5 Standardisation, exemplars, and the Auchinleck manuscript

1 Introduction

Samuels’ 1963-article “Some applications of Middle English dialectology” situated the first steps in the formation of present-day Standard English in fourteenth- and fifteenth-century London, the home of three of his four incipient standards. The orthographic forms respectively selected by Scribes 1 and 3 of the Auchinleck manuscript, National Library of Scotland, Edinburgh, MS Advocates’ 19.2.1, dated c. 1330–40, exemplify the earliest of the London-based types, Type II. This type dies out suddenly c. 1380 and is replaced by Type III, which in turn is equally suddenly replaced by Type IV half a century or so later. Samuels’ four types have been influential; for example, Kane and Donaldson (1975) explicitly selected Cambridge University Library, MS B.15.17 as the base text for their edition of the B version of Piers Plowman on the grounds that it is written in Type III. The types have come under fire in recent years but they continue to prove resilient despite the complete absence of contributions countering the criticisms. For example, the types are unconditionally accepted in a textbook on manuscript studies focusing on the late Middle English period (Kerby-Fulton et al 2012: 67), go entirely unquestioned in a widely used undergraduate linguistics textbook (Horobin and Smith 2002), and are reproduced in as many as three of the fifteen chapters in a recent handbook on Middle English (Brinton and Bergs 2017), including in the chapter specifically devoted to standardisation. It is time to lay the types to rest.

To fulfill this goal, this paper adds to the criticisms by questioning the basis for Type II. What follows details my methodology for orthographic analysis, which is able to discriminate the six scribes of the Auchinleck manuscript and the hands who produced the immediate exemplars. Relating how the exemplar hands are distributed to the manuscript’s codicology strongly suggests the exemplars were obtained from local sources which also produced them. A later section discusses orthographic standardisation because there is evidence that the orthographic forms selected by Scribes 1 and 3 are no more similar than the forms selected by the manuscript’s other scribes, contrary to what would be expected of a standard even at a very early stage in its formation. The final section summarises.

Note: Thanks are due to Laura Wright, Lawrence Warner, and two anonymous reviewers.

https://doi.org/10.1515/9783110687545-006
2 Perplexity distribution in the Auchinleck manuscript

How orthographically similar are two texts and why? A standard means of answering this question is to collect from both texts the orthographic forms they use for various pre-selected lexical, morphological, or phonological items. The items are ones that are very likely to occur in any text, such as function words and common lexical words, and analysis of their forms will typically concentrate on a specific part of them that is known to distinguish texts from each other. An example of such a part is how the third-person singular present indicative verbal suffix is represented, since its representation is known to differ between southern and northern texts. The number of forms considered is sometimes rather low, and the forms may have been collected from extracts. Any observed differences between the two texts in their inventory of forms and/or the forms’ relative frequencies are explained as the result of variables having selected those forms. A variable often and correctly invoked is dialect, which, however, must be carefully distinguished from localisation. The latter represents a conflation of the total set of variables into a single one and cannot be assumed primarily to reflect the former. Williamson (2000) explains the distinction as the placing of a real-world locality on a map (‘geographical localisation’) versus the fitting of a text’s orthographic profile into a typology (‘linguistic localisation’). It is none the less standard to express linguistic localisation by reference to geographical space, a practice followed in this paper.

My methodology does not rely on pre-selected items and does not focus on any specific part of an orthographic form at the expense of other parts. Instead it takes into account every single orthographic form comprising the texts. I ask how well a probabilistic language model trained on all the forms found in one text is able to account for all the forms found in another text. Language models, including probabilistic ones, are a stable in natural language processing and are at the core of many applications involving pattern recognition, such as for example machine translation and optical character recognition. An extended example will clarify. A traditional linguistic profile of text A for a single item gives all the orthographic forms for that item found in text A. It may also give the orthographic forms’ respective frequencies in absolute or relative terms. For example:

SUCH: <such> 28x, <suche> 23x, <swilk> 4x, <suylk> 1x

---

1 Whether a text’s geographical localisation reflects the place where its scribe’s received his training or the place where he copied the text is another matter.
A profile of text B contains the forms <such>, <sucH>, and <swylk> for that same item. Disregarding frequency, it is reasonable to infer a high level of similarity between texts A and B from a comparison, since two forms are identical and the texts additionally share the features <swv> and <lk>, where V represents <i> or <y>; it is common for medieval scribes to use these two letters interchangeably. The level of similarity established, it is up to the researcher to put forward the variables that best explain how the similarity has come to be. Texts A and B might share exemplars if both are copies of the same literary work, or the form <lk> might suggest that they share a connection to a county like Lincolnshire where /l/ had not vocalised and /k/ had not palatalised.

A language model is an exhaustive version of the traditional linguistic profile, and “to train a model on text A” is just another way of saying “to compile a profile of text A”. The unit that a model records could in principle be anything from the single letter to the whole word. Single letters record too little information to be useful for comparison of orthography. Units of two or three letters capture orthography well, whereas larger units capture less orthography and more lexicon. If the unit is three letters, a 3-gram, the recorded orthographic forms for SUCH in text A are:

#su, #sw, suc, swi, suy, uch, che, wil, ylk, yyl, ch#, he#, lk# (where # means beginning/end of word)

The linguistic profile for the item SUCH in text A corresponds to a 3-gram language model:

<#su> 51x, <#sw> 4x, <suc> 51x, <swi> 4x, <suy> 1x, <uch> 51x, <che> 23x, <wil> 4x, <ilk> 4x, <uyl> 1x, <ylk> 1x, <ch#> 28x, <he#> 23x, <lk#> 1x

To estimate the level of similarity between texts A and B is to calculate the probability of encountering the text B orthographic forms in text A. I do not give the equations here. Suffice it to say that they involve multiplying by the frequencies of the 3-grams that make up the forms, which has a consequence that text B’s <swylk> will be assigned zero probability since <swy> and <wyl> are both unattested in text A. However, it can be inferred from the presence of both <swyk> and <suylk> that <swylk> is a possible form in text A, as was argued above, although it ought to receive a lower probability than <such> and <sucH>.

