A LINGUISTIC ATLAS OF EARLY MIDDLE ENGLISH

INTRODUCTION

PART I: BACKGROUND

CHAPTER 1

PRELIMINARIES

Margaret Laing and Roger Lass

1. Background[1]

> There are no such things as dialects. Or rather, 'a dialect' does not exist as a discrete entity. Attempts to delimit a dialect by topographical, political or administrative boundaries ignore the obvious fact that within any such boundaries there will be variation for some features, while other variants will cross the borders (Laing and Lass 2006: 417).

1.1 Historical Dialectology

A dialect atlas is, at least in part, a set of maps showing the distribution of linguistic features in space.[2] Whether modern or historical, it is not made up, however, of static displays of dots or boundary markers on regional maps. Rather, it shows a continuum of overlapping distributions, where 'dialects' are assemblages of regional features that vary from map to map both spatially and temporally.[3] This is because language change, like biological change, occurs differentially in different spatial settings. Change over time involves the vagaries of language use and thus introduces a necessary social dimension. Dialectology therefore operates on three planes: space, time and speech community.

We have no direct record of spoken language before the development of sound recording in the 19th century. For any earlier period our only source material is therefore written language. In other words, the 'native speakers' of past stages of a language are writers and copying scribes. For a medieval historical atlas therefore, 'text languages' (Fleischman 2000) must take the place of the live informants of a modern dialect survey. We cannot question our sources directly nor ask them for data that do not already appear in their surviving output. These are the three potential operational problems: (a) we have to work with the written rather than the spoken language; (b) our data are limited by the contingent survival of texts; (c) our knowledge of the personal details of the producer of any particular text is limited: medieval scribes are mostly anonymous. All we know is that they were literate and, at our early Middle English period, mostly adult male Catholics.[4]

1.2 Early Middle English

---

[1] In this chapter, we have not attempted to make a complete restatement of all the material already published in *A Linguistic Atlas of Late Mediæval English* (*LALME*) or its associated publications. Nor do we recap in detail the already published papers that contribute to an understanding of how *LAEME* fits into the tradition established by *LALME*. This chapter is a summary of the most important contextualising information for the rest of the Introduction and Manual. Parts of this chapter, however, are closely based on material from Laing (1993: 1–10 and 2000) and Laing and Lass (2006), where some of the theoretical issues are treated in more detail. For an informal account of the making of LAEME see Preface and Acknowledgements.

[2] This atlas (see further Part II) also contains a tagged corpus, concordancing software and other non-cartographic utilities.

[3] For some discussion and further references see Benskin 1981 and 1994: 169–173.

[4] Other variables may in some cases be recoverable from clues in the manuscripts or their contents: e.g. affiliation to a particular religious order and sometimes association with a specific house or scriptorium. This will only be true of productions from religious houses. For our period there will also have been non-monastic writers trained in town schools and universities. Other than the fact that university books tend to be written in particular types of hand, extra-linguistic information enabling us to associate non-monastic manuscripts with particular local backgrounds is in general lacking.

By early Middle English we mean all written English from about 1150–1325, after which the language may be termed 'late Middle English'. *A Linguistic Atlas of Early Middle English* (*LAEME*). is designed to cover the English that was written during the three or four generations before the period already covered in *A Linguistic Atlas of Late Mediæval English* (*LALME*). It aims, as far as the surviving material allows, to capture, display and analyse the written dialect continuum of this earlier stage in the language.[5]

How do we delimit the period 'early Middle English'? Periodisation of languages is to a large extent arbitrary:[6]

> The 'early'/'late' division is largely ... a matter of convenience. The language of the 12th century and that of the 15th are different enough to justify separate identifying names. But language differences in time, like those in space, form continua: there are no sharp temporal boundaries. With respect to 'archaism' and 'modernness' Middle English before 1300 is certainly 'early' and that after 1350 'late'; the language written between is perhaps 'transitional', but this property exists only by virtue of the strong differences at both ends (Laing and Lass 2006: 418–19).

In the case of English, however, there is one very clear break in its written history, which enables selection of a *terminus a quo*. This break was one of the consequences of the Norman Conquest in 1066. Before the Conquest, the Anglo-Saxons, typically for Germanic rather than Romance cultures, used the vernacular (alongside Latin) to record their legal and administrative documents, as well as for religious, historical and literary works. Thereafter, the conquerors imposed their own native practice, Latin soon replacing English as the language of government[7] and of literature. The settlers also introduced their own vernacular literature. Alongside Latin, French became established as the other high status written language.

