This is an Accepted Manuscript of an article published by Taylor & Francis in Studia Neophilologica *on 2 October 2013, available online:* http://dx.doi.org/10.1080/00393274.2013.834107.

Studia Neophilologica 2014, Volume 86, Supplement 1: 157–170

Graphetic variation within one scribal hand as evidence on manuscript production

Mari-Liisa Varila

Introduction

Palaeography is a field fundamentally important to manuscript studies, offering evidence for example on the date, provenance and production of a manuscript or a group of manuscripts, or information on certain scribes and their habits. Scribal habits have been studied widely in the field of palaeography. However, the results of palaeographical analyses often influence philological and textual research as well: for example, for the purposes of a scholarly edition it is usually relevant to find out how many scribes worked on the manuscript(s) and where the shifts take place. In 1997, Jeremy Smith noted that 'one of the most inspiring developments over the last few years has been the way in which the disciplines of philology and paleography have been brought into closer collaboration' (1997: 133–134). Such collaboration, I believe, is worthwhile for both fields. As the textual and the physical are intertwined in manuscripts, both textual and physical evidence should be taken into account when studying manuscripts.

This pilot study explores the potential outcomes of analysing graphetic variation, thereby testing a method drawing heavily from palaeography but potentially applicable to philological and textual research on a more general level. My primary material comes from two mid-sixteenth-century manuscripts (New Haven, Yale University, Medical Library MS 45 and Beinecke Rare Book and Manuscript Library MS 558). The selected sections were copied by the same scribe and contain texts that are partly shared. I analyse graphetic variation by looking at the distribution of word-final <s> forms. I will first briefly discuss some previous studies on orthographic variation and scribal behaviour, then introduce my material and methods, and finally move on to the results and possible applications of the method.

Graphetic variation and scribal practices

Peter Robinson and Elizabeth Solopova (2006: 2) distinguish four levels of manuscript transcription: *graphic* (every mark in the manuscript is represented in the transcript); *graphetic* (every letter-type is represented); *graphemic* (every spelling is retained, but letterforms are not distinguished); and *regularised* (spellings are 'normalised' according to editorial principles). They note that transcriptions are often mixtures of these categories (2006: 2). To put it briefly, a *grapheme* is an abstraction that is realised in the form of a *graphete*: for example, the grapheme /a/ may be realised as <a> or <*a*>. On a graphemic

level, there is thus no distinction between the forms, and one graphete, or variant, may be selected to stand for the grapheme in the transcription. In a graphetic transcription, however, the graphetes would be distinguished.[1] When analysing individual handwriting, the analysis may be even more fine-grained. Tom Davis, addressing the question of handwriting identification in forensic science and palaeography, describes the hierarchy as follows:

> The grapheme /a/ is the letter considered independently of any particular realization of it. An allograph is an accepted version of that grapheme: 'a' and '*a*' are allographs of /a/. An idiograph is the way (or one of the ways) in which a given writer habitually writes /a/. A graph is a unique instance of /a/, as it appears on a particular page. (Davis 2007: 255)

Scribal identification is a highly relevant question in palaeography and manuscript studies. A well-known example is the attribution of the most famous *Canterbury Tales* manuscripts to Adam Pinkhurst, whose hand (c. 1400) has also been identified in various other documentary and literary manuscripts (see e.g. Mooney 2006). When comparing scribal hands, graphetic variation is one factor to consider. If a scribal hand systematically employs a certain type of a grapheme (e.g. one-compartment <*a*>), and the hand it is compared to only contains tokens of another type (two-compartment <a>), this could be used as one criterion for distinguishing between the two hands.

Less than forty years ago, Angus McIntosh noted that 'no systematic attempt has yet been made to characterise the variations of mediæval lettershapes and the like in terms of discrete categories such as could then be handled by some form of numerical taxonomy' (1975: 223). McIntosh suggested compiling graphetic profiles of scribes to complement the information gained from linguistic profiling (1975: 223). Since then, progress has been made regarding research into scribal habits, for example in the form of the ongoing *Late Medieval English Scribes* project (Mooney et al.). While no single questionnaire is used to create graphetic profiles across the field, electronic databases offer valuable evidence on the diagnostic letterforms that may be helpful in scribal identification. However, as Smith notes, a scribal profile is not exactly equal to a fingerprint: there are material, spatial and temporal constraints and conventions affecting and changing scribal behaviour (1997: 136).

Modern forensic handwriting analysis is not directly comparable to the study of professional scribes who could produce several hands. However, the two fields share similar methodological problems. One of these is determining the influence of internal and external constraints on handwriting. Davis notes that the hand a writer produces is a compromise between the internalised model hand and internal and external constraints such as the writing materials, the medium, the writing environment and the neurophysiology of the writer; furthermore, there may sometimes be a conscious effort by the writer to produce a different hand from the internalised model (2007: 260–261). Forensic document analysis presupposes that a certain writer produces idiographic writing with characteristics that may be ascertained by experts, that remain relatively constant in different writings, and that are unique to that individual (Davis 2007: 261). Davis (2007: 265) stresses the fact that the various components present in the act of writing are highly interdependent: for example, the writing implement essentially affects the resulting text and should therefore be considered in handwriting identification.

[1] For the purposes of this paper, the terms *graphete*, *graphetic variant* and *allograph* are used as synonyms.

