samuels (1983) and Smith (2019). Some parts of the abbreviation system, such as the hook \circ for <er, re, e> or final loop f for plural <is, ys, es> proved useful for vernacular copyists. As scribes became increasingly fluent in using them and more and more vernacular texts were copied, their frequency increased. If this kind of process is found to take place it could be labelled "Anglicisation" or maybe just "vernacularisation" of the abbreviation system. But what kind of developments can we observe for various abbreviation types in the *Polychronicon* corpus?

In order to study how the frequencies of various abbreviations change over time, we will apply a type of diagram commonly used in variationist sociolinguistics: the S-curve. The S-curve is useful in illustrating how the variant forms of a linguistic variable change diachronically (see, e.g., Kroch 1989; Labov 1994: 65–67; Kretzschmar and Stenroos 2012; Nevalainen 2015). To compare mean frequencies, the twelve English samples were divided into four periods, each consisting of three manuscripts/books.¹⁹ Mean frequencies were counted for each period.

¹⁸ This corresponds to what Laing (2013: 3.4.5.1) notes for the *Linguistic Atlas of Early Middle English* (LAEME).

¹⁹ This division has some problems: Firstly, the two first sub-periods have three manuscripts, but each period represents the output of only two scribes. However, both of them also contain one outlier, that is, a manuscript containing exceptionally many abbreviations, which balances them out. Secondly, the third sub-period is longer than the others, covering 50 years. However, in here the results do provide some useful generalisations. The B scribe is considerably more conservative than the G scribe, even though Waldron (2004) dates both of these manuscripts to s. xv^{med} . On the whole, this division into sub-periods did provide some significant results using the s-curve.

Early manuscripts: M, C, H 1400–1425: F, A, J 1425–1475: B, G, T Printed books (1484–1527): Cax, Wor, Tre

To keep the number of variables manageable, the abbreviation types were grouped into six categories:²⁰ (1) the ampersand (&), (2) the macron (*circūsicon* 'circumsicion'), (3) the hook (*eu⁹ych* 'every'), (4) other brevigraphs (*puynce* 'province', $\mathcal{P}tie^{s}$ 'parties'), (5) superscripts (b^{t} 'that', $abrah^{a}m$ 'Abraham'), (6) the strikethrough (crossing the ascender in certain letters such as ħ or ł, sometimes indicating a final <e>, sometimes otiose). The reason for treating the macron and the hook on their own is their high frequency. These two abbreviations could be highly productive: together with the ampersand, these abbreviations represent much more than half of all the abbreviations in the data. On the other hand, there are numerous other brevigraphs, which is an umbrella term for Latin-based abbreviations which sometimes resemble the letters they replace and sometimes have an apparently arbitrary shape (Petti 1977: 23). As there are several fairly low frequency brevigraphs, treating them as separate would only show too many very low frequency items to be of use. Superscript abbreviations cause the same problem. Moreover, as they are part of the Latin alphabet, they are an openended category, which can lead to a high number of types. Consequently, these six categories treat categories 1 to 3 as individual types, while categories 4–6 are amalgams of several less frequent types. Together they offer an illuminating overview of the diachronic developments that took place from the late fourteenth to the early sixteenth century.

The S-curves reveal how the slower decrease, evident in the density diagrams, progresses with respect to abbreviation types. Figures 9.5 and 9.6 illustrate how these take place. In English, there is a major overall decrease between the second period and the third period (after 1425). In Latin, however, all types remain fairly frequent even in the final manuscript period (1425–1475), but there is a drop in the frequencies of all types except brevigraphs and macrons.

Some abbreviation types show a surprising increase in the early period. One of these is the macron, which actually increases in English from the first period to the second. There are two possible explanations for this. One is related to

²⁰ These categories are based on a system of taxonomies that can be found in many palaeographical handbooks. The system was first introduced by Chassant ([1845] 1970), and very influential versions of it are presented by Cappelli (1990) and Petti (1977). For an account of various abbreviation categories as well as their treatment in handbook literature, see Honkapohja (2013: sections 1 to 4).



Figure 9.5: Diachronic changes in mean number of abbreviation types in English.²¹

Frenchisisation. Camps (2016: cclix) notes that, while there is a drop in the diversity of abbreviation types after the twelfth century, the density of a few common abbreviation types, especially the macron, may increase both in Latin and French. If an increase in the use macron in English in this period is corroborated by other studies we could indeed speak of the "Anglicisation" or "vernacularisation" of the abbreviation system. On the other hand it has to be kept in mind that our corpus is fairly small and the same developments could also simply be due to scribal preferences.

Indeed, there is support for the explanation that the reason can be found in scribal preferences. The "Polychronicon Scribe" of C and M does not use the macron very much, whereas "Scribe Delta" (A and J) and the "Trevisa-Gower Scribe" (F) both use it frequently (see Table 9.2). A similar jump can be observed also for the parts that are in Latin, as the mean number of macrons and hooks increases from Hu to the early English manuscripts. This is because the "Polychronicon Scribe" uses the macron for Latin, although he does not use it for English. Moreover, he and the Harley scribe (H) make much use of the convention of writing Latin endings in superscript. These increases are consistent

²¹ The figure omits two categories: ampersand and strikethrough. For ampersand, see Figure. 9.7. Strikethrough, on the other hand, is very infrequent: the mean numbers are (4.3333, 3.3333, 1.6666667, 0).



Figure 9.6: Diachronic changes in mean number of abbreviation types in Latin.²²

with the fact that the abbreviation density for Latin is higher in English manuscripts before 1425 than in Hu (see Figure 9.3 above).

Both English and Latin abbreviation repertoires experience change and reduction from manuscript to print. All types of abbreviation are present in the third period, but the repertoire of abbreviations is much reduced in printed books. In addition to the macron, which is sometimes used in both languages, there are two specific developments for English and Latin. In Latin the main type of abbreviations that remain in use are brevigraphs. In English, the emergence of SPSA causes a slight increase for superscript abbreviations from late manuscripts to printed books.

Figure 9.7 illustrates the mean density of abbreviating 'and', 'the' and 'that' in an S-curve. It shows how, in the early period, both & 'and' and y^t 'that' are the major variant, used more than 70 per cent of the time. The development partly corresponds with what Samuels mentions of the use of the abbreviated form of 'that' in two *Canterbury Tales* manuscripts. However, our results do not support his idea that y^t is caused by the shift from thorn to taking place

²² For the sake of comparison we used only the part of Latin that is in Latin even in the English MSS.



Figure 9.7: Mean density of three SPSA abbreviations.

between 1400 and 1420. In the *Polychronicon* corpus, spellings with thorn remain the majority form until manuscript witnesses G and T, which are from the middle of the century (see Table 9.5 for 'the' and Table 9.9 for 'that'), whereas a drop in the number of the abbreviated forms happened by 1425.