For *LAEME*, the most important effect of this disruption in the use of written English is the relative paucity of surviving material for our period compared with the more extensive sources available for *LALME*. It is likely in any case that more written texts will have been lost from a period so much earlier, but this problem is exacerbated by the fact that English was scarcely used as a written medium for the century immediately following the Conquest, and thereafter only gradually became reinstated. Until well on in our period, much of what does remain may be short or fragmentary texts written in manuscripts containing mainly Latin and/or French.[8]

However, writing in English did not cease completely after the Conquest, since Old English texts continued to be copied and studied (and in the case of legal documents used) as historical objects themselves. Some of these copies, from before 1150, already sporadically display linguistic features we associate rather with Middle English than with Old English.[9] However, they are still recognisable as versions of an

---

[5] The period covered by *LALME* is roughly 1350–1450 in the Midlands and the North and 1325–1425 in the South. For the southern area of survey *LALME* also includes several texts from the late 13th and from the early 14th century, so some useful overlap between *LAEME* and *LALME* is assured.

[6] The choice of fixed dates to delimit a period of language is in any case problematic, since linguistic change is a continuous process (cf. Lass 2000).

[7] Some post-Conquest documents were written wholly or partly in English (Pelteret 1990) but their language is largely based on earlier documentary models.

[8] On the post-Conquest linguistic situation and the changing relative status of English, French and Latin including bilingualism and trilingualism, see Berndt 1965; Brooks 1982; Clanchy 1979; Clark 1992; Davis 1976; Hunt1991; Richter 1979; Rothwell 1968, 1975–76, 1978, 1983; Shelley 1921; Short 1980, 1992, 1996; Wilson 1943; Woodbine 1943.

[9] For studies of some of these changes see Malone 1930. For a recent study on English in the transitional period between Old and Middle English see D. Scragg *et al.*, *An Inventory of Script and*

older stage of the language rather than as something that their copyists would have written spontaneously. There was apparently very little new composition in English in the century after the Conquest, apart from a few late additions to the Anglo-Saxon Chronicle. It is only towards the end of the 12th century that new writings in English begin to be attested in any quantity. These texts are linguistically clearly distinct from the copied Old English ones, indicating that the spoken language, as we might have expected, had continued to change throughout the post-Conquest period.

1.3 Time-span and coverage

We take the work of the scribe of the second continuation of the Peterborough Chronicle, writing in 1154/5, as our *terminus a quo* for early Middle English. This text represents the earliest surviving example of truly Middle English language — that is, it reflects how the spoken language of the scribe's region of origin had developed in the preceding century. After this there is a gap in attestation: there are very few other surviving texts in newly composed English that can be firmly dated earlier than the last quarter of the 12th century, most of these being 'ca 1200'. Often manuscripts can be dated only from palaeographic evidence and this method is always approximate. Dates judged from script and decoration are usually given as lying somewhere within a quarter-century range. Were it not for the survival of the Peterborough Chronicle second continuation, we could reasonably say that our *terminus a quo* would be ca 1200. This is relevant to the selection of our *terminus ad quem*.

The *terminus ad quem* is less easily determinable — we do not have a last text in the same way as we have a first text. In addition, the existence of *LALME* creates certain responsibilities: we would like the two atlases to give as coherent a picture as possible of Middle English 'as a whole' without excessive duplication. A linguistic atlas displays how different forms of language change through time and how they vary across space. This leads to a further consideration: selection of too broad a time-span for a dialectal survey makes it more difficult to assess the relationship between diachronic change and regional variation. This is because over an extended period, the results of linguistic change are likely to present us with an intractably large typological range. On the other hand, choice of too narrow a time-span incurs the risk that the surviving texts will not provide sufficiently dense geographical coverage to constitute a continuum.

For *LALME*, a period of 100 years was considered optimal. For *LAEME*, in practice it is necessary to select a rather longer time-span[10] to allow maximal regional coverage in a period from which fewer texts survived. In other words, wider temporal extension is a heuristic that increases our chances of encountering written material originating from different regions, which is especially important for early Middle English when the survival of texts is geographically very patchy. Moreover, dependence on data from an accidentally surviving sample means that our optimal time-span is not the same across the country. Even in *LALME* the 100-year period was different for the two areas of survey, North and South:

Spellings in Eleventh-Century English http://www.arts.manchester.ac.uk/mancass/C11database/. The production and use of written texts in English between 1060 and 1220 is now the subject of an AHRC-funded project at the Universities of Leeds and Leicester, directed by Mary Swan and Elaine Treharne. See further http://www.le.ac.uk/english/em1060to1220/index.htm.

[10] This greater extent is in part illusory, as the Peterborough Chronicle second continuation in effect occupies a 25-year band on its own, so the 1150–1325 we give as our time-span is not in fact continuous.

> From before about 1350 there are very few sources for northern or North Midland English, so that, for those areas at least, the *Atlas* can go back no earlier; but for certain parts of southern England the evidence from texts of the earlier fourteenth century is so valuable that they could hardly be excluded, and the same applies even to a few late-thirteenth-century texts... Conversely, since the spread of written Standard English was earlier in the south than in the north, dialectal texts in the south become rare at a correspondingly earlier date. For the south, therefore, the 'core' evidence here utilised should be regarded as falling within the period *ca.* 1325–1425, rather than that *ca.* 1350–1450 which applies for most of the Midlands and the north (*LALME* I, §1.1.2).