Jacob Thaisen suggests that scribal orthography can be analysed with regard to a cline 'with a "short" pole characterised by scarcity of space, a speedy execution of graphic shapes, and a low level of formality, and a "long" pole having the opposite characteristics' (2011: 84). Thaisen argues that such an analysis is copy-specific rather than exemplar-related: the same scribe could have preferred 'long' forms in one copy, given enough time and space, and 'short' forms in another copy, if such resources were lacking (2011: 84). Thaisen (2011) studies the occurrence of 'long' and 'short' forms in Adam Pinkhurst's hand – for example, whether the word <and> is written in full or abbreviated as <&>. He concludes that some such differences are better explained by 'spatial, temporal, and functional constraints' influencing Pinkhurst than by a chronological shift in scribal habits (2011: 84).

In addition to the factors already mentioned, the exemplar used by the scribe may also have an influence on the resulting copy. McIntosh briefly comments on the influence of the exemplar on graphetic features of the copy:

> A good deal of material survives in which this kind of influence can be examined in detail, e.g. where a scribe has copied two or more texts from obviously different sources and where he varies correspondingly in such matters as his choice of letter forms (e.g. the three main shapes of <r>). One interesting characteristic of such an influence is that it tends to assert itself less and less as the scribe proceeds with his work. (McIntosh 1975: 225n)

McIntosh further notes that if the scribe is influenced by the exemplar, the influence is usually not strong enough to conceal the graphic habits of the scribe (1975: 225). Unfortunately, McIntosh does not give any specific references to such material. This is also the only time he refers to the usefulness of analysing graphetic variation in relation to the timeline of manuscript production. Furthermore, McIntosh discusses the Middle English period and scribes working with manuscript exemplars. My material is datable to the mid-sixteenth century on the basis of internal evidence, and there is a strong argument for some of the texts to have been copied from printed exemplars (see Varila 2010: 105–111). This direction of textual transmission has not been studied as extensively as its opposite, printing books from manuscript exemplars. Although this issue is beyond the scope of the current paper, I hope to investigate further the processes and influence of copying from print to manuscript in the future.

McIntosh acknowledges that the same scribe could have produced several scribal styles or *modes*, varying in their degree of formality (1975: 226). Whereas the difference between two modes may be rather striking, stylistic variation may also appear within one mode (1975: 226). Some graphetic variation can probably be explained in such terms, for example when a cursive letterform occurs sporadically in a more formal, palaeographically more elaborate environment. McIntosh stresses the importance of graphetic/graphemic analysis especially in the case of 'variations which seem to reflect no parallel variations in spoken systems' (1975: 235). Allographs could be a case in point here.

It has been suggested that research on the graphetic level might be useful from the point of view not only of paleography but also of linguistics. McIntosh et al. believe that such work should ideally be conducted by combining the 'crafts of the ('written') linguist and the palaeographer' (1986: 6). They note that graphetic variation could be considered parallel to phonetic variation:

> [Details of graphetic variation] have a potential relevance to the dialectology of written Middle English comparable to that which phonetic variations have in the study of spoken dialects. They can be of use in localising the work of individual scribes, and there can be little doubt that it would be profitable to study and plot their distributions. For our knowledge of the graphetic details of the hands in which our specimens are written can be incomparably more precise than our knowledge of the phonetic details of Middle English can ever hope to be. (McIntosh et al. 1986: 6)

Much of the work on scribal behaviour has concentrated on medieval manuscripts. There is less research on the early modern period. Smith notes that even in Middle English texts, spelling 'is often dismissed as accidental, not of interest to the literary or textual critic' (1997: 141). Spelling variation still occurs during the early modern period, but it is perhaps less useful a feature from the point of view of scribal profiling than in medieval material: while variation occurs for example between <u> and <v> or <i> and <j>, such variation does not reveal much about the scribe's dialect. Analysing graphetic variation might provide an additional tool for a period during which the English language becomes more standardised, handwriting becomes generally more cursive, and the number of people capable of writing increases: palaeographic idiosyncrasies may in some cases reveal more than linguistic idiosyncrasies.

Some researchers have developed digital tools in order to answer palaeographic questions. Terras and Robinson (2004) describe a markup system for XML encoding Old Roman Cursive texts on the stroke level, that is, instead of individual letters, every individual stroke is represented. Their corpus consists of approximately 1,700 characters altogether, and according to their estimate, 300 hours of annotation work was needed (Terras & Robinson 2004: 407). Annotating lengthy manuscripts on such a level of precision would thus take thousands of work hours.

The digital tools developed for linguistic research are not necessarily suitable for drawing palaeographic conclusions. Electronic corpora, much like editions, treat orthographic variation in different ways: some corpora heavily edit and modernise their source, while others preserve graphemic distinctions, for example between a word and its abbreviated form (e.g. <and> vs. <&>). Even the corpora that are highly sensitive to variation on the graphemic level rarely – if ever – preserve the variation on the graphetic level, that is, between the various realisations of a single grapheme, e.g. one- and two-compartment /a/. The only such graphetic transcription mentioned by Robinson and Solopova is Hans Fix's transcription of the Saarbrücken Version of the Old Norse *Grágás Konungsbók* (1984), although they state that 'the advocacy by McIntosh (1974; 1975) and Benskin (1990) of "scribal profiles" implies graphetic transcription of Middle English manuscripts' (2006: 2). Thaisen's study, in turn, is based on electronic transcripts which, while retaining graphemes, do not distinguish between allographs (2011: 78). Another example is the *Reference Corpus of Late Middle English Scientific Prose*, where the distinctions between <u> and <v> and <y> and <i> are retained but graphetic variation is not. The editors note that in their material, the choice between variants is dependent on the position of the letter in the word (Calle-Martín et al. 2012: 427). The authors explain that 'graphetic transcription has been discarded on account of the research interests which lie behind the edition itself' (Calle-Martín et al. 2012: 427n). The inclusion of graphetic variants would therefore not have been helpful for their specific research purposes. Wiggins (2004) uses electronic texts for scribal profiling and stresses the

importance of using whole-text data. However, the facsimile the study is based on does not distinguish between letters on a graphetic level (Burnley & Wiggins 2003, see their Editorial and Transcription policy). To the best of my knowledge, no linguistic corpora currently exist that could be used to study graphetic variation.