The proportion of abbreviating 'and' shows a fairly steady decrease up to printed books, in which the word is still abbreviated 12.16% of the time. The two thorn-based abbreviations are not quite as frequent, partly because Caxton does not use them (see Table 9.4 above), but both show the interesting development in which the mean frequency of the abbreviated forms increases with printing. The superscript variant of the definite article, y^e , is a particularly interesting case as it is not abbreviated by the scribes of the early manuscripts and only once between 1425 and 1475, but is clearly part of the repertoire of the typesetters of De Worde (13 tokens) and Treveris (19 tokens). To shed light on what is causing this curve, we will next examine the actual forms.

Table 9.5 shows the distribution of spellings for the definite article. It clearly deviates from the expected development in which a gradual elimination of variants takes place. The two early scribes, H and M, are completely consistent in spelling the word expanded and with an initial thorn: *be*. However, the scribes working in the next period, 1400-1425, use several spelling variants, including as many as four different ones used by the scribes of F and J. Nevertheless, abbreviations of the definite article remain the minority form, as only one scribe (A) abbreviates the word more than ten times

Early manuscripts	м	be	235						
	Н	þe	242						
1400–1425	А	þe	219	þ ^e	18	the	3		
	F	þe	215	the	20	þ ^e	5	þþe	2
	J	þe	223	þ ^e	7	the	7	þhe	1
1425-1475	В	þe	243						
	G	the	238						
	Т	the	227	þ ^e	1				
Printed Books	Cax	the	253						
	Wor	the	246	y ^e	13				
	Tre	the	220	y ^e	19				

Table 9.5: Distribution of variants of 'the'.

(18 tokens). The next period, 1425–1475, sees a step toward standardisation and the emergence of spellings, but only a single token of the abbreviated variant. However, De Worde and Treveris make occasional use of the abbreviation as an alternative to the already modern spelling. In Latin manuscripts, a bigger set of abbreviations remained in use.

One of the characteristic features of the Latin abbreviation system were the so-called brevigraphs, many of which first emerge in the fifth and sixth centuries (cf. Lindsay 1915: 3; Hasenohr 2002: 80). A number of brevigraphs were also applied to English, such as the ubiquitous hook, as well as p 'pro' and p 'per/par', which were useful for Romance loan words such as *proporcion*, *profit*, *departed*, *persones*. However, the number of brevigraphs used for English was no match for the great range of variety found in Latin. Figure 9.8 below shows the mean number of brevigraphs (other than hook), for each period, in English and Latin.²³ Despite the ten times higher word count for English in our data, these types of abbreviations are more common in Latin. They also remain in use by early printers, especially in Latin but occasionally also in English.

Two printers, De Worde and Treveris, use brevigraphs for the Latin bibliographical section. Figure 9.9 shows the brevigraphs and macrons used by

²³ The Latin point of comparison is with Hu_bibl, that is, the part that remains in Latin in the English translations.



→ Lat brevigraph → Eng brevigraph

Figure 9.8: Diachronic changes in mean number of brevigraphs.

tins. Sultin Enfetus in hydona teifeilainta aif Ji. lunt ibn. Chri tona teifeilaitia troptrita air fya fut audores Eufebnus Beronuh? et Throdonus epifopus. Edugufti nus de cuitate der potifiume "thu. et potif. Dollas bripatrita in bin. aifti is preiture in lubo te oinetha singit Upponus bripatrita in lubo. aifti mologit. Solin?de nurabilit/mai ou. Christias bantyndoo Elefidia ou?. L'Eurodi? do britona romas ya. Walte? Foruntis artipatao? "L Paulus opacan? do britona lon gobardor. Calinobe buretacento tofaunt?. Calinobes de gittes undoanties un britona ber gittes undoanties undo britona ber gittes undoanties undo britona britona te

tunus pentenciarus piripape in cos nicis luis de imperatorité i e pontitaciaus. E Partian (vonce, 10 villetar) malmetburentis monach? de gettus Kega indire to pontificar filorenz cues buggen inopachus die jo annég imputatorona (ni Data tano pontifimet junt fectures.

Cprefacio fecunda ad hyftoriam La pitulum fecundum.

Annote the series of the serie

Figure 9.9: Latin abbreviations used by Caxton and De Worde (Left: *The Polychronicon*, William Caxton, 1482 (STC 13438). © British Library Board (G.6011-12, f. 7r). Image published with permission of ProQuest. Further reproduction is prohibited without permission. Right: *The Polychronicon*, Wynkyn De Worde, 1495 (STC 13439). © British Library Board (C.11.b.2., f. 5r). Image published with permission of ProQuest. Further reproducts. Further reproduction is prohibited without permission. Images produced by ProQuest as part of *Early English Books Online*. www.pro quest.com.).

Caxton and De Worde. Perhaps surprisingly, De Worde uses them very frequently. Individual examples such as the Latin word *triptita* 'three-part' (illustrated in the figure) and the word *venemo9* 'venomous' (see Table 9.4 above) reveal, however, that these symbols were included in Caxton's printing type sets. It is possible that De Worde uses more abbreviations because of the twocolumn layout. Therefore, we will next investigate the effects of material limitations on abbreviation frequencies.

4.3 Developments in book production

One of the major questions this study aims to answer is how far the abbreviations were conditioned by technology and how far by standardisation, as the fifteenth century also saw the so-called paper revolution, which made book production cheaper and thus may have reduced the need for abbreviation. There are also changes in the *mise-en-page*, since some of the manuscripts have a double-column layout and some a single-column one. Moreover, there is a shift from Anglicana hands to Secretary hands. The purpose of this section is to examine these changes.

The first change to be discussed is paper. The English versions of the *Polychronicon* are high-grade manuscripts, and continue to be copied on parchment until the late fifteenth century. None of the manuscript copies included in this study is on paper, whereas all of the three printed books (Cax, Wor and Tre) are. The major reduction for English abbreviations happens from the first quarter of the fifteenth century to the mid-fifteenth century (see Figures 9.4 and 9.5) and the major reduction of Latin abbreviations from manuscripts to print (see Figures 9.3 and 9.6). While the present data do not allow us to draw conclusions of the influence of writing support (as opposed to the influence of print technology), it would appear that the switch from parchment to paper is not an important dividing line, as the number of English abbreviations begins to decline in the manuscripts copied on parchment.

The second change comes with the script used to copy the body of the text. The majority of the manuscripts in the data are copied in varieties of the Anglicana script, mainly Anglicana formata, which was a common book script used in four-teenth-century England (see Parkes 2008 [1969]: xvii). The manuscripts dating from the period 1425–1475 display Secretary or Secretary-influenced hands: B has two scribes working on the *Polychronicon*, one of whom writes Anglicana and Secretary, the other Secretary, a script which mixes Secretary with Textura influences in the style of the French Bastarde script (Parkes 2008 [1969]: xxi). The choice of script depends on the preferences of the commissioner and the scribe, but also the current fashion. The higher abbreviation frequencies seem to correlate with Anglicana. With the limited data it is difficult to say, however, if the script is

a deciding factor or if the low frequencies attested in the Secretary manuscripts are due to a general decrease in the number of abbreviations in this period.²⁴

A general question is whether abbreviations are simply used to save space, as this was always one of the main reasons to use an abbreviated variant instead of spelling the word out (see Samuels 1983: 59; Thaisen 2011: 84). Petti (1977: 22) notes that abbreviation could also be used "for keeping lines of writing of equal length, rather like 'justifying' in printing". If they are used for this purpose, one would expect them to be more frequent closer to the right margin. Indeed, Shute (2017a) discovered that early modern printers used abbreviations as one of two main strategies for right-margin justification: that is, when they needed to fit a word on a page, they could use an abbreviation to make it fit (see also Camps 2016: ccli–ccliii).