To make the transition from a language model, which in essence is indistinguishable from a traditional linguistic profile, to a probabilistic language model is to incorporate two measures to simulate that inferential process. The one
measure is to substitute the frequencies of the 3-grams <swy> and <wyl> with the frequencies of the constituent 2-grams <sw>, <wy>, and <yl>, and if this still leads to multiplication by zero, that is, if they too are unattested, then with the frequencies of the constituent 1-grams <s>, <w>, <y>, and <l>. This measure is termed “interpolation” or “backoff”. The other measure is to add infinitesimal probability to every gram in the model, say by adding .01 to every frequency so that <#su>'s frequency is raised from 51 to 51.01, <#sw>'s from 4 to 4.01, and so forth. An unattested gram will thus be assigned the frequency .01, but notice that since every form is composed of several grams, two unattested forms will not receive the same probability. This measure is termed “smoothing”, and while the example of adding a constant illustrates the principle, what is added to the frequencies in professional applications is not a constant but a variable, and there is some debate in the literature about what variable leads to the closest simulation of human inference. Scaling up by adding up all the probabilities of all the individual forms gives the level of similarity between texts A and B. If that probability is 1:69, the perplexity is just the denominator, which is always a positive integer. The lower the perplexity, the more similar are the two texts.

Note the asymmetry of these similarity metrics. The perplexity of a probabilistic language model trained on text A when tested on text B will not equal the perplexity of a probabilistic language model trained on text B when tested on text A.² Note also that the volume of training data affects a model’s accuracy, for the greater this volume is, the less interpolation and smoothing will be required to handle unattested forms.

Quantification does not in and of itself explain why two texts are similar at a given level. Above, I pointed to shared exemplars and shared geographical localisation as possible explanatory variables. In what follows I elucidate methodology in a more technical and detailed manner, before establishing how perplexity is distributed in the Auchinleck manuscript and discussing a number of possible explanatory variables one by one. An important finding is that similar segments tend to be similar because they share a scribe.³ “Scribe” must be the strongest

² Most textbooks on Natural Language Processing that have a chapter on statistical language models will explain more exhaustively. Thaisen (2012) gives a fuller example of how perplexity is calculated for N-gram models of Middle English orthographic data.

³ In Thaisen (2009) I built separate probabilistic language models of all 50+ manuscript copies of two of the tales from Chaucer’s Canterbury Tales and established the perplexity of every model on every tale. I found a strong tendency for the two tales by the same scribe always to have lowest perplexity relative to other combinations of tales and scribes/manuscripts. This tendency, then, meant that a shared scribe was the strongest predictor variable in that dataset.
explanatory variable, for the perplexity distribution closely matches the scribal stints that paleographers have established.

I took the following steps to build probabilistic models of the Auchinleck manuscript. All punctuation was removed from a transcript of the manuscript downloaded from the Oxford Text Archive (Burnley and Wiggins 2003), and all emphatic letter shapes were made nonemphatic (lowercased). Tags were introduced to mark word- and line-boundaries. The transcript was then segmented at every 200th line, and the SRI Language Modelling Toolkit (Stolcke 2002) built a separate model for the gram length 3 for each segment. The toolkit assigned probability to every such gram based on its frequency in the segment.

3-grams spanning the space between consecutive words were excluded in order to reduce unwanted lexical effects, that is to say, in order to minimise the capture of a segment’s lexicon in the process of modelling its orthography. If every line comprises thirty 3-grams and there are an average six words to a line, discarding the five 3-grams spanning the space between consecutive words within the line amounted to discarding around one-sixth of the data (30 less 5 is 25). Roughly speaking, the basis for the models, then, were 58,000 lines x 25 3-grams/line = 1,450,000 3-grams.

The models were smoothed according to the method devised by Witten and Bell, which weights the frequency-derived probability assigned to a gram according to the number of unique contexts in which it is attested so as further to reduce the undesirable lexical effects. The models were also linearly interpolated, that is to say models were built also for gram lengths 1 and 2. Any final segment shorter than 200 lines was ignored.

The toolkit next tested every segment on every model, returning a separate perplexity for each model on each segment. The resulting perplexity distribution had no noteworthy skew so no further action was taken to normalise it. The mean perplexity on all the segments and its standard deviation were established for every model and visualised by means of a scatterplot. The scatterplot revealed the perplexity distribution in the Auchinleck manuscript to comprise several groups of consecutive segments with similar means and standard deviations and such groups to be interrupted by abrupt shifts. The scatterplot is given in Figure 5.1 below.

The process of first segmenting the transcript, then modelling every segment separately, testing every model on every segment, and eventually identifying groups of models with similar perplexity was repeated with 2-gram models, with several other segment sizes, and with odd and even lines modelled and tested separately to ensure that the groups were no artefact of the method or property of the segments’ lexicon. Their existence verified, the proposed groups were further isolated through exclusion of every 200-line segment.
falling at a transition between two groups. The R software environment for statistical computing subjected the groups to a one-way ANOVA test in conjunction with Tukey’s Range Test, the latter with Kramer correction to compensate for the unequal number of segments in each group, to test the independence of the groups. Groups found not to be independent of each other were collapsed as sets, and the tests for independence repeated on the sets. Table 5.1 gives the final analysis. The groups are labelled alphabetically from “a” to “l” reflecting their order in the Auchinleck manuscript.

Table 5.1 shows a division into six sets. Groups a, c, g, i, and k form a single, non-consecutive set, as does groups b, e, and l. It can be seen from the p-values given in the table that the sets are clearly discriminated. The table identifies the scribe responsible for all text in a set, for the sets strongly correlate with the scribal stints established by paleographers. These statistics, then, robustly show “scribe” to be a salient predictor variable in explaining the perplexity distribution. There is

Figure 5.1: Perplexity distribution in the Auchinleck manuscript, based on 200-line segments.

What is visualised in Figure 5.1 is the mean perplexity on all the segments and its standard deviation for every model. These data were the input to the grouping procedure so that when a group of models was tested for independence from other groups, it was so on the mean of the mean perplexity of the models in the group. As a result, each group does not subsume an equal number of models – in particular groups b, e, f, and l subsume fewer models than the other groups – which could potentially render them incomparable, as one of the anonymous reviewers has rightly pointed out. However, Kramer correction, which was applied, is the variant of Tukey’s Range Test designed for an unequal number of segments. An alternative strategy would have been to segment the transcript afresh along the boundaries of the proposed scribal stints, train models on the resulting segments, and repeat the analysis. I acknowledge that this strategy would have been a viable one.

Thaisen (2013), Table 1 gives the p-values for the groups prior to their merger into sets.
one exception: The statistics do not show the respective sets corresponding to the stints executed by Scribes 3 and 6 to be statistically significantly different from each other.