Robinson and Solopova (2006: 3–4) experimented with the idea of graphetic transcription during the *Canterbury Tales Project*. They noted some problems with that approach: firstly, it is difficult to decide how many forms one should distinguish; secondly, graphetic forms of different graphemes may look similar (for example forms of *o* and *a*); and thirdly, that there is no well-established model for graphetic transcription, whereas best practices of graphemic transcription have already been extensively discussed (Robinson & Solopova 2006: 4). All these problems should be given due consideration when formulating a methodology for analysing graphetic variation. There have been some attempts to use digital tools for differentiating between forms and categorising them, but such an approach is not unproblematic either (see e.g. Ciula 2005). While choosing the graphemic level for their work, Robinson and Solopova call for further work on the methodology of graphetic transcription:

> The way to graphetic transcription lies through refinement of scribal profiles based on selected features of individual manuscripts [...] and through progressive test transcriptions in controlled circumstances (e.g. of parts of different manuscripts written by one scribe; of a single shorter work in relatively few manuscripts). Through this work, a methodology may develop that would permit graphetic transcription of the type not attempted here. (Robinson & Solopova 2006: 4)

The present paper attempts to answer Robinson and Solopova's call by discussing graphetic variation in stretches of text written by one scribe in two manuscripts. I argue that analysing graphetic variation within text produced by a single scribe may prove fruitful not only for palaeographical but also for philological research. The variation pattern may reveal where there are breaks in the chronological sequence of copying, or help show where one scribal stint ends and another one begins. Such evidence can then be used alongside textual evidence and codicological information gathered from the analysis of the material object itself, resulting in a fuller picture of the manuscript as a whole.

Material and methodology

For this pilot study, I analysed stretches of two mid-sixteenth-century English manuscripts: New Haven, Yale University, Medical Library MS 45 (henceforth ML 45) and New Haven, Yale University, Beinecke Rare Book and Manuscript Library MS 558 (henceforth B 558). ML 45 is a paper manuscript of 149 leaves, measuring 300 mm x 205 mm. B 558 (89 leaves) is mixed paper and parchment and measures 200 mm x 150 mm. The analysed sections, the first quire of B 558 and the last quire of ML 45, contain shared material and are written by the same scribe (Voigts 1989: 100n; Birkholz 2003: 36). The quires mainly comprise short astrological prognostications, some of which are found in both manuscripts. The first quire of B 558 consists of twenty leaves (approximately 16 500 words) and the last quire of ML 45 currently has 21 leaves (approximately 20 000 words). The difference in physical size makes it more difficult to compare the handwriting in terms of general aspect. There is, however, very little spelling variation between the shared texts in the two manuscripts.

The scribe in question has not yet been identified. On the basis of internal evidence, B 558 can be connected to the Stalon/Butler merchant family of Harlow, Essex (see Birkholz 2003). According to his testament from 1556 (Essex Record Office D/AMR 1/55), Thomas Butler owned a book 'that Isaac wrote'. I have not yet managed to identify 'Isaac', but it is possible that the hand in the manuscripts, analysed here, belongs to him.

Both manuscripts survive in later bindings and contain several items written in several hands. It is therefore challenging to try to establish the chronological relationship between the texts and to find out which parts of the manuscripts share a common history of production. Based on internal evidence, it seems that the bulk of the two manuscripts was copied within approximately a 10-year period. F. 120r of ML 45 reads: 'Wryten Anno Domini 1551'. B 558 has been dated between 1547 and 1554 (see Birkholz 2003: 11n). The exact relationship between the manuscripts seems complex. There is internal evidence suggesting that at least some texts might have been copied from ML 45 into B 558, but more work is needed to determine the relationship of the manuscripts more closely[2].

My approach is qualitative rather than quantitative, although the results are quantified to an extent to form a starting point for the qualitative work. My intention in the present paper is not to identify or distinguish several hands but rather to explore the variation occurring within one scribal hand and to bring this information to bear on codicological and textual evidence. My criteria for choosing the grapheme to be discussed were as follows:

1) All variants or allographs should occur in both manuscripts in the same hand. This allows for recognizing changes in preferences (as opposed to different scribes having different preferences).
2) The forms should take up approximately the same amount of space, to limit the influence of material constraints (for example, a scribe may use abbreviated forms in line-final position in order to save space and fit the word on the line).
3) The forms should be easily distinguishable. They should, for instance, have a different number of strokes per letter, so that the tokens can be counted without having to take measurements. Ideally, the forms should be distinguishable even in black-and-white or microfilm images. This makes an initial analysis possible without access to the original manuscript. If the forms are analysed by hand, it takes a relatively long time to analyse a representative sample; being able to perform the analysis before exploring the physical object saves a great deal of time.
4) The variation should not be caused by capitalisation. For example, the scribal hand in ML 45 and B 558 contains both one- and two-compartment /a/ forms in word-initial position, but it is sometimes difficult to say whether or not the two-compartment <a> is intended by the scribe to stand for a capital letter. For this reason, word-initial variation is not ideal.
5) Both forms should be found in the same palaeographic context to exclude the influence of the preceding and/or following letterform(s) as a decisive factor in selecting the variant.