If *LAEME* and *LALME* together constitute an 'atlas of Middle English', in principle we would want to minimise the chronological gap between the *LAEME* sources and the *LALME* sources. In practice, this means that we too have adopted different *termini* for different parts of the country. Because of the small numbers of surviving texts of northern and North-Midland origins from pre-1350, we have included the few early texts that do survive, even if they are as late as the first half of the 14th century. In the southern half of England, survival of early texts is also uneven. There is no representation for the Central Midlands, very little for the true South, rather more for the East Midlands and relatively dense coverage for the South West Midlands. For the southern area of survey, the inclusion of some late-13th-century texts in *LALME* means that a *terminus ad quem* of 1300 for *LAEME* allows for sufficient overlap, with some late-13th-century texts appearing in both atlases.[11]

One text that is an apparent exception to our usual procedure is *Ayenbite of Inwyt*. Unusually for a medieval text, a colophon in the hand of its scribe gives his name (Dan Michel), origins (Canterbury) and the date of the text's completion (1340). There is very little early Middle English surviving from the South East, but Dan Michel himself refers to his work as being in *engliss of kent*. We know that he was probably an old man in 1340[12] and can therefore take the language of his text to represent Kentish of the last quarter of the 13th century rather than of the mid 14th century. Accordingly this text appears in *LAEME* as well as in *LALME*.

1.4 LALME: Theoretical principles and praxis

In this section we outline the principles and methodology that underpinned the construction of *LALME*, and in the next section (§1.5) we identify the modifications that have been required for *LAEME* and introduce the resulting new methodology and data outputs.

1. Written language should be studied in its own right, not just as a representation of spoken language (cf. McIntosh 1956). Written language should be regarded as an autonomous linguistic system, if one with a special relation to phonology (see further chapter 2).

2. Regional dialects do not have discrete geographical boundaries. Interlectal variation normally constitutes a continuum (see above §1.1).

---

[11] Note, however, that the very different methodology employed by *LAEME* means that the linguisitic information presented is not all of a directly comparable kind and therefore amounts to more than just repetition. See further §1.5.4 and Part II, chapters 3 and 4 below.

[12] See Gradon 1979:12

3. Anchor texts and the fit-technique.[13] In surveys of modern dialects, localisation of informants is transparent: for any survey point one identifies a local witness and then plots his usage on the maps. Unfortunately, most medieval witnesses are more elusive: scribes, especially those of literary manuscripts, typically do not supply information about their local origins. As *LALME* felicitously puts it (1, §2.3.1): 'It is rather as if the compilers of a modern dialect atlas had access to any number of speakers, all willing to be interviewed but very few of whom divulged where they came from'.

Some manuscripts, however, can be localised on extra-linguistic grounds. 'Local documents' such as personal correspondence, local records or legal instruments,[14] constitute the one large body of texts whose places of origin or execution are for the most part explicitly indicated. Such texts, which are referred to as *anchor texts*, are the necessary starting point for a historical linguistic atlas. They enable the setting up of a matrix of defined localities, within which texts not carrying such information can be localised. This is possible because of a central empirical claim incorporated into the *LALME* tradition: 'it is fundamental that dialect changes in, for the most part, an orderly way over space, and that the domains of the various dialectal characteristics (again, for the most part) overlap differentially (Benskin 1981: xxx). For *LALME*, the language of local documents provided valuable data itself, but its importance was far greater as the means by which the unlocalised literary manuscripts, which formed the greater part of the *LALME* source material, could also be utilised. The methodology is known as the fit-technique. 'Fitting' is done by comparing, map by map, spellings particular to an unlocalised text with those already placed in the localised matrix. For each map, areas where those or similar spellings are *not* found are then eliminated, until (in the ideal case) only a single, well-defined location is left where the whole assemblage of spellings could plausibly occur. For exemplification and further discussion see Benskin (1991).

4. The questionnaire.[15] The traditional analytical tool for the dialectologist is the questionnaire. This involves the preselection of items, which the investigator has reason to believe are likely to be dialectal discriminants. 'The term *item* [is] used to denote the heading for a collection of different *forms* that are regarded as equivalent in function and/or meaning, and may therefore, potentially at least, differentiate dialects' (*LALME* 1, §2.1.1). Constructing a suitable questionnaire out of these functional equivalents is a matter in the first instance of informed experimentation. Ideally it must be large and varied enough in the range of linguistic phenomena it elicits to provide a conspectus of the language under investigation. Choice of a questionnaire item requires not only that its forms show variation but also that it occurs in most of the source texts. The combination of formal variation and frequency of attestation results in what we might call the discriminatory yield. Maximal discriminatory yield may be elicited by different items in different regions. It is for this reason that the two areas of survey in *LALME* used somewhat different,[16]

---

[13] See further *LALME* 1, §2.3.