The trouble with similarity metrics such as these is that all one can demonstrate is that the groups and sets exist. As was hinted at in the opening section, one can show that two texts have similar orthography but one cannot immediately tell why they are similar; this is so with traditional linguistic profiles too, of course. One cannot attribute the similarity to predictor variables other than by pointing to correlations, and one cannot gauge the strength of a variable by varying it while keeping all the other variables constant. While Table 5.1 provides a strong case for “scribe” as being a predictor variable that explains much of the variation in the perplexity distribution, it should be considered what other variables may explain or confound the metrics. Some can be ruled out. The sets do not correlate with any division of the Auchinleck manuscript by booklet, quire, ink, paraph, textual item, or theme-based grouping of textual items. There is no change of scribe within any textual item, and the genre-based koiné evidenced by rhyming usage does not show in these metrics either, but rhyming usage was not modelled separately from the rest of the line. I am unaware of obvious codicological or paleographical traces of temporal constraints having operated on any of the six scribes, and there are no marked shifts in formality between different parts of the manuscript. Either variable could plausibly have influenced a scribe’s selection of orthographic form at a given location of text.

By contrast, a variable for which a case can be made is spatiality. A need to cram the text into a preset space may, perhaps unsurprisingly, lead a scribe to select shorter orthographic forms among the alternatives available to him, while a perceived need to fill a generous such space may have the opposite effect (Peikola 2011; Thaisen 2011, 2013). This variable could explain the perplexity distribution in the three textual items copied by Scribe 2 of the Auchinleck manuscript (Speculum Guy of Warwick; The Sayings of the Four Philosophers;
The Simonie). Since one of them, The Sayings, is a mere 98 lines long and occupies a position between Floris and Blancheflour by Scribe 3 and The Battle Abbey Roll by Scribe 4, it constitutes a transitional segment in relation to the data summarised in Table 5.1 and so is omitted from them; but when segmentation is by textual item, the perplexity of the model trained on The Sayings does differ statistically significantly from the respective ones trained on Scribe 2’s other two items. If this is not simply due to the sparsity of the data The Sayings provide, it could be because Scribe 2 writes The Sayings on folios ruled by Scribe 1 and compresses his script to fit the ruling, whereas no such compression characterises his other two items. In fact, Scribe 2 writes the Speculum Guy on pages ruled for fewer lines than the standard 44 lines and enlarges his script to fill them, and he writes The Simonie on folios laid out in single columns that provide ample space for accommodating its long lines; it would have required considerable squashing to fit the text within the double column format that is standard elsewhere in the manuscript.

Spatiality could also explain why the greatest variation between segments copied by a single scribe is found within Scribe 3’s stint. The respective models trained on his first five 200-line segments all have higher perplexity than any of those trained on other segments copied by him. This difference is clearly visible in the illustrative scatterplot based on the 200-line segments (Figure 5.1) but it is not statistically significant. It stays non-significant when the methodology is repeated separately for each scribe’s contribution so as to keep the predictor variable “scribe” constant. The five segments relate to The Assumption of the Blessed Virgin, which is laid out with generous spacing in the horizontal dimension as compared with the other three items. The result is that the scribe has filled up the lines by selecting longer orthographic forms from among the alternatives available to him.

Another variable for which a case can be made is language. The language of the Auchinleck manuscript is almost exclusively English but it does contain text in other languages. The probabilistic models are sufficiently robust for a Latin word or two to matter little to the similarity metrics but Scribe 2’s Sayings contain as many as ten lines written in French. This amount of text may not be insignificant in a 200-line segment, let alone a 98-line poem. Similarly, Scribe 4’s sole item, the Battle Abbey Roll, is an enumeration of 551 Norman surnames and as such does not follow English graphotactic rules.

It is well-known that scribes may spell non-identically in verse and prose. This variable is, however, often not perpendicular on spatiality, as verse and prose may be laid out differently. The distinction is not salient as a predictor variable in relation to the Auchinleck manuscript. There never was any potential for it to be so, for the simple reason that the manuscript does not contain any prose,
save for the *Battle Abbey Roll* whose contents are not discursive. End-rhymed verse dominates the manuscript among the verse forms, but alliteration is found, in whole or in part, in *The Thrush and the Nightingale*, *Sir Tristrem*, *The Four Foes of Mankind*, and *The Simonie*. Three of these four textual items are excluded from the similarity metrics summarised in Table 5.1 on account of their falling at transitions between groups in the perplexity distribution for the models of the 200-line segments. The fourth item, *Sir Tristrem*, does none the less not constitute an independent group within the output of Scribe 1 according to the one-way ANOVA test. Nor do lines 1–474 form an independent group within *Sir Beves of Hampton*, despite the verse form changing from tail-rhyme to couplets at line 475, which is not a significant juncture in the narrative. A change in poetic form is present in the corresponding location in the Anglo-French version on which the English text is loosely based.

If the agreement between orthography and paleography indicates that “scribe” is the strongest predictor variable, repetition of the methodology on text exclusively written in a single scribal hand will reveal the scribes behind the exemplars. Accordingly, I divided all text in the hand of Scribe 1 into 200-line segments, modelled each segment separately, computed each model’s perplexity on each segment by means of the Toolkit, established each model’s mean perplexity on all the segments and the standard deviation, visualised the perplexity distribution by means of a scatterplot, identified groups and transitional segments, eliminated transitional segments, and tested the independence of the groups and ultimately sets of groups by means of a one-way ANOVA and Tukey’s Range Test.

I repeated these steps separately for each of the remaining scribes, resulting in Table 5.2 as the final analysis. It can be seen from the p-values given in the table that the sets are clearly discriminated. Each set is described as comprising one or more textual items just like Table 5.1 lists its sets by scribe. This is because I have not found a statistically significant change in perplexity within any of the textual items. Note that thirteen textual items do not appear in the table, for example both *The Sayings of the Four Philosophers* and *The Simonie* by Scribe 2. Their omission from consideration is due to them either occupying a transitional segment or constituting a stint less than 200 lines long that is sandwiched between stints by other scribes in the manner of *The Sayings*.

With these reductions in the amount of text considered, Table 5.2 strongly suggests that Scribe 1 worked from exemplars in four hands, Scribe 3 from ones in three hands, and Scribe 5 from ones in two hands, while no more than a single hand was behind the respective exemplars separately for each of Scribes 2, 4, and 6. The separation of scribes for the purpose of Table 5.2 does not permit one to determine any possible identity between an Auchinleck scribe and an exemplar hand, nor any possible instance of two of the scribes drawing on the
Table 5.2: Pairwise comparison of mean perplexities for sets of 200-line segments from the Auchinleck manuscript, in p-values.