[2] A third manuscript that can be connected to ML 45 and B 558 is New Haven, Yale University, Medical Library MS 26. The manuscript group will be discussed in more detail in my doctoral dissertation.

The shared hand in ML 45 and B 558 is a formal secretary hand with three different variants for /s/. The 'long' or 'tall' <s> consistently occurs in word-initial and word-medial position. Correspondingly, the two other variants only occur in word-final position. The tall <s> is not discussed in this paper, as its usage is systematic. The variation that I am interested in is that of the word-final <s> types: the kidney-shaped two-compartment <s> typical of secretary hands, and the sigma-shaped or 6-shaped <s>, found in Anglicana hands.[3] The terminology used to describe letterforms is not consistent across the field, especially concerning early modern handwriting. In this paper, I will refer to the two variants shown in Figure 1 as the 'kidney-shaped' and the 'sigma-shaped' <s> respectively.

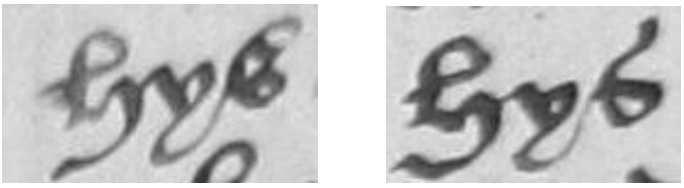


Figure 1. B 558 f. 3r, l. 7 and l. 14 respectively. Beinecke Rare Book and Manuscript Library, Yale University.

All the five criteria listed above were met by the variation occurring between these two word-final /s/ forms. They appear in both ML 45 and B 558 in the same hand. The two variants of /s/ take up approximately the same amount of space. The forms are distinguishable even in black-and-white images, and as they are word-final variants, capitalisation does not influence the variation. Both /s/ forms are found in the word 'hys' (*his*), even on the same page (B 558, f. 3r), as illustrated in Figure 1. This shows that the selection of one variant over another is not dependent on the immediate orthographic environment; that is, the preceding letter does not dictate the selection of the variant following.

Jane Roberts notes that 'it would seem that the most frequently found Anglicana alternatives in reasonably self-consistent Secretary hands are the round <e> and the sigma-shaped <s>, that is, the less time-consuming letter-forms among those distinctive of Anglicana when compared with Secretary' (2008: 212). The sigma-shaped <s> has a lower level of formality, as it can be executed in one continuous pen-stroke and is thus faster to write than the kidney-shaped variant, which consists of more than one stroke. Roberts discusses the hand of Harley 2255 (a collection of Lydgate's poems from c. 1460–1470), stating that the 'Secretary kidney-shaped **s** is used at the end of words, as if a choice made to mark the hand's formality' and that the scribe only 'misses' one word by using an Anglicana variant there (2008: 212, 219, Plate 49). She also notes that by the end of the fifteenth century, the features of Anglicana and secretary scripts are intermingled to such a great extent that the question arises whether it is relevant to distinguish the letterforms on that basis (2008: 212). For the purposes of this paper, the underlying script is not as important as distinguishing between the variant forms. Similarly, the ductus (order of tracing) of the kidney-shaped <s> may vary slightly depending on the preceding letter, but the emphasis

[3] A further grapheme representing word-final /s/ is the abbreviation for *-es*, which looks like an *e* with the curved stroke extending below x-height. It would be possible to compare the abbreviated *-es* forms to the non-abbreviated forms, but as the abbreviated form can be fitted into a smaller space, the spatial constraints may inform the selection of variant here.

here is on the resulting graph and its degree of formality: in the material analysed here, the kidney-shaped <s> always consists of more than one stroke.

For my analysis, I selected one quire of about 40 pages from each manuscript (B 558, ff. 1r–20v; ML 45, 129v–149r) and counted all the instances of word-final /s/ from microfilm images. Some tokens were unclear in the images, usually because of a smudge on the top of the letter. However, these unclear cases only amount to 36 of 2,348 cases; in other words, 98.5 per cent of all the word-final /s/ forms on these pages could be placed in either the kidney-shaped or the sigma-shaped category. I calculated the number of kidney <s> tokens and sigma <s> tokens page by page. It should be noted here that 'page' is, of course, a rather arbitrary unit. However, the purpose of this study was not to produce a full-scale statistical analysis of the variation but rather show whether, and where, the pattern of /s/ usage changes. Therefore, the page can be used as a basic unit: Figure 2 and Figure 3 (shown below) reveal the pattern breaks, which could then be analysed more carefully both in their immediate context and in their wider, codicological context. Final /s/ forms occurring in the rubrics and in other words written in a more formal script (textualis with secretary influence) were ignored, as the script used places a constraint on the repertoire of variants that can be used. Blank pages were excluded from the graphic representations in Figures 2 and 3 but included as evidence for the qualitative analysis.[4] After conducting the analysis of variants based on microfilm images, I studied the original manuscripts in order to compare the results of the /s/ form analysis to the codicological structure of the manuscripts.

Results

The results show that both forms of word-final /s/ are found in both manuscripts, but the balance of preference is different. The kidney-shaped <s> is more common throughout the analysed section of B 558, whereas in ML 45 the sigma-shaped form is preferred on most pages. The results are presented in the form of two figures (Figure 2 and Figure 3). The number of tokens is shown in relation to 100 per cent of tokens of word-final /s/ on each page to facilitate comparison across pages. The actual numbers of the tokens are shown under the graph to increase the comparability of the results, as the number of /s/ forms per page varies.