In order to investigate whether we can reproduce the effect discovered by Shute, we calculated the proximity of <am> tags to <lb> tags that signify a line break. Each line was divided into five bins. The bins were summed up.²⁵ If abbreviations were used for justification, we would expect the number in the rightmost bin to be significantly higher than others. If they were not, then the variation would not likely be statistically significant.

The results reveal that right-margin justification is an important conditioning factor in manuscripts, too, but also that it does not apply all the time (see tables 9.6 and 9.7). In most cases, the number of abbreviations is highest in the rightmost bin (see Figures 9.10 and 9.11). However, neither Latin nor English shows large enough differences to be statistically significant for the four earliest manuscripts (Hu, C, M, H). The differences become overwhelmingly significant in English manuscripts dated to the first quarter of the fifteenth century. This change corresponds with the change to a two-column layout (see Table 9.7). With regard to the manuscripts from the latter half of the century and the early printed books, the results are less clear. The results are not statistically significant for T or, perhaps surprisingly, Wor. However, the token counts for printed books are so much lower for the latter period that the element of random chance cannot be ruled out. In Latin, the differences are not statistically

²⁴ A similar question can be raised regarding the printed books: to what extent do characteristics such as the size of a typeface influence the compositors' abbreviation practices, in addition to the obvious limitations imposed by the selection of types? Pursuing this question is, unfortunately, outside the scope of the present study.

²⁵ For this calculation, abbreviations that were part of headings or marginal headings were omitted, which means that the numbers are slightly lower than the ones reported in overall counts. In addition, a number of exceptionally short or long lines were omitted.

MS	Date	Matr	Cols	bin 1	bin 2	bin 3	bin 4	bin 5	p-value	
Hu	s. xiv	parch	one	244	255	282	258	303	0.076738	NO
С	s. xiv/xv	parch	one	25	33	34	32	35	0.740386	NO
М	s. xiv/xv	parch	one	29	30	32	32	34	0.975033	NO
н	s. xv(in)	parch	one	42	47	57	43	60	0.245256	NO
A	s. xv(1)	parch	two	18	22	27	30	27	0.453781	NO
F	s. xv(1)	parch	two	34	26	22	35	32	0.381171	NO
J	s. xv(1)	parch	two	16	19	26	16	23	0.419709	NO
В	s. xv(med)	parch	two	22	30	25	29	38	0.277467	NO
G	s. xv(med)	parch	two	12	10	13	13	20	0.378859	NO
т	s. xv(3)	parch	two	12	10	14	15	19	0.511199	NO
Cax	1482	paper	one	0	0	2	4	5	0.0053	YES
Wor	1495	paper	two	8	11	10	11	16	0.540058	NO
Tre	1527	paper	two	2	6	5	4	14	0.008399	YES

Table 9.6: Right margin justification, Latin.



Figure 9.10: Right margin justification in three early manuscripts.



Printed books: right-margin justification

Figure 9.11: Right margin justification in three early printed books.

MS	Date	Matr	Cols	bin 1	bin 2	bin 3	bin 4	bin 5	p-value	
Hu	s. xiv	parch	one							
С	s. xiv/xv	parch	one	64	59	55	55	61	0.904595	NO
М	s. xiv/xv	parch	one	34	26	40	38	33	0.490888	NO
Н	s. xv(in)	parch	one	115	113	109	108	122	0.893628	NO
A	s. xv(1)	parch	two	34	34	35	50	80	6.26E-08	YES
F	s. xv(1)	parch	two	76	85	83	85	156	3.25E-09	YES
J	s. xv(1)	parch	two	45	48	40	53	85	9.97E-05	YES
В	s. xv(med)	parch	two	28	30	43	29	55	0.004758	YES
G	s. xv(med)	parch	two	15	10	11	18	16	0.511199	NO
т	s. xv(3)	parch	two	13	14	20	13	28	0.047464	YES
Cax	1482	paper	one	3	3	7	2	6	0.34538	NO
Wor	1495	paper	two	8	13	19	11	18	0.178522	NO
Tre	1527	paper	two	12	12	6	11	21	0.050751	NO

Table 9.7: Right margin justification, English.

significant, except in the early printed books – even though the right hand bin is always the biggest (see Table 9.6).

The fact that abbreviations are most frequent in the rightmost bin, but the difference in the earlier manuscripts is so small that it could be due to random chance, suggests that while the economy of space was always an important conditioning factor for abbreviating, there were also other reasons to use them in the fourteenth century. When the abbreviation and suspension system began to fall out of use, these results seem to suggest that it still continued to be used for line justification purposes. Shute's results suggest that the system was adapted to this end by the early printers (cf. Shute 2017a). The genre may also affect the results, since the two-column layout in *de luxe* quality required the scribes to produce very even margins; to achieve this they used the abbreviations as a justification device in the manner noted by Petti (1977:22).

From the point of view of standardisation, however, our results suggest that, as the system began to disappear, the function of justification was among the last to survive, as this use was appropriated by early printers. One major question remains, however: Was the disappearance of abbreviations an independent process or did the abbreviations disappear as a part of the elimination of spelling variations? In the next section, we compare how abbreviation density corresponds with spelling variation.

4.4 Spelling variation

The final topic we examine in this study is how the reduction in the number of abbreviations corresponds with reduction in the number of English spellings. The main research question here is whether we are dealing with two sides of the same phenomenon or two separate phenomena. This is important, as our theoretical framework presumes gradual elimination of variation in the course of the fifteenth century towards proto-standard English. Moreover, statistical approaches require independent populations to establish that the observed reduction in abbreviation (sections 4.1 and 4.2) is not simply a byproduct of an overall drop in variation at this point. A secondary research question is to determine whether the spelling variants data are distributed as in a complex system.

In order not to confuse the two figures into the same data, we performed the counts twice, both including and excluding the abbreviations. The reason for this approach was that on the one hand several scholars argue that the abbreviated

spellings should be treated as legitimate spelling variants of their own (Lass 2004; Driscoll 2009; Rogos 2012). On the other hand, we would not be dealing with two independent samples if reduction in abbreviations showed as reduction of spelling variants or vice versa, which might potentially confound the results.