<table>
<thead>
<tr>
<th>Scribe 1 Item</th>
<th>1 2 22 28 30 31</th>
<th>3 4 33 37 38 40</th>
<th>7 11 12 13 41 42</th>
<th>5 6 23 26 43</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 22 28 30 31</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>3 4 33 37 38 40</td>
<td>.0000001</td>
<td>—</td>
<td>.0000001</td>
<td>—</td>
</tr>
<tr>
<td>7 11 12 13 41 42</td>
<td>.0000001</td>
<td>.0000001</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>5 6 23 26 43</td>
<td>.0000816</td>
<td>.0000001</td>
<td>.0000001</td>
<td>—</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scribe 2 Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scribe 3 Item</th>
<th>16</th>
<th>17 18</th>
<th>19</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17 18</td>
<td>.0000001</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>.0000001</td>
<td>.0015620</td>
<td>—</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scribe 4 Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scribe 5 Item</th>
<th>24</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>.0000494</td>
<td>—</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scribe 6 Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
</tr>
</tbody>
</table>


Note: The following items are omitted (scribe indicated in square brackets): 8 [1], 9 [1], 14 [3], 15 [3], 20 [2], 21 [4], 27 [1], 29 [1], 34 [1], 35 [1], 36 [1], 39 [1], and 44 [2].
same exemplar hand (but see below). Every one of the hands who supplied Scribe 1 with exemplars did so for at least five textual items. Such recurrence of hands would be very unlikely if independently circulating individual textual items were collected and appropriated as exemplars. In addition, the absence of any statistically significant change in perplexity within a textual item shows that there was no change of hand within the exemplars for any individual textual item. This, in turn, implies that each scribe received the exemplars for each separate textual item as an integral whole.

The similarity metrics further show that Scribe 1 drew on exemplars prepared by the same hand on more than one occasion. This must be so in light of where the exemplar hands occur relative to the quiring, for their distribution preclude the possibility that Scribe 1 exhausted the exemplars written in one hand before turning to those written in other hands. For example, it is impossible to reconstruct a consecutive progress of copying for the first exemplar hand which reconciles its appearance in *The Legend of Pope Gregory* and *The King of Tars* as the first booklet’s first two items with its appearance in *A Penniworth of Witte*. This is because the latter is situated in the middle of the fifth booklet following a stint by Scribe 5 (*Sir Beves of Hampton*), a stint by Scribe 1 that is not based on the first exemplar hand (*Of Arthur and of Merlin*), and another stint by Scribe 1 that is a based on an undetermined exemplar hand (*The Wenche that Loved a King*). The same argument holds for the second exemplar hand, since it appears in the middle of the first booklet (*Seynt Mergete*) but also opens the eighth and ninth booklets (*King Alisaunder* and *Sir Tristrem*, respectively). It can be appreciated from Table 5.2 that only Scribe 1 among the six scribes alternated between exemplar hands.

It is conspicuous that there is support for different exemplar hands for *Seynt Mergete* and *Seynt Katerine* since scholars disagree on their possibly shared authorship (Bliss 1956, Görlach 1981); that *Lay Le Freine* and *Sir Orfeo* do not share their exemplar hand since they may share their author (Pearsall and Cunningham 1977: xi); and that *Of Arthur and of Merlin* and *King Richard* do not share their exemplar hand with *King Alisaunder* since Smithers (1957: 41) suggested there is linguistic evidence for a shared London authorship for them; and that *The Anonymous Short English Metrical Chronicle* groups with other textual items (second exemplar hand), since it has been proposed that Scribe 1 actively revised its text (M. Fisher 2012: 146–87). Scribe 1 appears on at least one other occasion to have largely resisted this revising impulse, for he transmitted *The Four Foes of Mankind* (unknown exemplar hand) “in a form which leaves no doubt as to its northern origin. Indeed the amount of change which it underwent at his hands is so minor that in its final form it must still have had as alien a linguistic flavour to a reader or listener in the London of the time as
did the speech of the northern clerks in Chaucer’s Reeve’s Tale later in the century” (McIntosh 1978: 138).

3 The immediate exemplars for the Auchinleck manuscript

Codicologists agree that Scribe 1 played a key role in the production of the Auchinleck manuscript. He may have acted as stacionarius coordinating the work of all the scribes and interacting with the customer, as the manuscript was definitely compiled to order. The evidence not only is that he copied almost three-fourths of the total text (72 percent) but also that his hand appears outside the stints copied by himself; specifically, he is responsible for supplying paraphs and quire signatures throughout most of the manuscript, as well as for ruling the folios in the case of Scribe 2’s The Sayings, as previously mentioned. He also furnished the textual items with item numbers and titles after the manuscript was illuminated. Add to this evidence that the manuscript is characterised by a general uniformity of layout with double columns and a fixed number of lines to the page. Such uniformity bespeaks some measure of advance planning, a supervisory presence (in the shape of Scribe 1), and ready access to exemplars for the textual items.

Disagreement has pivoted around the exact nature of the scribes’ cooperation and continues to do so, for there are several interruptions to the general scheme. “[M]aybe the organisation was not very sophisticated”, Pearsall and Cunningham comment (1977: ix). The order of copying is not settled; the four paraphers may have received the text of the manuscript piecemeal, in scribal stints or booklets (Shonk 2016); and someone other than Scribe 1 appears to have introduced corrections throughout the manuscript, including in Scribe 1’s stints (Vaughan 2016).

One view, now abandoned in its strongest form, considers the Auchinleck manuscript as an example of booklet production. This mode of production would have entailed the various scribes producing booklets practically independently of each other at the behest of Scribe 1. One would, consequently, expect of a new scribe for him to have started a new textual item on the first recto of a new booklet and for this booklet to have a separate ruling pattern or be free-standing in other ways, such as by ending in a blank. The scribes would probably have worked on them each in their own workshop, before returning them to Scribe 1 for him to assemble the manuscript. He would have copied shorter items into the blanks at the ends of booklets at the time of assembly so as to fill the booklets up and create smooth transitions between them.
Booklet production of any manuscript gives the assembler flexibility in deciding the order of already copied materials and the possibility to include new material as it becomes available. It is theoretically possible in this production model for one or more booklets to have been copied prospectively, for several scribes each to have worked simultaneously on a booklet, and for exemplars to have been received piecemeal, resulting in a protracted copying process—in fact, there is good reason to think it was regular practice among scribes also to produce and stock exemplars as booklets (Hanna 1986, 1996; cf. Robinson 1980, Gillespie 2011).