The resulting figures showcase some interesting features. Figure 2 shows three points within the analysed stretch of ML 45 where the pattern of /s/-form preference changes. Firstly, the kidney-shaped form is preferred on the first 7 pages of the section, after which the preference changes to the sigma <s>. Based on an earlier foliation present in the manuscript and watermark evidence, at least two leaves have been lost between current ff. 132 and 133. The current f. 132 ends in the middle of a short electionary text on the activities suitable for each day of the week. The current f. 133 begins in the middle of a weather prognostication. The first change of preference in /s/ forms thus co-occurs with missing leaves and a gap in contents.

The second change of preference, on f. 137v, co-occurs with a textual border. The top of the page still conforms to the sigma <s> pattern of preference. Below that, approximately two thirds of the page are taken up by two texts, titled *Thinges to be obseruyd by th(e) course of {Moon}* and *Off bloud lettyng after digge(s) saieng(es)*, in both of which the kidney-shaped variant is preferred. Closer inspection of the original manuscript also shows a difference in

[4] In ML 45, ff. 129r, 146v and 149v are blank. In B 558, ff. 5v–6r and 15v–16r are blank.

ink between the first text on the page and the two texts below that. On the preceding three (136r–137r) and following eleven (138r–143r) pages, only the sigma <s> is used. This suggests that *Thinges to be obseruyd* and *Of bloud lettyng* have been added at a later stage, probably to make use of blank space on the page. The general aspect of the handwriting in the two texts matches the texts on ff. 129v–132v in the beginning of the section, where the kidney <s> is also preferred.

The third page with an interesting distribution of /s/ forms is f. 149r. This page seems to show genuinely mixed usage of the two forms throughout the text. This is the last folio of the quire and also of the manuscript in its current state. The variation could perhaps be explained by the text having been added at a later stage to fill the quire, with less attention paid to the consistency of the hand.

Although Figure 2 shows some clear changes in preference, the leaves 129–149 actually form a single quire. The layout within the quire seems consistent. Ff. 129r and 149v, the first and last pages of the quire, are blank (except for rules and later annotations). Ff. 129r–v and 149r–v, the outer leaves in the quire, have double rules in all margins in diluted ink, and ff. 130r–148v have single rules in the same ink. The margins in this quire are generally narrower than elsewhere in ML 45. The changes in <s> usage thus do not co-occur with quire boundaries here. Instead, they may offer evidence on the sequence of copying within the quire. It is possible that the more formal variant was used in the beginning of the work, and at some point the preference changed in favour of the sigma variant, as it is speedier to execute than the kidney <s>. The increased degree of formality in the addition on f. 137v can be explained as representing a different writing stint.

In B 558 (cf. Figure 3), the kidney-shaped variant is preferred on every page within the section analysed (ff. 1r–20v). As noted, there is some evidence for some of the shared texts to have been copied from ML 45 into B 558. The addition on f. 137v of ML 45 might be used to support the suggestion of a chronological shift in the scribe's preferences. However, if the kidney <s> was preferred by the scribe at a chronologically later time than the sigma <s>, this would also seem to suggest that the first pages of the quire of ML 45 analysed were copied later than the larger part of the quire. This is perhaps not very plausible, as the layout of that quire seems consistent and planned. The most striking changes in /s/ form usage in ML 45 are perhaps best explained simply by a temporal break between scribal stints, whether days or years.

As stated, the kidney <s> is preferred throughout B 558. As with ML 45, the analysed section forms a single quire (1r–20v) – the first quire of the manuscript in its current state. There is some variation in /s/ usage, and occurrences of sigma <s> are concentrated in the beginning and in the end of the quire. After the blank pages 5v–6r, variation is visible on f. 6v, where more than a fifth of all final /s/ forms are sigma-shaped. All five sigma forms occur in the first text on the page. There are, however, only 2 occurrences of final /s/ altogether in the latter text. Interestingly, this page showing notable variation in /s/ forms contains the same items as those found in the addition on f. 137v of ML 45, where the items stand out from their surroundings in terms of their general aspect and the preference for kidney-shaped <s>. F. 6v is followed by a rather consistent stretch of kidney <s> usage, as can be seen from Figure 3. Pages 15v–16r are blank; the final 7 pages after the break again contain instances of the sigma-shaped variant. Most of these occur within a single text, titled *Judgment upon the*

12 houses. It is possible that there was a production break between copying this text and the preceding items.

While the analysed sections in B 558 and ML 45 both show a preference for one or the other variant when the total numbers of word-final sigma-shaped and kidney-shaped <s> forms are compared, the method used here reveals shifts in preference that may be used as a starting point for a more detailed analysis. The mechanisms of graphetic preference could perhaps be compared to Smith's (1997: 139–140) discussion of *fixed* and *focussed* usage of language: there is some room for variation despite script 'standards', such as Anglicana or secretary. Analysing the patterns of preference or focus may then be used as additional evidence for tracing the timeline of manuscript production.

Considering that the kidney-shaped variant is more formal than the sigma <s> in terms of strokes needed, it is perhaps surprising that it is the smaller manuscript which shows a consistent preference for the use of kidney <s>. Whereas Thaisen (2011) found that the short forms tend to occur when resources (time and space) are lacking, in my material the 'short', one-stroke form occurs in the larger (folio-sized) manuscript whereas the 'long', or palaeographically more complex, form occurs in the smaller (quarto-sized) manuscript with less marginal space. The more complex allograph thus dominates in what is the more modest manuscript of the two in terms of size. However, the writing support of ML 45 is paper, whereas B 558 is mixed paper and parchment. The quire of B 558 analysed in this paper is itself mixed: leaves 1, 7–14 and 20 are parchment, the rest are paper. It is potentially relevant that the most consistent stretch of kidney <s> usage overlaps with the parchment leaves in the middle of the quire.