When making the decision, we followed the following "rules". Words like eu⁹ech, eu⁹eche, eu⁹iche, eu⁹yche 'every' have four spelling variants with abbreviation and four without abbreviation, because the variation is not caused by the abbreviation. Words like grace, g^ace 'grace' have two variants including abbreviation, but only one excluding it. Words like hundred, hudrid ('hu(n)drid') have two variants including abbreviation and two variants excluding abbreviation, because even though there is both an abbreviated and an expanded form, there is also variation of the graphs and <e>. Finally, the words *hert*, *herte* were counted as having two including abbreviation and two variants excluding abbreviation, even though neither of the forms is abbreviated. Capital letters were not counted as spelling variants, but the doubling of letters (ffrom, from) was counted even when it represents a larger initial or a littera notabilior. The reason for this is that it can in other instances be a significant spelling variant, and we try to avoid exercising too much editorial interpretation. To keep morphological variation separate from orthographic variation, the singular and plural of each word were counted as separate words (age and ages = two different words) as were different verb forms. The results are shown in Figures 9.12, 9.13 and 9.14.²⁶

Regardless of whether one counts abbreviation characters as variants or not, the results show a much slower reduction for spelling variation than for abbreviations. Figures 9.12 and 9.13 display the average number of spelling variants per words with two or more tokens, excluding abbreviation and including abbreviation. Figure 9.14 illustrates the same data using a different method of calculation: it shows the number of words in each manuscript that have three or more spelling variants, both including and excluding variation. These tables reveal that, while the proportions may vary, the relative amount of spelling variation stays the same between the various witnesses.

There is no decrease from the early manuscripts to the mid-fifteenth century. It could even be argued that there is a slight increase, as the two early scribes, H and the "Polychronicon Scribe", spell very consistently. The scribe of H in particular is the most consistent speller before early printed books, especially excluding abbreviations. The impression of a slight increase is augmented by the fact that one of the scribes in the 1400–1425 period, F, stands out as using more spelling

²⁶ The Cotton manuscript, C, was not included in this part of the study, because it was considered too damaged to provide reliable evidence for spelling variation.



Figure 9.12: Average number of spelling variants excluding abbreviations.

variants by three out of four ways for calculating spelling variation, and as using the second highest number of variants (see Figure 9.14) after scribe A. There is, however, a slight drop from manuscript to print.

The secondary research question is to find out whether spelling variants of individual words are distributed as predicted by the linguistics of speech. To recapitulate, if this was the case, one would expect to find "a few very common variants, many uncommon variants" (Kretzschmar and Stenroos 2012). More specifically, we would expect the distribution to follow the so-called A-curve. When applied to scribal data, one would expect "that clerks wrote a majority form for a given feature, a substantial minority form, and then a tail-off of several minority forms at very low rates" (Wright 2013: 64). Moreover, we would expect some reduction towards a proto-standard form.

The results do reveal an A-curve distribution for some words, but there are major differences in the number of spelling variants for individual words (see Table 9.8). Some words are spelled very consistently even in the earliest manuscripts, some have varied forms in early printed books. A good example of consistency is how the preposition 'of' is spelled entirely consistently already in the earliest manuscript witnesses (see also 'the' in Table 9.5).²⁷ To some extent this is

²⁷ The spelling with a double in MS G almost certainly indicates a littera notabilior.



Figure 9.13: Average number of spelling variants including abbreviations.



Figure 9.14: Number of words with three or more spellings.

		OF	FROM							
Early manuscripts	W	of 163	fram 3	5 frā	<u>a</u> 4	f ^a m 2	ffram 2	fra91		
	т	of 166	frō 39	fr	om 2	ffrō 1	ffrom 1	fro 1		
1400-1425	А	of 157	from 2	9 fr	am 9	fr <u>o</u> 3	ffram 1	ffrom 1	fro 1	
	ш	of 162	from 2	t fri	0 10	fram 5	frā 3	fro 1	ffrom 1	ffrō 1
	_	of 160	from 3	5 fri	0 3	fro 2	fram 1			
1425-1475	в	of 165	from 3	1) El	rom 2	frō 2	fro 1			
	ŋ	of 158 Of	f 1 from 2	t fr	0 18	frō 1				
	Т	of 161	from 2	5 fr	0 16	frō 1				
Printed books	Сах	of 170	fro 32	fr	om 10	ffrom 1	ffro 1			
	Wor	of 169	fro 33	fr	om 11					
	Tre	of 161	fro 32	frı	om 7	frome 3				

Table 9.8: Spelling variants for 'of' and 'from'.

what one would expect, as some studies suggest that very frequent words are likely to have fewer variants (Evans 2012; Shute 2017b: 113). However, there are also fairly frequent function words that do exhibit variation, which follows the expected A-curve pattern. A good example of these is the preposition 'from', as it still has two variant spellings used by all printers *fro* and *from*, and an additional variant *frome* used by Treveris. On top of that, the word is often abbreviated with a macron or spelled with a double initial. If one includes these in the count, the distribution of variants resembles the characteristic A-curve found in complex systems still in early printed books. It is thus clear that reduction of variation does not happen at the same rate for these two words.

While words can retain spelling variants, many individual spellings show developments towards proto-standard forms. Table 9.9 illustrates two cases which go through qualitative changes towards forms that are closer to eventually standard variants: 'that' and 'to be', 3rd person plural. The scribal forms for 'that' show how thorn is replaced by , starting in the mid-fifteenth century (G and T). As we know, this was a gradual change subject to variation, which could be conditioned by such factors as text type or recipient (see chapters by Hernández-Campoy and Gordon in the current volume). The date of the change in our current data corresponds roughly to what one would expect. What is interesting for

		THAT		TO BE, 3rd person pl	
Early manuscripts	М	þ ^t 37	þat 23	buþ 9	
	Н	þ ^t 52	þat 12	beþ 9	
1400–1425	А	þat 54	þ ^t 2	beeþ 7	beþ 1
	F	þat 34	þ ^t 23	beþ 9	
	J	þat 57	þ ^t 1	beeþ 8	beþ 1
1425–1475	В	þat 60	þ ^t 2	biþ 6	beþ 3
	G	that 54	þ ^t 1	beth 5	ben 3
	Т	that 54	þat 2	ben 9	beth 1
Printed books	Cax	that 65		ben 9	
	Wor	that 59	y ^t 6	ben 8	
	Tre	that 57	y ^t 4	ben 7	

Table 9.9: Spelling variants for 'that' and the verb 'to be'.

our research questions is that a major drop in abbreviation frequencies happens earlier than change from thorn to in all classes of words.

Spellings with thorn lingered longer in function words (see, e.g., Gordon: 206). In the present data thorn is replaced by , word-initially, in 'that' and, word-finally, in forms of 'to be'. However, the drop in frequency of the abbreviated spelling occurs right after 1400. In the two earliest manuscripts (M and H), the abbreviated spelling, p^t , is the major variant. In later manuscripts (A, F, J, B, G), it is used as an occasional minor variant. These manuscripts are the ones in which margin justification is highly significant, so it is very likely that the scribes use it when they need to save space. The scribe of T in the third quarter of the fifteenth century does not use the abbreviated form, but spells the word twice with thorn as a minor variant to the Present-Day English spelling *that*. The abbreviated form, however, becomes part of the SPSA leading to the familiar situation in which y^t is an occasional variant for *that*, for example, when right-margin justification demands it.