The reason that this view is no longer considered entirely valid is that the Auchinleck manuscript does not fully meet the criteria, especially in relation to the roles of the six scribes. Heavy loss of folios hampers complete analysis but it can be confidently stated that scribal stints, textual boundaries, and quire/booklet boundaries do not always coincide. Codicologists recognise twelve booklets of between one and nine quires each with a regular quire size of eight folios. The second booklet spans quires 7–10. Scribe 2 opens this booklet and the stacionarius Scribe 1 concludes it, as expected in the booklet production model; but Scribe 2’s stint, the Speculum Guy, ends already in quire 8 and what follows it, Amis and Amiloun, is not a short item helping to fill up a quire. The third booklet is started by Scribe 3 in the expected manner and does end in short items, only those items, The Sayings and the Battle Abbey Roll, are not supplied by Scribe 1 as one would expect but rather by respectively Scribes 2 and 4, who appear successively on fol. 105. A third example relates to the first of Scribe 5’s two textual items, which are consecutive but straddle a booklet boundary. This item finishes a booklet started by Scribe 1, the fourth booklet, as if it was Scribe 5 who was the stacionarius.

Scribe 3, then, may have worked independently. So may Scribe 6, for the entire text of Otuel, his sole stint, occupies the seventh booklet, which is made up of the ten-folio quire 38, which is the manuscript’s sole irregular quire. While Scribe 2 appears elsewhere in the manuscript as described in the preceding, it is possible that he prepared the twelfth booklet independently as it is a single-quire booklet with no other contents than The Simonie and does not have the alternating red and blue paraphs found elsewhere; however, this textual item ends imperfectly through loss of quires and it cannot be determined how the booklet ended.

Another abandoned view, this one fully abandoned, argues for readily available exemplars and closer contact between the scribes than the booklet theory allows. This view, associated with Hibbard Loomis (1942) and accepted by both Pearsall and Cunningham (1977) and Taylor (2003), holds that all the various artisans worked under one roof, including not only the scribes and illuminators but also the translators and versifiers who produced the exemplars and the binders who fixed the order of the booklets by binding the final manuscript. Professional
lay scribes, in other words, organised themselves and their working methods along the lines of monastic scriptoria. A shared location would explain the codicological support for some form of collaboration among the six scribes, and the evidence that there are frequent thematic and other links between the textual items at the level of contents, such as verbal echoes. Very many of the items are translations from French and appear uniquely or in their earliest attested form in the Auchinleck manuscript. A good many of them are romances, including recently composed ones. It is unlikely that versions of all of them were simultaneously available anywhere else in England than London, be it in Parisian French, Norman French, or English.

It is widely, if not universally, accepted today that there were no organised scriptoria in London in the early fifteenth century, some two or three generations later. The basis for this view is a codicological examination of Trinity College, Cambridge, MS R.3.2 of John Gower’s *Confessio Amantis* conducted by Doyle and Parkes (1978). The five scribes who each contributed booklets to that manuscript variously entered into joint ventures with other scribes and book artisans for the production of other manuscripts, suggesting the norm was *ad hoc* collaboration among them. The model is known as distribution copying. Each artisan plied his trade out of his joint living and working quarters, which were situated in close proximity to those of his peers but were none the less separate from them, except an apprentice could work in the same workshop as his master. Muniments relating to Paternoster Row and adjoining streets confirm such living and working arrangements for some book artisans and suggest the quarters’ not-very-generous physical dimensions (Christianson 1989, 1990). This neighbourhood abutted the northern wall of St Paul’s Cathedral but was not the sole locus of scribal activity in late medieval London. It tallies well with the distribution copying model that the Auchinleck manuscript’s Scribe 3 appears to have worked more independently of its Scribe 1 than, for example, its Scribe 2. And it tallies well with the model that the manuscript’s Scribe 6 may in fact be its Scribe 1 copying a separate booklet at a moment in time before the manuscript was even conceived.

The absence of organised lay scriptoria two or three generations after the Auchinleck manuscript was produced starkly contrasts with the view of it as emanating from one and is why this view of its making is no longer current. However,

---

6 The six scribes of the Auchinleck manuscript do not reappear in other manuscripts, except that Scribe 2 possibly furnished British Library, London, MS Egerton 1993 with parahps (Marshall 2010). There are more certain links to other London manuscripts in the decoration. The four surviving miniatures are in the same style found in British Library, London, MS Royal 2 B VII (the Queen Mary Psalter) and may or may not be by the same artist (Dennison 1990; cf. Shonk 2016: 179).
later scholarship such as Pearsall (2016: 12) has tended slightly to misrepresent Hibbard Loomis’s pioneering paper, perhaps because the misleading term “bookshop” appears in its title. The paper, which was written during World War II when less was known about medieval book production, in fact carefully avoids committing to any specification of the shared location as a single scriptorium, entire neighbourhood, or something in between. The specific term “bookshop” is introduced “[f]or convenience [to mean a] hypothetical lay center where went on, whether under one roof or not, the necessarily unified and directed work of compiling, copying, illuminating, and binding any book” (1942, 597; my emphasis). There may, then, on balance, have existed a community of independent lay professionals akin or identical to the one centred on Paternoster Row already in early to mid-fourteenth century London. These professionals did not work together in as unified and directed a fashion as Hibbard Loomis envisioned. They were competitors on the market working together on an ad hoc basis, and one such collaborative effort resulted in the Auchinleck manuscript (Shonk 1983, 1985, 2016). This conclusion challenges the notion that London emerged late in the Middle English period as a centre of book production in comparison with the South-West Midlands, East Anglia, and Yorkshire.