More research is needed in order to determine the relationship between material constraints, chronological developments, and the patterns of graphetic variation. The preference pattern in the analysed sections could be compared to the rest of the manuscripts, or their relevant sections. The analysed quire in ML 45 stands out in terms of its layout which separates it visually from the other quires. It seems that B 558, in its current state, is not strictly speaking a product of intentional planning, as it consists of quires of different sizes and in different hands. Comparing the quire analysed here to the other sections written by the same scribe might shed more light on the production history of the manuscript.

It is also worth pointing out here that only choosing a few pages from both B 558 and ML 45 might have produced the misleading argument that in ML 45 the scribe only uses the sigma <s> and in B 558 only the kidney-shaped form, whereas both forms are actually found in both manuscripts, although the dominant variant is different in the two. McIntosh (1975: 221) notes that the size of a sample for scribal profiling should be of 'reasonable length (say not less than a thousand words, but no fixed figure can be given)'. If one were to take a sample from ML 45, ff. 138r–143r, that sample would consist of approximately 6,000 words without showing a single example of kidney <s> usage. This demonstrates the importance of a representative sample in gaining a reasonably reliable picture of a scribe's repertoire, even within one style or mode of writing. The aspect of the hand or the spelling conventions may remain similar throughout, while a different pattern can be detected by analysing the sample on a graphetic level.

Conclusion

In this pilot study, I explored one way in which palaeographical analysis on the graphetic level could be used to provide clues on manuscript production. Although the present paper discussed two manuscripts, the same idea could be employed when studying a single manuscript showcasing variation within the hand of a single scribe. This approach could be tested further and a diagnostic set of graphemes and their variants could be formed. Other graphetic and visual variants could also be examined to test the method further. In addition to other cases of graphetic variation, the repertoire of abbreviations employed by the scribe, punctuation, or the spelling conventions could be analysed.

This method could be useful from the point of view of scribal identification, which is still a very difficult task. Graphetic variation could be quantitatively analysed and the results used as an additional criterion when trying to identify hands. Such an analysis might also shed light on the development and change of a single scribe's practices over time and offer one more tool for placing manuscripts or their sections on a chronological continuum. Mapping the patterns of preference both within and between scribal hands could contribute to forming a more detailed picture of scribal work in the medieval and early modern periods.

The results of this pilot study suggest that an analysis of select graphetic variants may also be helpful in determining the textual structure of manuscript material, for example establishing the chronological sequence of manuscript production. Comparing letterforms quantitatively and analysing the results qualitatively could thus be used as a help in solving textual and codicological problems. The method could therefore be used to support philological analysis of manuscripts and their text.

One benefit of the method presented is that it offers a measurable variable: in addition to noting that the general aspect of handwriting on a certain page seems different compared to its surroundings, one can pinpoint at least one quantifiable variable contributing to that difference. Another benefit is the possibility of performing the initial analysis on the basis of a microfilm copy or digital images of the original manuscript. However, due to the lack of suitable electronic corpora, analysing graphetic variation in representative samples of text is currently rather time-consuming.

Finally, it should be noted here that this method cannot replace a codicological analysis of the original manuscripts, a textual collation, or a full-scale palaeographical analysis of the scribal hand. Instead, it is intended to complement the results gained by such well-established methods. It is to be hoped that new resources currently under construction, such as the Digital Resource and Database of Palaeography, Manuscripts and Diplomatic (DigiPal), will facilitate new methodological developments at the intersection of palaeography and philology.

Acknowledgements

I am grateful to the anonymous referee for their valuable comments and to Langnet (The Finnish Graduate School in Language Studies) for supporting my research.

School of Languages and Translation Studies, Department of English
University of Turku
FI-20014 University of Turku
Finland
E-mail: mari-liisa.varila@utu.fi

REFERENCES

Birkholz, David. 2003. The Vernacular Map: Re-charting English Literary History. In David Lawton, Rita Copeland & Wendy Scase (eds.), *New Medieval Literatures, Volume VI,* 11–77. Oxford: Oxford University Press.

Burnley, David & Alison Wiggins (eds.). 2003. The Auchinleck Manuscript Project. http://auchinleck.nls.uk

Calle-Martín, Javier, Laura Esteban-Segura, Teresa Marqués-Aguado & Antonio Miranda-García. 2012. The *Reference Corpus of Late Middle English Scientific Prose.* In Jeremy Jancsary (ed.), *Empirical Methods in Natural Language Processing: Proceedings of the Conference on Natural Language Processing 2012*, 424–432. http://www.oegai.at/konvens2012/proceedings.pdf

Ciula, Arianna. 2005. Digital Paleography: Using the Digital Representation of Medieval Script to Support Palaeographic Analysis. *Digital Medievalist* 1. http://digitalmedievalist.org/journal/1.1.

Davis, Tom. 2007. The Practice of Handwriting Identification. *The Library*, 7th series, 8(3): 251–276.

Digital Resource and Database of Palaeography, Manuscripts and Diplomatic (DigiPal). http://www.digipal.eu

McIntosh, Angus. 1975. Scribal Profiles from Middle English Texts. *Neuphilologische Mitteilungen* LXXVI(2): 218–235.

McIntosh, Angus, M. L. Samuels, Michael Benskin, Margaret Laing & Keith Williamson. 1986. *A Linguistic Atlas of Late Mediaeval English. Volume One: General Introduction, Index of Sources, Dot Maps.* Aberdeen: Aberdeen University Press.

Mooney, Linne. 2006. Chaucer's Scribe. *Speculum* 81(1): 97–138.