[14] For a full list of the kinds of texts covered by the term 'local documents' see *LALME* I, §2.3.2.

[15] See further *LALME* 1, §2.1.

[16] For instance, the 'southern' item KIND *etc* which elicited spellings of OE *-ynd* words in order to capture *-e-/-u-/-i-* regional variants, was not collected in the northern area of survey because only *-i-* would be expected. Conversely, the item A, O and others that contain OE *ā* were not collected for the southern region because at this period variation between *-a-* and *-o-* spellings would not be expected there.

though (for obvious comparative reasons) partly overlapping, questionnaires. The combined questionnaire consists of 280 items, some of them subdivided, of which about a third form the common core. A completed questionnaire is processed to form a Linguistic Profile (LP) of the scribal language. If the localisation of the scribal language is known, the forms recorded for each Linguistic Profile may then be entered at the relevant place on the dialect maps. In *LALME* there are two types of map: Item Maps (vol. 2) that display actual manuscript spellings, and Dot Maps (vol 1) 'that highlight the distributions of specified variants or classes of variant' (*LALME* vol. 1: 297): e.g. 'kir-' and 'kyr-' types of the item CHURCH (*LALME* vol. 1: Dot Map 388).

For *LAEME* we elected not to use a questionnaire, so the problem of constructing differential regional questionnaires does not arise. For explanation of the alternative methodology adopted see §1.5.4 below.

5. Scribal practice. Very few surviving Middle English texts are holographs. Most by far are scribal copies at some remove from their archetypes. Before the work of the late Middle English dialect survey that resulted in *LALME*, it had been widely believed that the language of all such copied texts must represent dialect mixture. It was assumed that a copied text must be made up of a combination of elements from the authorial language and whatever linguistic forms could be ascribed to the series of copyists between the original and the surviving version in question. Early in the investigation, however, Angus McIntosh (1973 [1989: 92])[17] observed that most copied Middle English texts were in fact in language that was dialectally homogeneous: some scribes 'translate', or convert the language of their exemplar wholesale into their own variety. If a copyist is a consistent translator his language overrides that of his exemplar. However far removed the scribal language might be from that of the original, it nevertheless represents the usage of an individual and therefore has independent value as belonging to a native informant. In what might be called the '*LALME* tradition' — as opposed to more conventional forms of textual study — the emphasis is shifted away from a putative 'authorial' language to the actual language of a copy and its scribe. The output of a translating scribe is thus always useful for mapping.

Not all scribes were translators. McIntosh identified two other possible scribal copying strategies, the outputs of one of which are also potentially valuable as witnesses for linguistic geography. Aside from translating, a scribe may copy a text exactly, or he may only partially translate. The first strategy is known as literatim copying, and as long as the language of the exemplar is itself homogeneous, it will result in a homogeneous copied text language.[18] The second strategy, and any literatim copy of its results, will always result in a so-called *Mischsprache*, A true *Mischsprache* is defined as linguistic output containing two or more elements that are mutually incompatible: that is, from non-contiguous areas within the established dialect continuum (see further Benskin and Laing 1981:76–77). It is obvious therefore that a *Mischsprache* is not as such mappable. Further analysis may make it possible to

---

[17] Cf. also the account in Benskin and Laing 1981: 56.

[18] We use the term literatim copying to refer to spelling only. It need not include imitation of the script of the exemplar. Here we use it as a linguistic rather than palaeographical term, although in matters of interpretation the two of course overlap (cf. chapter 2). In the early Middle English period there are proportionally more scribes who are literatim copyists by habit or training than there are in later Middle English times. This has the useful corollary that, for early Middle English, a single scribe may provide us with several different text languages.

isolate the disparate linguistic elements and thus to map subsets of the scribal usage. In *LALME* such processed subsets were not used on the maps — there was sufficient homogeneous source material to render this unnecessary. For *LAEME*, we are not so fortunate, and for some cases we have had to devise strategies to make use of data from dialectally mixed texts (see §1.5.5 below).

1.5. LAEME

1.5.1 Preliminaries

Our survey of linguistic variation in the period 1150–1325, while being firmly in the tradition of *LALME*, has developed an independent identity. When the *LAEME* project was initiated 20 years ago, we were aware that the new investigation would have its own problems and difficulties, but we thought that the principles for constructing a medieval dialect atlas had already been established and just needed to be adapted in relation to the new data. In the event the very different nature of the available materials for early Middle English drove us to modify the established principles and to create a new conceptual and procedural model. The modifications resulted in the adoption and development of a radically new methodology and also prompted an expansion of what we thought was the original remit of an atlas.

The principles and praxis adopted in the creation of *LALME* require discussion in relation to the rather different demands of *LAEME*.