The forms used for the third-person plural of 'to be' show an even clearer development from regional to proto-standard. The earliest scribe (M) spells the word *bub*, which is likely close to Trevisa's own form (cf. Waldron 1991). It gets replaced by the Southern *-th* paradigm and eventually the Midland paradigm *-en*. This is proto-standard English and the same form as observed by Wright (2013) for London Bridge clerks. It is not the form that eventually became standard (*are*). There is also the change from thorn to *<*th>. The Midland spellings first appear in G in the mid-fifteenth century as a minor variant (*beth 5, ben 3*). In T, they are already the major variant (*ben 9, beth* 1), and in the printed books the sole variant. Based on this evidence, it therefore appears that there is both quantitative and qualitative development towards proto-standard English in the latter half of the fifteenth century. The change progresses at a different rate for different words, as some words still retain variation in printed books.

The overall conclusion is that abbreviation decreases faster than spelling variation, and thus these two processes are separate. Figures 9.15 (a) and (b) give an overall impression of the rate of change for the reduction of spelling variation and abbreviation. If anything there is a slight increase in variation from the earliest English manuscripts to the mid-fifteenth century. The fourteenth-century professional scribes wrote in a consistently spelled local dialect, and the number of variants used by an individual scribe was not bigger than for scribes writing in the mid- or late fifteenth century – we did not find quantitative evidence of reduction of variation before a slight drop towards the early printed books.

It is possible that the increase from early manuscripts to ones copied between 1400–1425 can be explained by the differences between regional variants and a London melting pot. The earliest manuscripts of the *Polychronicon* are



Figures 9.15 (a) and (b): Mean abbreviations and mean spelling variants. S-curves for the reduction of spelling variants, reduction of abbreviations in Middle English.

copied in a Gloucestershire dialect close to Trevisa's own. Moreover, as Waldron (1991: 67) points out, Trevisa and his patron Sir Thomas Berkeley appear to have hoped for the work to contribute to lay education, and the example of King Alfred's educational program is invoked in Trevisa's *Dialogue between a Lord and a Clerk*, which comments on the translation work. Perhaps an educational aim would have contributed some uniformity to spelling. The later copies, on the other hand, are copied in varieties closer to London, which was a major commercial hub, in which "both provincial sellers and foreign buyers interacted" (Wright 2013: 68). As a result, its language was "an urban amalgam drawing on non-adjacent dialects" (Kitson 2004: 71). If this is true, the pool of variants available to the early fifteenth-century London scribes would be slightly more extensive than that available to provincial ones. But as our data are fairly limited, it is impossible to conclude this with certainty.

5 Discussion and conclusions

This chapter has presented a quantitative look at the reduction of abbreviation with the arrival of standardisation, and provided outlines on the rate of disappearance for the abbreviation and suspension system. We uncovered a number of promising results that need to be examined with other data in different genres. It remains to be seen how many of our discoveries are specific to high-end manuscripts such as copies of the *Polychronicon*, and how many are typical of abbreviations in other types of texts.

The data from the *Polychronicon* show that in English the reduction happens between the late fourteenth and mid-fifteenth century, while the scribes continue to use abbreviations in Latin until the third quarter of the fifteenth century. There is, however, a great deal of variation in the rate with which individual abbreviation types disappear. In the early copies, small function words strongly contribute to the raw abbreviation density. Moreover, a few popular types of abbreviation, the macron and the hook in particular, could be applied to many different types of words and their overall frequency fluctuates depending on individual scribal preferences. As soon as scribes begin to spell small function words and stop using the hook, the number of abbreviations decreases rapidly. Brevigraphs, on the other hand, remain more constant, showing a slow and steady decrease. They are always more common in Latin and are still used by the later printers.

The quantitative approach also reveals that although the trend is towards less use, the process was not always one of decrease. There is a general increase in the abbreviation density from Caxton to the two later printers, which may be related, on the one hand, to the small but commonly used set of abbreviations favoured by the printers, here called the SPSA. The SPSA forms must have been part of the pool of variants in proto-standard English in other works at the time, as they are all very common types, but the later manuscript copies of the English *Polychronicon* do not contain these abbreviations. According to the linguistics of speech approach, variation consists of change of frequency of items, so this is likely to be a shift in the "ratios of a given feature" in favour of "a feature found in a majority elsewhere" (cf. Wright 2018: 348) rather than disappearance as such. As our 3000-word samples are fairly small, it is also possible that they exist as low frequency minor variants and a bigger sample from these manuscripts would uncover a handful of tokens. Nevertheless, the present results show that the SPSA were not used by professional scribes responsible for mid- to late fifteenth-century manuscript copies of the Polychronicon.

Our study also found one conditioning factor for abbreviations from the mid-fifteenth century to printing. Abbreviations tend to be more common closer to the right margin. When the abbreviation and suspension system began to fall out of use, it remained in use as an alternative when the scribes were pressed for space and needed to produce a neat right-margin justification. Yet our study reveals that even though the majority of abbreviations occur near the right-hand margin, the difference is not statistically significant for either the early period or the printed editions. This means that, even though abbreviations were partly motivated by the need to save space, there were also other uses for the abbreviation system.

Finally, our study suggests that although both abbreviation and spelling variation eventually disappeared, the processes were separate. The loss of abbreviations happens earlier than the loss of spelling variation. Our data show only a slight reduction in the number of variants towards early printed books; the idea of one word, one spelling was still not the norm. Qualitatively, we did uncover changes towards proto-standard spellings, even if, from a wider perspective, they can be explained by shifting ratios in a complex repertoire of forms available at the time. It is not completely certain to which events this corresponds. Multilingual systems of accounting were still in use at the time (cf. Wright 2018: 352; Alcolado Carnicero 2013: 217). The writing support in all of the manuscripts is parchment. One possibility is the shift from Anglicana to Secretary scripts, as in our data manuscripts copied in Secretary have a lower abbreviation density than ones copied in Anglicana. Whatever the reason, it would appear that professional scribes, working on *de luxe* commissions like these *Polychronicon* copies, were mainly expanding their function words by the mid-fifteenth century, whereas their language only shows a shift towards proto-standard spellings after the mid-fifteenth century.

This chapter has described what happened to abbreviations in a single work from the late fourteenth to the early sixteenth centuries. There are, of course, many things it did not cover. These include some of the more multilingual sources of the time. The medieval abbreviation and suspension system has been linked to hiding morphological endings in a multilingual society. Abbreviations could function as visual diamorphs, that is, language independent elements which can potentially be expanded in several languages, including Latin, English and Anglo-Norman French. This usage is mentioned already by Hector (1958: 37), who notes that English proper names in Latin documents could be "terminated by a mark of suspension to preserve the fiction that they were declinable Latin words". The phenomenon is investigated further by Wright, who notes it has the effect of suppressing morphological endings and highlighting stems in mixed-language writing (Wright 2011; ter Horst and Stam 2018: 223–242). The functions of abbreviations as visual diamorphs in highly mixed-language data remain yet to be described quantitatively.