It is other, later passages than the one cited which make clear that Hibbard Loomis held translation and versification of texts to have been routine tasks for “bookshops”. She further held that the artisans who engaged in these activities with a view to the Auchinleck manuscript introduced modifications to the textual items in the process so as to integrate them with one another. Pearsall and Cunningham (1977: ix–xi) follow suit. However, the many verbal echoes advanced by Hibbard Loomis as evidence of the integrating modifications are a universal characteristic of the romance genre and do not in themselves provide good evidence of close contact between the individuals engaged in the translation and versification activities (cf. Wiggins 2002: 98–102). Parallels of many of the textual items occur individually in other manuscripts and include the verbal echoes. In addition, both Shonk (1981: 34–35) and Mordkoff (1981) argue for translation and versification to have preceded the production of the Auchinleck manuscript.7

The distance, stemmatically and chronologically, between the translator-versifiers and the Auchinleck scribes must none the less have been short given that the Auchinleck copy is typically or always the earliest known one. Pearsall (2016) operates in the space between the root of the stemma for the English-language version and the immediate exemplars for the Auchinleck copy of it. He

7 Olson (2012: 101, 103) appears to accept the idea that the translator-versifiers may be identical to the Auchinleck scribes, including Scribe 1 for rewriting Guy of Warwick (stanzas).
argues for exemplars—meaning a copy falling somewhere in that space—for some textual items to have been put into circulation by their authors, others to have been specifically commissioned for the compilation of the Auchinleck manuscript, and still others to be the work of minstrels-cum-authors. The metrics presented in Table 5.2 cannot situate in time and space the production of what was at the root of the stemma for the English-language version of any of the individual textual items, as they measure orthographic similarity between the immediate exemplars.

Scribe 1 of the Auchinleck manuscript somehow enjoyed continual access to exemplars written in a mere four hands, although they contained no less than 23 textual items. His access to them must have been continual, since he was able to alternate between the hands. The alternation means either (a) that he obtained the exemplars piecemeal, textual item by textual item, or (b) that exemplars for all the textual items were available to him from the outset. The non-consecutive progress of copying rules out the latter explanation, and the low number of hands strongly suggests nearby sources. So, the likelihood is that the community of book artisans probably did not translate or versify any of the textual items present in the Auchinleck manuscript, but it did copy out the text of these items when they came to hand so as to produce exemplars for them and these exemplars it stocked. The community will have adopted these practices for the production of other manuscripts too. The various scribes may each have kept a repository of their own in their separate workshop and have exchanged exemplars with each other on an as-needed basis. If so, there is no guarantee that exemplars for every textual item necessarily passed through the hands of the stacionarius coordinating the scribal work on a manuscript. A scribe commissioned independently to copy one or more textual items into a booklet may have drawn on his own repository or that of a nearby peer for the exemplars for them, like Scribes 3 and 6 of the Auchinleck manuscript possibly did. An exchange of exemplars may plausibly have taken place against payment but there does not survive any evidence in that regard.

4 Type II as an incipient standard

Samuels (1963 [1989: 71]) posited four “incipient standards”, which he respectively labelled Types I–IV. One of them, Type I, is not London-based and will be ignored in the present discussion. The other three are London-based and follow each other chronologically. His article repeatedly stressed that there is considerable variation in form within the types, in particular within Type III about which he remarked
(1963 [1989: 71]) that “any form of written standard is conspicuous by its absence”, for a fully codified standard language will possess minimum variation in form (Haugen 1966). Samuels thus repeatedly emphasised the incipiency of the types in his article and is reputed to have continued doing so orally throughout his professional life. Much later scholarship has none the less tended to play the incipiency down and has instead portrayed the types as fulfilling more of the criteria for a standard language than Samuels intended and than the evidence warrants. But, what is the actual evidence for Type II?

Samuels (1963 [1989: 70]) advised that the forms of Scribe 1 of the Auchinleck manuscript “may be taken as typical of Type II” and in a footnote listed those of Scribes 1 and 3 as well as six scribes of other manuscripts as defining the type – this association of Scribe 3 with a standardising variety would have been surprising a generation early, for Brunner (1933) held him obviously to be a Norman not fluent in English.\(^8\) Hanna (2005: 14–15) updated the list with a few more scribes and manuscripts identified by himself and other scholars. Scribes 1 and 3 are the earliest of them, by as much as half a century in the majority of cases, although allowance must be made for the fact that the vagaries of time have not ensured the survival of much relevant material in fourteenth-century English.

Haugen’s discussion implies that selection of orthographic forms to make the basis for a future standard must occur early in the standardisation process (cf. Ayres-Bennett 1994: 55), and the basis for connecting the Type II scribes is a shared core of orthographic forms linguistically localisable to Essex. Internal migration may have brought about a change in the kinds of orthographic forms in use in London, a shift in dialectal composition, that saw Type III succeed Type II. This is what Samuels (1963) proposed to account for the incongruity of Types II and III, for the texts manifesting the latter share a core of orthographic forms linguistically localisable to centrally in the Midlands,\(^9\) all the while that the manuscripts housing them too were produced in London. His article offered the flimsy evidence of fewer than twenty contrastive orthographic forms for a mere twelve items in support of a distinction between Types II and III, which it backed up by stressing how the forms did not exhaust the evidence but were the best representatives. Samuels further held that the shift must have been

---

\(^8\) Runde (2016: 71–72; cf. Runde 2010) notes how Scribe 3 is just as comfortable with English as Scribe 1 and so cannot have been a Norman and, more importantly, how he must have been a translating scribe rather than a \textit{literatim} copying one since his orthography is consistent throughout his stint.

\(^9\) Some literature capitalises the word “central” and speak of localisation to the Central Midlands. However, there is no geographic entity or traditional dialect area known as the Central Midlands.
sudden, for the manuscripts manifesting Type II all date from before 1380, whereas those manifesting Type III all date from after that year.