1.5.2 Comparability and coverage

Any historical linguistic survey has to address the fact that different kinds of text will prompt different linguistic choices. Pulling forms out of their linguistic context and taxonomising them by means of a questionnaire creates a comparability that irons out some of these variables. That indeed, is partly what a questionnaire is designed to do. But for historical surveys such as *LALME* and *LAEME*, the dependence on the contingent survival of texts as linguistic witnesses means that some of these differences are in practice obvious. The questionnaires completed for *LALME* were processed to form taxonomic listings for each scribe, known as Linguistic Profiles (LPs).[19] A comparison of the LPs in volume 3 of *LALME* quickly reveals that some are considerably denser in their responses than others. 'Thin' LPs, with multiple null responses to particular questionnaire items, are mostly those derived from very short texts (such as some of the otherwise crucially important documentary 'anchors'). But comparative thinness or density may also be the result of lexical and morphosyntactic choices. For instance, a past tense narrative may not have examples of items that elicit present tense verb morphology,[20] while instruction manuals may not have examples of those that elicit past tenses.[21] The result is that for some items the linguistic maps in *LALME* are variously patchy in attestation. Where the matrix of survey points is

---

[19] An LP is essentially one text language's set of answers to the questionnaire posed to it.

[20] Items in the *LALME* questionnaire that might have null entries in such a case are: ARE, IS, ART, SHALL, WILL, THINK, 3sg present indicative, present plural, CAN, DO 2sg and 3sg, GIVE, GO 2sg and 3sg, GROW, HAVE inf, 2sg, 3sg and pl, HEAR, KNOW, MAKES contr, MAY, MON, RUN, SAY 3sg and pl, SEE 3sg and pl, SEEK, TAKES contr, WITEN sg and pl, YIELD.

[21] Items in the *LALME* questionnaire that might have null entries in such a case are: WERE, WAS, SHOULD, WOULD, MIGHT, weak preterite, strong preterite plural, BEGAN TO, CAME, COULD, DO pt, GIVE pt, HANG pt, HOLD pt, LEAD pt, LET pt, READ pt, RUN pt, SAY pt, SEE pt, SPAKE/BRAKE, WENT, WITEN pt.

dense this does not have such serious consequences for our view of the continuum as a whole. But where the matrix is itself thin, lack of attestation of forms for any particular item makes for a more gapped continuum.

There are many other variables that may affect the outcomes of linguistic analysis, such as whether a text is verse or prose or whether it is a translation from Latin or from French or from an older form of English itself. The problem of comparability cannot easily be solved for a historical dialect survey of any period, but it becomes more acute the thinner the corpus of witnesses (Laing 2000a: 103). For the early Middle English period, even when all the available texts are utilised, we are still left with less than a tenth of the number of texts that were chosen for analysis from the vastly larger surviving corpus of manuscripts available to the compilers of *LALME*.

For early Middle English we have information from certain areas from early in our period and from others only from late in the period. For some regions (see §1.3) there are no witnesses at all. We therefore have considerable gaps in both the space and time continua, so that *LAEME* is patchy in its basic coverage of survey points. Complete, coherent and cohesive dialectal patterns across the whole area of survey are not to be had. In some areas, however, such as the South West Midlands, there are enough local texts for something approaching a dialect continuum to be observable. Moreover, early Middle English dialects should not be studied in isolation from later Middle English dialects. The space continuum in *LAEME* is incomplete, but the time continuum is inextricably entwined with that of *LALME*. In order to allow empirically responsible comparison with the later period, and a conspectus of diachronic and diatopic variation for Middle English as a whole, we have produced text dictionaries as the equivalent of the *LALME* LPs, and a series of item and feature maps, which to some extent parallel the item maps and dot maps in *LALME*. The *LAEME* text dictionaries and maps are derived from a corpus of lexico-grammatically tagged texts,[22] which form the core of the *LAEME* database. The corpus methodology and outputs are described in part II. Other display types that have been made possible by our revised methodology are discussed and illustrated in part III.

1.5.3 Anchor texts and the fit-technique

In constructing *LAEME* we faced a further problem that is closely linked to the general problem of the gapped time/space continuum. For *LALME*, the numbers of surviving local documentary texts in English from 1350 onwards, though not completely even in their distribution, provided a reasonable matrix of localised anchors (*LALME* 1, §2.3.2). As indicated in §1.2 above, in the early Middle English period the vernacular was not used for the kinds of official document that provide explicit local origins.[23] The only official documents appearing in English in the late 12th and the 13th centuries are copies from pre-Conquest originals found in cartularies and monastic registers. Only very few of these were linguistically updated

---

[22] For explication of the term 'lexico-grammatically tagged' see chapter 4.