Another feature which should be subject to more quantitative work is whether there is a gradual adoption of certain features into the vernacular. This process is known as "Frenchisisation" in French philological traditions since the 1990s. There are also other interesting ideas and results; for example, according to Hasenohr (2002: 88–90; see also Camps 2016: ccl) the application of the Latin abbreviations to French seems to originate in the Anglo-Norman speaking territories of the Angevin empire and from there spread to the rest of France. Even though French and English scholarly traditions do not always interact much, manuscript abbreviations are an area in which much could be gained from such interaction.

To conclude, much of the elaborate system of abbreviations that developed over centuries of handwritten book production gradually fell out of use, which indeed largely happened during the period studied in the present chapter. Nevertheless some parts of it continued in use and have become parts of standardised language. Standardisation is applied, for example, in the form of a rule, which states that "contractions, where the last letter of the abbreviation is also the last letter of the word, should not be followed by a point, whereas suspensions should. [...] One should thus write 'Mr', 'Mrs' [...] but 'Feb.', 'Rev.' etc." (Driscoll 2002). Moreover, some abbreviations which were used in the Middle Ages continue to be used in Present-Day English. Two especially common ones are & and etc. The latter has even expanded from Latin manuscript culture to spoken language. Abbreviated Latin was especially important for bibliographic references, and these are among the ones that have survived. It is telling that abbreviations such as cf., et al., e.g., ibid. or viz. are still used in the academic register today, even if some of them are stylistically old-fashioned. Thus even though most of the medieval abbreviation and suspension system is long gone, parts of it are still with us and show no signs of imminent disappearance.

References

- Alcolado Carnicero, José Miguel. 2013. Social networks and mixed-language business writing: Latin/French/English in the wardens' accounts of the Mercers' Company of London, 1390–1464. University of Castilla-La Mancha: PhD thesis.
- Beal, Jane. 2012. *John Trevisa and the English Polychronicon* (Medieval and Renaissance Texts and Studies 437). Tempe, AZ: ACMRS.
- Benskin, Michael. 2004. Chancery standard. In Christian Kay, Carole Hough, & Irené
 Wotherspoon (eds.), *New perspectives on English historical linguistics*, vol. 2, 1–40.
 Amsterdam: John Benjamins.
- BMC 11 = Hellinga, Lotte & George D. Painter. 2007. Catalogue of books printed in the XVth century now in the British Library. BMC Pt. 11: England. 't Goy-Houten: HES & De Graaf.
- Bozzolo, Carla, Dominique Coq, Denis Muzerelle & Ezio Ornato. 1990. Les abréviations dans les livres liturgiques du XVe siècle: pratique et théorie. In Manuel C. Díaz y Díaz (ed.)
 Actas del VIII Coloquio Internacional de Paleografia Latina (Estudios y Ensayos 6), 17–27. Madrid: Joyas Bibliograficas.
- Camps, Jean-Baptiste. 2016. La Chanson d'Otinel Édition complète du corpus manuscrit et prolégomènes à l'édition critique. University of Paris Sorbonne: PhD thesis.
- Cappelli, Adriano. 1990 [1899]. *Lexicon abbreviaturarum dizionario di abbreviature latine ed italiane*. Milano: Hoepli.
- Careri, Maria, Christine de Saint-Pol Ruby & Ian Short. 2011. *Livres et écritures en français et en occitan au XIIe siècle: catalogue illustré.* Rome: Viella.

- Careri, Maria, Genevieve Hasenohr, Françoise Féry-Hue, Françoise Gasparri, Gilette Labory, Sylvie Lefèvre, Anne-Françoise Leurquin & Christine Ruby. 2001. *Album de manuscrits français du XIIIe siècle*. Rome: Viella.
- Carroll, Ruth, Matti Peikola, Hanna Salmi, Mari-Liisa Varila, Janne Skaffari & Risto Hiltunen. 2013. Pragmatics on the page: Visual text in late medieval English books. *European Journal of English Studies* 17. 54–71.
- Chassant, L-A. 1970 [1845]. Dictionnaire des abréviations latines et francaises usitées dans les inscriptions lapidaires et métalliques, les manuscrits et les chartes du moyen âge. Hildesheim: Georg Olms Verlag.
- Cummings, James. 2009. Converting Saint Paul: A new TEI P5 edition of The Conversion of Saint Paul using stand-off methodology. *Literary and Linguistic Computing* 24(3). 307–317.
- Da Rold, Orietta. 2011. Materials. In Alexandra Gillespie & Daniel Wakelin (eds.), *The* production of books in England, 1350–1500, 12–33. Cambridge: Cambridge University Press.
- Deumert, Ana & Wim Vandenbussche (eds.), 2003. *Germanic standardizations: Past to present*. Amsterdam & Philadelphia: Benjamins.
- Doyle, A. I. & M. B. Parkes. 1978. The production of copies of the *Canterbury Tales* and the *Confessio Amantis* in the early fifteenth century. In M. B. Parkes & Andrew G. Watson (eds.), *Medieval scribes, manuscripts & libraries: Essays presented to N. R. Ker.* London: Scolar Press.
- Driscoll, Matthew J. 2002. Stray thoughts on abbreviations in some modern European languages. In Jonna Louis-Jensen & Ragnheiður Mósesdóttir (eds.), *Grace-notes played for Michael Chesnutt on the occasion of his 60th birthday, 18 September 2002.* Copenhagen: Det Arnamagnæanske Institut. http://www.driscoll.dk/docs/thoughts.html (accessed 7 January 2019).
- Driscoll, Matthew. 2009. Marking up abbreviations in Old Norse-Icelandic manuscripts. In Maria Grazia Saibene & Marina Buzzoni (eds.), *Medieval texts – contemporary media: The art and science of editing in the digital age*, 13–34. Pavia: Ibis.
- Dutschke, C. W. 1989. *Guide to medieval and renaissance manuscripts in the Huntington library*, vol 2. San Marino, CA: Huntington Library.
- Evans, Melanie. 2012. A sociolinguistics of early modern spelling? A case study of Queen Elizabeth I. In Jukka Tyrkkö, Matti Kilpiö, Terttu Nevalainen & Matti Rissanen (eds.), *Outposts of historical linguistics: From the Helsinki Corpus to a proliferation of resources* (Studies in Variation, Contacts and Change in English 10). Helsinki: VARIENG. http:// www.helsinki.fi/varieng/journal/volumes/10/evans (accessed 7 January 2019).
- Fisher, John H. 1977. Chancery and the emergence of standard written English in the fifteenth century. *Speculum* 52. 870–899.
- Fisher, John H. 1979. Chancery standard and modern written English. *Journal of the Society of Archivists* 6. 136–144.
- Fisher, John H. 1996. *The emergence of Standard English*. Lexington: University Press of Kentucky.
- Given-Wilson, Chris. 2004. *Chronicles: The writing of history in medieval England*. Hambledon & London & New York: A&C Black.
- Hasenohr, Geneviève. 2002. Écrire en latin, écrire en roman: réflexions sur la pratique des abréviations dans les manuscrits français des XIIe et XIIIe siècles. In Michel Banniard (ed.), *Langages et peuples d'Europe: cristallisation des identités romanes et*

germaniques (VIIe–XIe siècle), 79–110. Toulouse-Conques: CNRS, Université de Toulouse-Le Mirail.