Mooney, Linne, Simon Horobin & Estelle Stubbs. Late Medieval English Scribes. Accessed November 23, 2012. http://www.medievalscribes.com. ISBN 978-0-9557876-6-9.

Roberts, Jane. [2005] 2008. *Guide to Scripts Used in English Writings up to 1500.* London: British Library.

Robinson, Peter & Elizabeth Solopova. 2006. Guidelines for Transcription of the Manuscripts of the Wife of Bath's Prologue. http://www.canterburytalesproject.org/CTPresources.html

Smith, Jeremy. 1997. Dialect and Standardisation in the Waseda Manuscript of Nicholas Love's *Mirror of the Blessed Life of Jesus Christ*. In Shoichi Oguro, Richard Beadle & Michael G. Sargent (eds.), *Nicholas Love at Waseda: Proceedings of the International Conference 20-22 July 1995*, 129–142. Cambridge: Brewer.

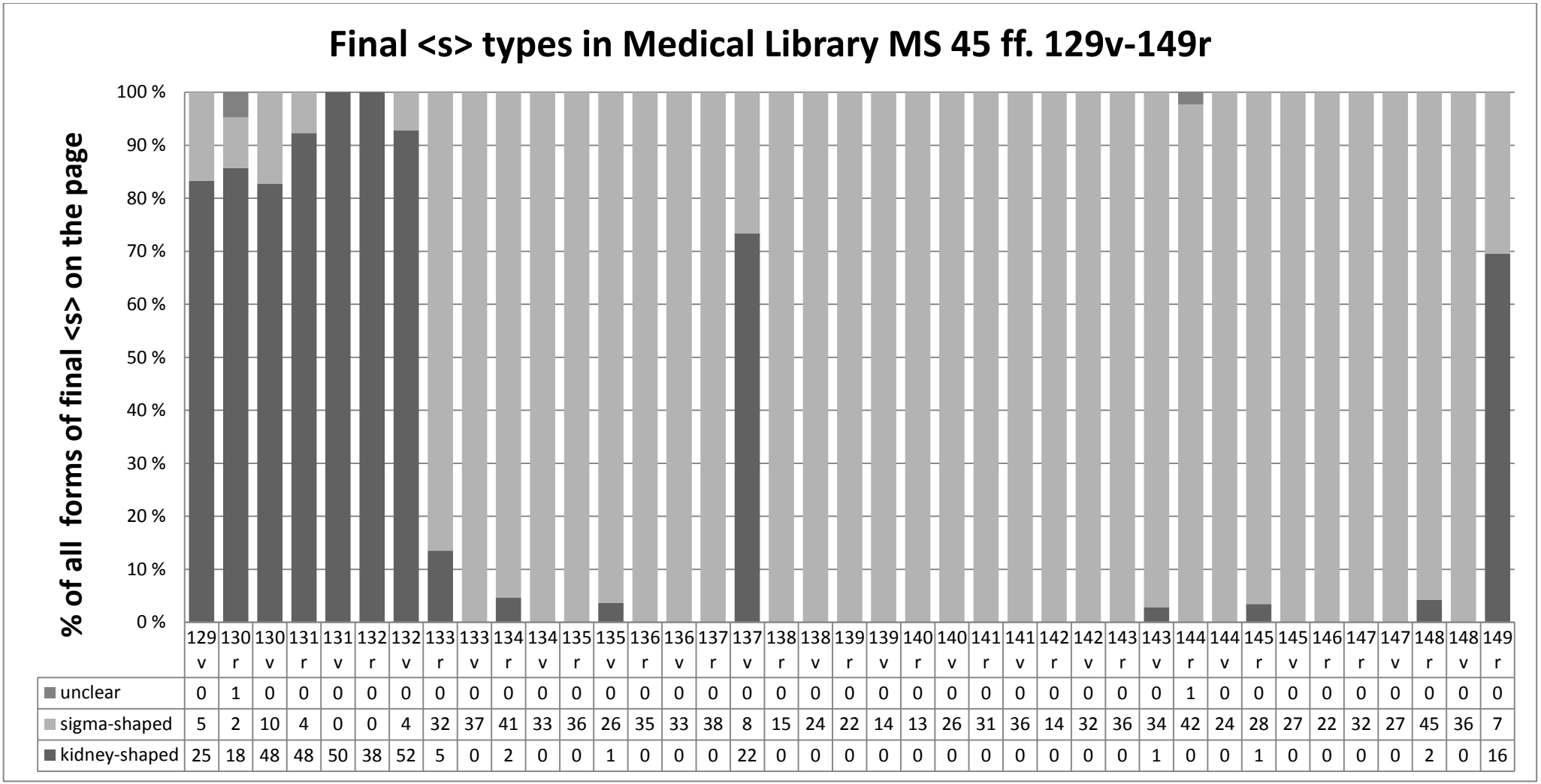
Terras, Melissa & Paul Robertson. 2004. Downs and Acrosses: Textual Markup on a Stroke Level. *Literary and Linguistic Computing* 19(3): 397–414.

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.

Varila, Mari-Liisa. 2010. Astrological Prognostications in MS Beinecke 558*: "*Sonday prime drye wether*"*. Master's thesis, University of Turku. https://www.doria.fi/handle/10024/59501.