[23] One famous exception is the 1258 Proclamation of Henry III. This was apparently originally drawn up in French and then translated into English and copied for distribution to every county in England. Only one of the single sheet English copies survives, that sent to Oxford (Oxford, City Archives, Town Hall, St Aldates, H 29). The official enrolment is London, PRO C 66, Patent Rolls, 43 Henry III, m. 15.40. The French original of the same date is enrolled on the previous roll. As mentioned in §1.2 above, documents in English did continue to be created for a time after 1066, but these were heavily dependent on pre-Conquest models.

by the copying scribes. Usually, presumably at least partly for reasons of 'authenticity', the language of these copies is preserved by the monastic registrars as unmodernised Old English (this is particularly common in the case of boundary clauses — Kitson 1995: 48–49). In spite of the date of copying, these texts therefore do not belong to the Middle English corpus. In some cases, the later copyists had only a very shaky grasp of Old English, so that their attempts accurately to transcribe their exemplars produce garbled versions that are equally useless as samples of early Middle English (Laing 1991: 38–40). When these are discounted, there remains only a very partial matrix of genuinely original or updated documentary texts (Laing 2000a: 104–106).

This documentary matrix is so exiguous that alone it would be virtually useless as an aid for fitting unlocalised material. Although we were aware from the outset of the almost total lack of original documentary material pre-1350, we had hoped that more of the Old English texts copied during our period would turn out to be of use. The recognition that almost all were not usable provided a further setback to our expectation of using *LALME* methodology to produce something like a parallel early Middle English atlas. Indications of local origins for early Middle English texts are not, to be sure, confined to documents. Literary texts may also be associated on non-linguistic grounds with particular places. There is however a hierarchy in such associative clues. Extra-linguistic local associations in documentary texts are for the most part more likely to be reliable than those for literary manuscripts. Local records or legal instruments were mostly drawn up by local scribes, so can usually be trusted to attest forms of language of their stated place of origin or of somewhere nearby. The few exceptions are likely to be recognised because the body of local documents for any place would normally constitute a tradition of scribal practice against which non-local deviation is obvious. Non-linguistic associations in literary manuscripts represent a much broader spectrum of localising evidence.

The single most convincing ascription of local origins in a literary manuscript in our corpus is Dan Michel's colophon in London, British Library, Arundel 57 containing *Ayenbite of Inwyt* (see above §1.3). This constitutes local association at least as convincing as any evidence for a documentary source, and accordingly the *Ayenbite* functions as a firm anchor. But other indications to be found in literary manucripts are variously less convincing.

The so called Worcester Tremulous Scribe was interested in Old English texts, and he provided Latin (and in three manuscripts also Middle English) glosses in at least twenty manuscripts known to have been in Worcester during the Middle Ages (Franzen 1991: 29–83). His work has been ascribed to the first half of the 13th century (Ker 1937) thus falling squarely in our period. The fact that he was working on manuscripts owned by Worcester Priory at the time need not entail that he was brought up and trained in Worcester; but lacking evidence to the contrary it is a good working assumption. Apart from his major work of glossing, he also wrote longer texts in English. These are preserved in Worcester Cathedral, Dean and Chapter Library F. 174: a copy of Ælfric's *Grammar and Glossary*, which is at least partly 'translated' from Old English to Middle English, and the so-called Worcester fragments which are probably copied from a late 12th-century exemplar (Moffat 1987). We have not included his Middle English glosses in the *LAEME* corpus, but the Tremulous Scribe's extended English work provides a literary anchor text for Worcester, especially as, allowing for date, its dialect appears consonant with Worcester material from the later period.

Further down the hierarchy lies a manuscript like Oxford, Bodleian Library, Digby 86. This is a trilingual manuscript (Latin, French and English) written in the last quarter of the 13th century (Tschann and Parkes 1996: xxxvii–xxxviii) by one main scribe with a second scribe contributing only a very small amount of French text. The manuscript contains obits (one written by the main scribe himself) of people associated with places in SW Worcs and N Gloucs (Laing 2000b: 524 and references there cited).

Less probative still, are manuscript associations provided by evidence for ownership, such as library pressmarks or *ex libris* inscriptions. Medieval library catalogues may or may not be able to establish the ownership of a particular manuscript at the time when it was actually produced. Later ownership does not guarantee that the scribe or scribes who contributed to its contents were local to its present home. Even contemporary ownership need not imply that the scribes themselves had not come from elsewhere. A cautionary example is provided by the British Library, Cotton Cleopatra MS of *Ancrene Riwle*, for which the earliest evidence of ownership is an inscription of ca 1300 indicating that it belonged to Canonsleigh Abbey in Devon (Dobson 1972: xxv). But we believe from linguistic evidence that of the three Middle English scribes contributing to it, two came from North Worcestershire or North Herefordshire, and one from West Norfolk (cf. Laing 2000a: 114).

Indications of provenance provided by manuscript contents, scripts or illuminations must be similarly treated with caution. However, the language of these texts may be taken, at least provisionally, to represent the language of the place of origin of the manuscript, subject to confirmation by other local material.