Hector, Leonard Charles. 1958. *The handwriting of English documents*. Ilkley: Scolar Press. HIMANIS 2018. https://www.himanis.org/ (accessed 31 July 2018).

- Honkapohja, Alpo. 2013. Manuscript abbreviations in Latin and English: History, typologies and how to tackle them in encoding. In Anneli Meurman-Solin & Jukka Tyrkkö (eds.), *Principles and practices for the digital editing and annotation of diachronic data* (Studies in Variation, Contacts and Change in English 14). Helsinki: VARIENG. http://www.hel sinki.fi/varieng/series/volumes/14/honkapohja/ (accessed 7 January 2019).
- Honkapohja, Alpo. 2017. *Alchemy, medicine, and commercial book production* (Texts and Transitions 9). Turnhout: Brepols.
- Honkapohja, Alpo. 2018. 'Latin in Recipes?' A corpus approach to scribal abbreviations in 15th-century medical manuscripts. In Päivi Pahta, Janne Skaffari & Laura Wright (eds.), *Multilingual practices in language history: English and beyond*, 243–271. Boston & Berlin: De Gruyter.
- Horst, Tom ter. 2017. *Codeswitching in the Irish-Latin Leabhar Breac: Mediæval homiletic culture*. University of Utrecht: PhD thesis.
- Horst, Tom ter & Nike Stam. 2018. Visual diamorphs: The importance of language neutrality in code-switching from medieval Ireland. In Päivi Pahta, Janne Skaffari & Laura Wright (eds.), *Multilingual practices in language history*, 223–242. Boston & Berlin: De Gruyter.
- Ingham, Richard. 2012. *The Transmission of Anglo-Norman: Language history and language acquisition*. Amsterdam: Benjamins.
- Kestemont, Mike. 2015. A computational analysis of the scribal profiles in two of the oldest manuscripts of Hadewijch's letters. *Scriptorium* 69. 159–175.
- Kitson, Peter. 2004. On margins of error in placing Old English literary dialects. In Marina Dossena & Roger Lass (eds.), *Methods and data in English historical dialectology* (Linguistic Insights: Studies in Language and Communication 16), 219–239. Bern: Peter Lang.
- Kopaczyk, Joanna. 2011. A V or not a V? Transcribing abbreviations in seventeen manuscripts of the 'Man of Law's Tale' for a digital edition. In Jacob Thaisen & Hanna Rutkowska (eds.), Scribes, printers, and the accidentals of their texts (Studies in English Medieval Language and Literature 33), 91–106. Frankfurt: Peter Lang.
- Kretzschmar, William A., Jr. 2009. *The Lingsuistics of Speech*. Cambridge: Cambridge University Press.
- Kretzschmar, William A., Jr. 2015. *Language and Complex Systems*. Cambridge: Cambridge University Press.
- Kretzschmar, William A., Jr. & Merja Stenroos. 2012. Evidence from surveys and atlases in the history of the English language. In Terttu Nevalainen & Elizabeth Closs Traugott (eds.), *The Oxford handbook of the history of English*. Oxford & New York: Oxford University Press.
- Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1(3). 199–244.
- Kytö, Merja, Peter Grund & Terry Walker. 2011. Testifying to language and life in early modern England: Including CD-ROM: An electronic text edition of depositions 1560–1760 (ETED). Philadelphia: Benjamins.
- Labov, William. 1994. Principles of linguistic change: Internal factors. Oxford: Blackwell.

- Laing, Margaret. 2013. *A linguistic atlas of early Middle English, 1150–1325*, version 3.2. Edinburgh: The University of Edinburgh. http://www.lel.ed.ac.uk/ihd/laeme2/laeme2. html (accessed 7 January 2019).
- LAOS = Williamson, Keith. 2008. A linguistic atlas of Older Scots, Phase 1: 1380–1500. The University of Edinburgh. http://www.lel.ed.ac.uk/ihd/laos1/laos1.html (accessed 7 January 2019).
- Lass, Roger. 2004. Ut custodiant litteras: Editions, corpora and witnesshood. In Marina Dossena & Roger Lass (eds.), *Methods and data in English historical dialectology* (Linguistic Insights: Studies in Language and Communication 16), 21–48. Bern: Peter Lang.
- Lindsay, Wallace Martin. 1915. Notae latinae: An account of abbreviation in Latin MSS of the early minuscule period. Cambridge: Cambridge University Press.
- Liira, Aino. 2020. Paratextuality in manuscript and print: Verbal and visual presentation of the Middle English *Polychronicon*. (Annales Universitatis Turkuensis B512). University of Turku: PhD thesis. http://urn.fi/URN:ISBN:978-951-29-8058-1 (accessed 22 May 2020).
- Lyall, Roderick J. 1989. Materials: The paper revolution. In Jeremy Griffiths & Derek Pearsall (eds.), *Book production and publishing in Britain, 1375–1475*, 11–30. Cambridge: Cambridge University Press.
- Matheson, Lister M. 1984. Historical prose. In A. S. G. Edwards (ed.), *Middle English prose: A critical guide to major authors and genres*, 209–248. New Brunswick, N.J: Rutgers University Press.
- McIntosh, Angus, M. L. Samuels & Michael Benskin. 1986. *A linguistic atlas of late mediaeval English*, vol 1. Aberdeen: Aberdeen University Press.
- MEG-C = Stenroos, Merja, Martti M\u00e4kinen, Simon Horobin & Jeremy Smith. 2011. The Middle English Grammar Corpus, version 2011.1. University of Stavanger. https://www.uis.no/ research/history-languages-and-literature/the-mest-programme/the-middle-englishgrammar-corpus-meg-c/ (accessed 7 January 2019).
- MELD = Stenroos, Merja, Kjetil, V. Thengs & Geir Bergstrøm (2017–). A Corpus of Middle English Local Documents, version 2017.1. University of Stavanger. www.uis.no/meld (accessed 7 January 2019).
- MENOTA = Medieval Nordic Text Archive. http://www.menota.org/EN_forside.xhtml (accessed 7 January 2019).
- Mooney, Linne. 2006. Chaucer's scribe. Speculum 81. 97-138.
- Mooney, Linne, Simon Horobin, & Estelle Stubbs. 2011. Late medieval English scribes. https://www.medievalscribes.com (accessed 7 January 2019).
- Nevalainen, Terttu. 2015. Descriptive adequacy of the S-curve model in diachronic studies of language change. In Christina Sanchez-Stockhammer (ed.), *Can we predict linguistic change?* (Studies in Variation, Contacts and Change in English 16). Helsinki: VARIENG.
- Nevalainen, Terttu & Ingrid Tieken-Boon van Ostade. 2006. Standardisation. In Richard Hogg & David Denison (eds.), *A history of the English language*, 271–311. Cambridge: Cambridge University Press.
- OxGarage. http://www.tei-c.org/oxgarage/ (accessed 7 January 2019).
- Parkes, M. B. 2008 [1969]. *English cursive book hands 1250–1500*. Aldershot/Burlington,VT: Ashgate.
- Petti, Anthony G. 1977. English literary hands from Chaucer to Dryden. London: E. Arnold.
- Ramsey, Roy Vance. 1982. The *Hengwrt* and *Ellesmere* manuscripts of the *Canterbury Tales*: Different scribes. *Studies in Bibliography* 35. 133–154.