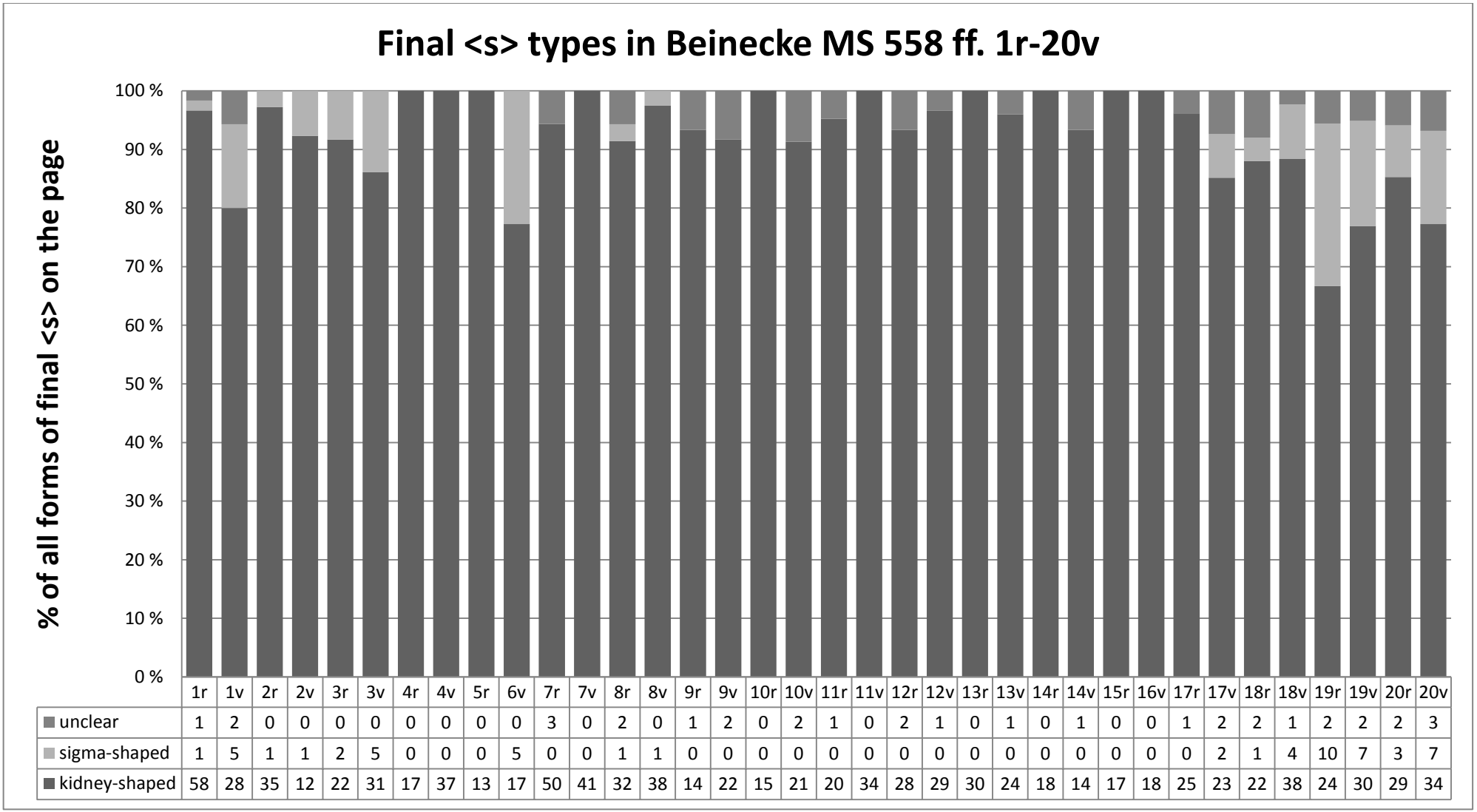
Voigts, Linda Ehrsam. 1989. The Character of the Carecter: Ambiguous Sigils in Scientific and Medical Texts. In A. J. Minnis (ed.), *Latin and Vernacular: Studies in Late-Medieval Texts and Manuscripts* (Proceedings of the 1987 York Manuscripts Conference), 91–109. Cambridge: Brewer.

Wiggins, Alison. 2004. Are Auchinleck Manuscript Scribes 1 and 6 the Same Scribe? Whole-Data Analysis and the Advantages of Electronic Texts. *Medium Aevum*, 73(1): 10–26.



	129v	130r	130v	131r	131v	132r	132v	133r	133v	134r	134v	135r	135v	136r	136v	137r	137v	138r	138v	139r	139v	140r	140v	141r	141v	142r	142v	143r	143v	144r	144v	145r	145v	146r	147r	147v	148r	148v	149r
unclear	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
sigma-shaped	5	2	10	4	0	0	4	32	37	41	33	36	26	35	33	38	8	15	24	22	14	13	26	31	36	14	32	36	34	42	24	28	27	22	32	27	45	36	7
kidney-shaped	25	18	48	48	50	38	52	5	0	2	0	0	1	0	0	0	22	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	2	0	16

Figure 2. Final <s> types in Medical Library MS 45 (ff. 129v-149r).



	1r	1v	2r	2v	3r	3v	4r	4v	5r	6v	7r	7v	8r	8v	9r	9v	10r	10v	11r	11v	12r	12v	13r	13v	14r	14v	15r	16v	17r	17v	18r	18v	19r	19v	20r	20v
unclear	1	2	0	0	0	0	0	0	0	0	3	0	2	0	1	2	0	2	1	0	2	1	0	1	0	1	0	0	1	2	2	1	2	2	2	3
sigma-shaped	1	5	1	1	2	5	0	0	0	5	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	4	10	7	3	7
kidney-shaped	58	28	35	12	22	31	17	37	13	17	50	41	32	38	14	22	15	21	20	34	28	29	30	24	18	14	17	18	25	23	22	38	24	30	29	34

Figure 3. Final <s> types in Beinecke MS 558 (ff. 1r-20v).