However, there are issues with this narrative. Firstly, there is no strong correlation between these geographical and linguistic localisations, for Ekwall’s (1956) survey of toponymic surnames mentioned in London records does not endorse any above-average volume of immigrants from central parts of the Midlands – Wright discusses the perceptual Midlands origin of Standard English in greater detail in Chapter 1. Secondly, Chaucerian verse copied by a single scribe, known as Scribe B, is over-represented in the manuscripts attesting Type III, which makes it a possibility that it is genre that is the strongest of the variables predicting linguistic localisation in the case of Type III.10 The Type II manuscripts similarly share their genre and in part also their textual items (Hanna 2005: 7).11 Type IV texts are also united by genre, and where it is end-rhymed verse that Types II and III tend to record, the chief evidence for Type IV is documents.12 Thirdly,

10 Genre may be the strongest grouping variable for the Type III items, but the argument that the defining manuscripts contain little else but Chaucerian verse copied a single scribe fails to convince. It rests on Horobin and Mooney’s (2004) assignment of Trinity College, Cambridge, MS B.15.17 of William Langland’s *Piers Plowman* to Doyle and Parkes’ (1978) Scribe B, the same scribe who copied National Library of Wales, Aberystwyth, MS Peniarth 392D (‘Hengwrt’) of Geoffrey Chaucer’s *Canterbury Tales* and Huntington Library, San Marino, MS El.26.c.9 (‘Ellesmere’) of the same poem. They bolster up their paleographical argument by fitting the *Plowman* manuscript into the chronology of Scribe B’s manuscripts established by Samuels (1983) from orthographic forms. It comes first. The orthographic forms, however, are ‘long’ and ‘short’ ones such as unabbreviated versus abbreviated THAT, WITH, PER-/PAR-, and PRO-, and glyphs of the grapheme ‘h’ with and without a cross-stroke. It can be demonstrated statistically that Scribe B selects ‘short’ forms when he is pressed for time and space (Thaisen 2011), and so it is the manuscripts’ production time and physical format (spatiality) that are the salient variables rather than Scribe B’s chronological age. In addition, since it is required of a standard that it transcends genres, it may be important that Samuels’ list of Type III items includes documents in addition to literary texts. The documents are unspecified ones from Chambers and Daunt (1931), presumably Thomas Usk’s appeal against John Northampton, Nicholas Brembre’s proclamations, and various London guild returns and wills.

11 Probabilistic modelling presupposes that all training and test text is transcribed according to the same protocol, or the perplexity metrics will be influenced by differences in transcription. This is a methodological limitation and a reason that the paper does not consider non-Auchinleck orthographic forms.

12 Type IV is dubbed “Chancery Standard” after Samuels (1963) but it was no standard employed by the Chancery (Benskin 2004). The first half of the label does not denote the entire national (royal) administration at Westminster as J. Fisher (1977) and J. Fisher et al (1984) appeared to maintain but rather a single department within it. Benskin (2004) emphasised that this department primarily issued documents in Latin and that the first departments to use English, or rather re-introduce this language after the Norman Conquest, were more properly
orthographic forms characteristics of Type II persist long after 1380 (Horobin 2003) and are found alongside Type III forms (and even Type IV ones), indicating a more gradual chronological shift than Samuels allows (Horobin 2003; Hanna 2005).

Fourthly, there are examples of London-based authors and scribes who did not select any of the types as their target norm. One of them is the poet John Gower whose orthographic forms have been linguistically localised partly to Kent, partly to Suffolk (cf. Horobin 2003, Mooney and Horobin 2003). Medieval London was a dialectal melting pot or, perhaps rather, patchwork quilt so the presence of a language user whose orthographic forms disagreed with any standardising tendency is not in itself evidence of absent standardisation. However, the examples of such users include scribes demonstrably in close contact with the users of Samuels’ Types II and III. Scribe 2 of the Auchinleck manuscript is one, for his orthographic forms linguistically localise him on the Gloucestershire/Worcestershire border (LP 6940) and its Scribe 6 may be another – see below. Yet another is the scribe of both Corpus Christi College, Oxford, MS 198 and British Library, London, MS Harley 7334, both of which are manuscripts of Geoffrey Chaucer’s Canterbury Tales. This scribe is known as Scribe D after Doyle and Parkes (1978) since he is the fourth of the five hands found in the distribution-copied Trinity College, Cambridge, MS R.3.2 of John Gower’s Confessio Amantis. Scribe B is the second of those hands, but despite them knowing each other, appearing together in at least one manuscript, and copying similar text-types by the same author, Scribe D’s orthographic forms are markedly different from those of Scribe B.

Fifthly, the level of similarity between the orthographic choices made by Scribes 1 and 3 of the Auchinleck manuscript does not in fact exceed that, established by the present analysis, between their choices and those made by the other the Privy Seal and Signet offices. Moreover, the Chancery cannot have been exclusively situated “at Westminster” because its staff numbers, upwards of 100, certainly exceeded the number of desks that could fit in Westminster Hall. It is rather the case that many clerks had their quarters elsewhere in the London-Westminster area. It has, further, become clear that royal clerks, whether attached to the Chancery or another department, frequently interacted with artisans engaged in the local book trade and sometimes took on writing tasks for other clients. Some were hired on a temporary basis. The clerks’ numbers, geographical dispersion within the general London area, and possible temporary attachment to the Chancery cannot have been conducive to anyone imposing norms on them for how to spell in English, let alone for a norm to have developed by more natural means within this specific institution. In line with this contextualisation of Type IV, which is pursued in much greater detail by Stenroos in Chapter 2, it resembles Types II and III in representing a single genre, and Wright (1996) has called attention to how a single-genre variety needs elaboration to be able to satisfy the requirements of a standard that can adequately serve many functions.
scribes. In particular, Scribe 3’s 200-line segments do not perplex the models trained on Scribe 1’s 200-line segments markedly less than Scribe 5’s or Scribe 6’s do, nor do Scribe 1’s 200-line segments perplex the models trained on Scribe 3’s 200-line segments markedly less than Scribe 5’s or Scribe 6’s do. Scribe 5’s segments in fact perplex them the least, albeit they do so by a narrow margin (not shown). Such comparable levels of similarity would be expected only if Scribes 5 and 6 also belong to the corpus of Type II scribes. There exists support for this possibility, for (a) *A Linguistic Atlas of Late Medieval English* (McIntosh et al., 1986), of which Samuels was one of the editors, linguistically localised the profiles for Scribes 1 (LP 6510), 3 (LP 6500), and 5 (LP 6350) adjacent to one another in the general London area; and (b) Hanna (2000, 2016) proposed on paleographical criteria that Scribe 6 is identical with Scribe 1, the differences being attributable to a chronological gap between the stints. On the other hand, the linguistic profile for Scribe 6 (LP 7820) belongs in Worcestershire, the perplexity distribution does discriminate Scribes 1 and 6, Wiggins’ meticulous comparison of a selected range of orthographic forms convincingly speaks against Scribes 1 and 6 being identical (Wiggins 2004), and Scribe 6’s stint is codicologically independent in the manner of Scribe 3’s.