Roberts, Jane. 2005. Guide to scripts used in English writing up to 1500. London: British Library.

- Roberts, Jane. 2011. On giving Scribe B a name and a clutch of London manuscripts from c. 1400. *Medium Aevum* 80(20). 247–270.
- Robinson, Pamela. 2014. Materials: Paper and type. In Vincent Gillespie & Susan Powell (eds.), *Companion to the early printed book in Britain, 1476–1558*, 61–74. Cambridge: Boydell & Brewer.
- Rogos, Justyna. 2012. Isles of systemacity in the sea of prodigality? Non-alphabetic elements in manuscripts of Chaucer's 'Man of Law's Tale'. https://www.isle-linguistics.org/assets/ content/documents/hogg/rogos2012.pdf (accessed 7 January 2019).
- Samuels, Michael. 1983 [1963]. Some applications of Middle English dialectology. In Angus McIntosh, Michael Samuels & Margaret Laing (eds.), *Middle English Dialectology: Essays on some principles and problems*, 81–94. Aberdeen: Aberdeen University Press.
- Samuels, Michael. 1983. The scribe of the *Hengwrt* and *Ellesmere* manuscripts of the *Canterbury Tales*. *Studies in the Age of Chaucer* 5. 49–65.
- Shepherd, Stephen. 1999. John Trevisa, Dialogue between the Lord and the Clerk on translation (extract) and Epistle to Thomas, Lord Berkeley, on the translation of Higden's Polychronicon. In Jocelyn Wogan-Browne, Nicholas Watson, Andrew Taylor & Ruth Evans (eds.), The idea of the vernacular: An anthology of Middle English literary theory, 1280–1520, 130–138. University Park, PA: The Pennsylvania State University Press.
- Shute, Rosie. 2017a. Pressed for space: The effects of justification and the printing process on fifteenth-century orthography. *English Studies* 98(3). 262–282.
- Shute, Rosie. 2017b. A quantitative study of spelling variation in William Caxton's printed texts. University of Sheffield: PhD thesis.
- Smith, Daisy. 2019. The predictability of {S} abbreviation in Older Scots manuscripts according to stem-final *littera*. In Rhona Alcorn, Joanna Kopaczyk & Benjamin Molineaux (eds.), *Historical dialoctology in the digital age*. Edinburgh: Edinburgh University Press.
- Stam, Nike. 2017. *A typology of code-switching in the Commentary to the* Félire Óengusso. University of Utrecht: PhD thesis.
- Stutzmann, Dominique. 2014. Conjuguer diplomatique, paléographie et édition électronique : les mutations du XIIe siècle et la datation des écritures par le profil scribal collectif. In Antonella Ambrosio, Sébastien Barret & Georg Vogeler (eds.), *Digital diplomatics: The computer as a tool for the diplomatist?* (Archiv für Diplomatik, Schriftgeschichte, Siegelund Wappenkunde 14), 271–290. Vienna/Cologne: Böhlau.
- Stutzmann, Dominique, Christopher Kermorvant, Enrique Vidal, Sukalpa Chanda, Sébastien Hamel, Joan Puigcerver Pérez, Lambert Schomaker & Alejandro H. Toselli. 2018.
 Handwritten text recognition, keyword indexing, and plain text search in medieval manuscripts. Paper presented at Digital Humanities 2018 conference, Mexico City, 26–29 June. https://dh2018.adho.org/handwritten-text-recognition-keyword-indexingand-plain-text-search-in-medieval-manuscripts/ (accessed 7 January 2019).
- TEI consortium. 2013. TEI P5: Guidelines for electronic text encoding and interchange. http://www.tei-c.org/Vault/P5/2.4.0/doc/tei-p5-doc/en/html/ (accessed 7 January 2019).
- Thaisen, Jacob. 2011. Adam Pinkhurst's short and long forms. In Jacob Thaisen & Hanna Rutkowska (eds.), *Scribes, printers, and the accidentals of their texts* (Studies in English Medieval Language and Literature 33), 73–90. Frankfurt: Peter Lang.
- Tristram, Hildegard. 1997. Latin and Latin learning in the Táin Bó Cuailnge. *Zeitschrift für celtische Philologie* 49–50. 847–877.

- Tyrkkö, Jukka. 2017. New methods of bringing image data into historical linguistics: A case study with medical writing 1500–1700. *Studia Neophilologica* 89, suppl. 1. 90–108.
- Varila, Mari-Liisa. 2016. In search of textual boundaries: A case study on the transmission of scientific writing in 16th-century England (Anglicana Turkuensia 31). University of Turku: PhD thesis.
- Waldron, Ronald. 1991. Dialect aspects of manuscripts of Trevisa's translation of the Polychronicon. In Felicity Riddy (ed.), Regionalism in late medieval manuscripts and texts: Essays celebrating the publication of A Linguistic Atlas of Late Mediaeval English. Cambridge: D. S. Brewer.
- Waldron, Ronald. 2004. John Trevisa's translation of the 'Polychronicon' of Ranulph Higden, Book VI: An edititon based on British Library MS Cotton Tiberius D.VII (Middle English Texts 35). Heidelberg: Winter.
- Warner, Lawrence. 2015. "Scribes Misattributed: Hoccleve and Pinkhurst". *Studies in the Age of Chaucer* 37: 55–99.
- Warner, Lawrence. 2018. *Chaucer's Scribes. London Textual Production 1384-1432*. Cambridge: Cambridge University Press.
- Wright, Laura. 2000. *The development of Standard English, 1300–1800*. Cambridge: Cambridge University Press.
- Wright, Laura. 2002. Standard English and the lexicon: Why so many different spellings? In Mari C. Jones & Edith Esch (eds.), Language change: The interplay of internal, external and extra-linguistic factors, 181–200. Berlin & New York: Mouton de Gruyter.
- Wright, Laura. 2011. On variation in medieval mixed-language business writing. In Herbert Schendl & Laura Wright (eds.), *Code-switching in early English*, 191–218. Göttingen: De Gruyter.
- Wright, Laura. 2013. The contact origins of Standard English. In Daniel Schreier & Marianne Hundt (eds.), *English as a contact language* (Studies in English Language), 58–74.
 Cambridge: Cambridge University Press.
- Wright, Laura. 2018. A multilingual approach to the history of Standard English. In Päivi Pahta, Janne Skaffari & Laura Wright (eds.), *Multilingual practices in language history: English and beyond*, 339–358. Boston & Berlin: De Gruyter.