We can therefore cautiously supplement our documentary matrix of anchors with literary anchors. Even so, the business of fitting text languages that have no extra-linguistic indications of provenance is still compromised by the unevenness and overall thinness of the early Middle English anchor matrix. In this situation we are fortunate to have the much more densely attested dialect continuum for later Middle English displayed in *LALME*. A dozen or so of our texts were already localised in *LALME*. More detailed analysis has led to modifications of some of these placings. In two cases, where the contributions of more than one scribe were amalgamated, these have been separated. In general, however, the localisations of these early texts hold good, whether as anchor texts or fitted texts. Where necessary, then, the *LALME* continuum can be used as scaffolding for the putative early Middle English continuum. In areas where there is little or no early Middle English attestation, we can interpolate unlocalised text languages of the earlier period into the *LALME* continuum in order to establish plausible localisations. But it must be understood that 'fitting' for LAEME is not anywhere near such a robust concept as it is for *LALME*. The *LALME* configuration has been used to help with some fittings, but in the sparsest areas any localisation is bound to be very approximate indeed and will always be subject to subsequent revision if more data or information becomes available. For much of LAEME, the display of linguistic data in map form at all is a convenient but highly generalised abstraction.

The apparently exact placings attempted for LAEME are a function of perceived patterning in relation to other texts of similar language within a kind of abstract linguistic space that also takes into account the time axis. They are also driven by the necessity, for mapping purposes, of putting in *some* specific place text languages that appear to be homogeneous and local. Where text languages exist in larger numbers, and the configuration is denser (e.g. in the SW Midlands) the concept

of linguistic space becomes even more important. In the early Middle English period, religious houses and a number of early-established town schools (Orme 1973: 295–325) would have provided opportunities for learning the art of writing and copying. For the SW Midlands, surviving early Middle English texts, in somewhat differing forms of language, outnumber the most likely places of origin of written local dialect systems. The complex of texts that include those in London, Lambeth Palace Library 487, London, British Library, Cotton Nero A xiv, London, British Library, Cotton Caligula A ix, part I (Laȝamon A), and part II (*Owl and the Nightingale*) are all very similar to each other, and also to the language of the Worcester Tremulous Hand and other material with Worcester associations. It is possible that varying Worcester language is what this complex may represent. But it would be very difficult to display the material cartographically all on one spot; so for the purposes of mapping, texts have sometimes been spread out, according to the usual criteria for fitting, across areas in which there were few or no contemporary centres of teaching and learning.

1.5.4 Corpus versus questionnaire

The advantage of a questionnaire is that the same list of items is used for every informant. This enables the dialectologist to address the key desiderata of dialectology: description and comparison. The questionnaire's two main weaknesses are that (a) it requires considerable prior knowledge of the language (or else a great deal of trial and error) before an effective questionnaire can be finalised; and (b) because its contents are by definition limited by the investigator's selection of items, its function as a linguistic net only ever results in a limited catch (Williamson 1992/3, p. 139).

There is a further disadvantage, which is that questionnaire analysis has to be done by hand. This makes its use particularly problematic for the English of our period. As part of the point of the survey of the earlier period was comparison with the later material, we would want most of the *LALME* items to be included also on the *LAEME* questionnaire. For very good methodological reasons (*chaque mot a son histoire*) the *LALME* questionnaire had not included many items eliciting phonological categories. For those that were there, it was operationally difficult for the analyst to collect consistently for them. For items such as A/O (reflexes of OE *ā)* or WH- (reflexes of OE *hw-*) one had to record for that item the segments not just from the relevant words that featured already on the questionnaire as items in their own right (for which one also had to record the full spelling in the relevant place), but one also had to note them for all the words not on the questionnaire already. For early Middle English, while fully acknowledging lexical specificity, we wanted to include many more phonological items that we thought would be of particular interest at the period, such as the reflexes of OE *y̆*. It would have been impractical to include as separate items all the possible lexical sources for this category. We were also interested in similarly 'catch-all' orthographical items such as the distributions of the letters thorn and edh, whether they were used initially, medially or finally.

During the early Middle English period, the language is in a process of rapid change from a state with still a considerable degree of inflection to one in which endings are being levelled. Prepositions are increasingly taking over the functions of some verbal prefixes and of inflexions. It is in principle likely that such changes were happening at different times in different places. We could not be content simply to

record the spellings of the stems of nouns and adjectives: there was a case for recording not only the singular and plural endings of each as categories in their own right, but also whether they had subject and object function in the sentence. The pronouns still show complex inflectional systems and even the definite article may still be marked for case number and gender. The relationships between linguistic form and function in these instances are potentially of dialectal significance. It will be obvious by now how impractically large and complicated the early Middle English questionnaire had become:

> The operational difficulty in noting each form and function of a word separately is compounded if one also tries to record separately its inflectional ending for comparison with others used in the same function. ... If *all* this potentially significant information is to be noted, a single word or morpheme in the text may have to be recorded in a number of different places in the questionnaire.
> It was apparent early in the investigation that the method of 'analysis by hand' described above was totally impractical. It was time-consuming, over-complex and the inherent likelihood of error and inconsistency was unacceptable (Laing 1993: 126).

In these circumstances we decided to use a different analytical tool for the investigation of early Middle English linguistic variation.