Sixthly, any orthographic alignment of Scribes 1 and 3 is unexpected from a codicological point of view. The reason is that the codicology does not suggest that they worked closely together, whereas it does suggest that Scribe 1 worked especially closely with Scribe 2 (cf. Shonk 1983, 1985, 2016), whose linguistic localisation disagrees with his. As Hanna (2016: 217) describes it, Scribe 3 is the only one of Scribes 2–6 “not demonstrably in touch with Scribe 1. Not only is Scribe 3’s hand isolated in the book as representing documentary rather than formal...”

13 While Bliss (1951) was the first to propose identity between Scribes 1 and 6, it is Hanna (2000, 2016) who has developed the argument. The principal paleographical argument for identity between them is that both scribes employ biting, since biting is a rare feature in early fourteenth-century Textura. Hanna explains other paleographical differences between them as differences in duct typical of stints produced at different times and for different purposes. The same explanation solves the paradox that the seventh booklet with *Otuel* is written in a western dialect, whereas all other text by Scribe 1 is written in an eastern dialect. Because the booklet was prepared independently, there was no need for its linguistic integration with the rest of the manuscript. Besides, examples of texts with western localisations such as those mentioned in this paper show that such texts were acceptable to a London clientele. *Otuel* could, therefore, be *literatim*-copied from western exemplars. Pearsall and Cunningham (1977: x) find “on the whole convincing” the argument (not discussed by Hanna) from literary clues that *Otuel* and the textual item preceding it, *Roland and Vernagu* may together represent a reworking within the Auchinleck “bookshop” of a single ancestral romance. *Roland and Vernagu* answers to the Scribe 1 first exemplar hand.
training; he appears uniquely estranged from the universal format of the book”. It will be recalled that Scribe 3 copied only the third booklet, which is concluded by Scribes 2 and 4 in succession and in which Scribe 1 does not appear. It will also be recalled that the models of Scribe 3’s first five 200-line segments have “odd” perplexities; it is these segments which do not adopt the general format of the manuscript, implying an absence of Scribe 1’s controlling hand.

Seventhly, there is no support in the perplexity distribution for coincident exemplar hands which could have conditioned Scribes 1’s and 3’s orthographic forms in the direction of convergence. It is a possibility that Scribes 1 and 3 could have copied from exemplars prepared by one and the same hand in view of how the immediate exemplars were locally produced and locally obtained, even if the present analysis has shown the various texts primarily to record the respective scribes’ own forms. There might even, at least theoretically, have been identity between an exemplar hand and an Auchinleck scribe. However, it can be stated with confidence that any support is absent, since no pairing of models and segments corresponding to a Scribe 1 exemplar hand with those corresponding to a Scribe 3 exemplar hand is associated with an appreciably lower perplexity than any other such pairing, including the textual items analysed by the Atlas.¹⁴ Those items were Seynt Mergrete, Seynt Katerine, Guy of Warwick (couplets), Sir Orfeo, and The Anonymous Short English Metrical Chronicle for Scribe 1 and The Seven Sages of Rome for Scribe 3. Three of the items in the list for Scribe 1 answer to the second exemplar hand that this scribe drew upon according to the results of the present analysis. The first and fourth exemplar hands are represented by one textual item each, and the third exemplar hand by none. The sole item in the list for Scribe 3 answers to the second of the three exemplar hands distinguished for this scribe.

5 Conclusion

Summing up, Samuels investigated texts that certainly are geographically localisable to London since they are housed in manuscripts known on extralinguistic criteria to have been produced there. He was probably in search of reference

¹⁴ The two Scribe 5 exemplar hands do likewise not differ noteworthily in perplexity relative to the Scribes 1 or 3 exemplar hands. In addition, the Atlas lists as the basis for Scribe 5’s LP 6350 the title of Sir Beves only but the folios for both Reinbroun (fols. 167ff) and Sir Beves (fols. 176ff). If Reinbroun is included, both exemplar hands discriminated for Scribe 5 are presumably represented in the profile.
texts for the London area for A Linguistic Atlas of Late Medieval English, the compilation of which was still far from completed in the early 1960s. Samuels was no doubt aware of how London was a meeting place for texts with all manner of incompatible linguistic localisations, such as Gower’s and Scribe D’s, but he noticed a tendency for some of the texts to group in respect of the orthographic forms they use for a dozen items: he posited a first class, Type II, the members of which share a linguistic localisation of those orthographic forms to Essex and pre-date 1380, and a second class, Type III, whose members post-date 1380 and share a linguistic localisation to centrally in the Midlands.

Later scholarship has suggested from the membership of not only Types II and III but also Type IV that it is genre that is the true predictor variable that gives rise to their separate linguistic localisations, rather than forms gaining wider currency in the London area and scribes targeting them. Later scholarship has additionally noticed that there exist still other London-produced texts that each question the basis for Types II and III by containing forms characteristic of both and not observing the 1380 boundary date.

If it is unclear or undetermined what variables truly unite the Type II texts so is their relationship with standardisation, apart from their extra-linguistic association with London as the presumed locus of the standardisation process. At the chronological beginning of this process, Haugen’s selection and codification stages are tantamount to scribes and other writers beginning to converge in their orthographical representation of lexical items. Whether this tendency toward convergence is better termed “focused variation”, “incipient standardisation”, or “purging of grosser provincialisms”, it is not in evidence in how perplexity is distributed in the Auchinleck manuscript, for the distribution does not group the two Type II scribes against the non-Type II scribes considered separately or in combination. The distribution at best lines up Scribes 1, 3, 5, and 6 against Scribe 2 (with Scribe 4 and his Battle Abbey Roll yielding no data) in discrepancy with their respective linguistic localisations.

Middle English is characterised by an unprecedented amount of orthographic variation. A need for scholars to have fixed reference points when addressing this variation might explain why Samuels’ types continue to prove remarkably robust, despite evidence undermining them as incipient standards having mounted over the past one or two decades. The evidence presented in this paper takes, I believe, the discussion past the point of no return. The types may be conclusively laid to rest.

Last, the paper has discriminated scribes at the levels of both the Auchinleck manuscript and its immediate exemplars, and by this means shown that Scribe 1 copied more than twenty textual items from exemplars written in a mere four hands. This finding aligns well with the proposed existence of a community of
book artisans in early to mid-fourteenth century London whose collaboration extended to them exchanging exemplars with one another. It adds to previous scholarship that this community did not simply store exemplars obtained from elsewhere but must also have produced them.

References

Primary sources


Secondary sources


Bliss, A. J., ‘The Auchinleck “St Margaret” and “St Katherine”’, *Notes and Queries* 201 (1956), 186–88.