1.5.5 Corpus-based dialectology

Instead of using a questionnaire, we elected to transcribe all the early Middle English texts (or extensive samples of very long texts) and to put them on disk in a form enabling electronic processing: [24]

> The advantage of this method is that *all* the linguistic data can be subjected to analysis without the investigator being committed to a pre-selected set of dialectal discriminants. The results of the analysis may then inform the selection of items for linguistic profiles and dialect mapping (Laing 1994:127).

The resulting corpus of lexico-grammatically tagged texts is a database that can be sorted and analysed (using software written for the purpose by Keith Williamson). From each tagged text is derived a text dictionary, which is the taxonomised inventory of each text language. Any pair or *n*-tuple of these may be compared electronically. Such a taxonomised inventory is what we call an assemblage, which is a proper subset of a given scribe's total usage. Thus the database, in association with the processing software, becomes itself the instrument of selection. Because of its open-endedness, it develops a heuristic function impossible from self-limiting procedures such as the questionnaire. Further, the corpus allows for extended use (see chapters 3 and 4): rather than being solely a dialectological resource, it also becomes a powerful research tool for studies in historical phonology, morphology, syntax, semantics and pragmatics.

---

[24] For detailed description and explication of the corpus and its methodology see part II.

1.5.6 Scribal practice

For late Middle English, McIntosh observed that many scribes were translators, and many also perpetrated *Mischsprachen*, but that literatim copying was comparatively rare. The situation seems to have been different in early Middle English, where there were far fewer translating scribes and proportionally more literatim copyists. As a result, we find more manuscripts containing linguistically composite texts. Considering the relative paucity of data at this period, we do not have the luxury of rejecting a large proportion of our source material as being in mixed language. It is therefore necessary to subject the texts to more detailed analysis so as to establish their stratigraphy. It has proved possible in a number of cases to isolate within a text in mixed language one or more linguistic layers which can be taken to represent genuine regional usages.

Benskin and Laing (1981) details for late Middle English the copying strategies of scribes, and the kinds of *Mischsprachen* and pseudo-*Mischsprachen* that result from phenomena such as relict usage and constrained selection. Relicts, by definition, are confined to the output of translating scribes. A relict is a form from the usage of the exemplar that is alien to the dialect of the scribe and which he has allowed to slip through his translating net. This may happen either by accident as the translating scribe is working in to his task, or perhaps sometimes deliberately because it helps the text retain an exotic flavour. Constrained selection invokes the concept of a scribe's passive repertoire of spellings. A translating scribe will normally translate all exemplar forms that are alien to his own usage — his active repertoire. If, however, a spelling not his own is familiar to him because it is to be found in neighbouring areas, it forms part of his passive repertoire, and he may well not feel the need to translate it. In this case his own usage is constrained by that of his exemplar, and his copied text, while not containing any forms alien to his area of origin, may show skewed frequencies for the variant spellings of particular items.

Following the terminology established in Benskin and Laing (1981), Laing (2004) describes the types of scribal complexity, both spatial and temporal, to be found in early Middle English texts, and the major analytical techniques for disentangling them. References to specific published studies on early Middle English texts undertaken for the creation of *LAEME* may be found under individual manuscript entries in the Index of Sources, while notes on particular textual interpretations are often to be found in the tagged texts themselves.[25]

This chapter has introduced and defined the concept 'early Middle English text'. Chapter 2 addresses the main theoretical issues concerning orthographical and phonological interpretation of the LAEME materials.

---

[25] For explanation of textual notes in the tagged text see chapter 3, §4.

References

Benskin, M. 1981. The Middle English dialect atlas. In: Benskin and Samuels, eds. 1981: xxvii–xli.

Benskin, M. 1991. The 'fit'-technique explained. In: Riddy 1991: 9–26.

Benskin, M. 1994. Descriptions of dialect and areal distributions. In: Laing and Williamson 1994:169–187.

Benskin, M. and Laing, M. 1981. Translations and *Mischsprachen* in Middle English manuscripts. In: Benskin and Samuels 1981: 55–106. Substantially reprinted in *LALME* 1, §3.

Benskin, M. and Samuels, M.L. eds. 1981. *So meny people longages and tonges: philological essays in Scots and mediaeval English presented to Angus McIntosh.* Edinburgh: The Editors.

Berndt, R. 1965. The linguistic situation in England from the Norman Conquest to the loss of Normandy (1066–1204), *Philologia Pragensia* 8.145–63.

Brooks, N.P. ed. 1982. *Latin and the Vernacular Languages of Early Medieval Britain* . Leicester University Press.

Clanchy, M.T. 1979. *From Memory to Written Record: England 1066–1307*. London: Edward Arnold.

Clark, C. 1992. The myth of the Anglo-Norman scribe. In: Rissanen *et al.* 1992: 117–29.

Davis, R.H.C. 1976. *The Normans and their Myth*. London: Thames & Hudson.

Dossena, M and Lass, R. eds. *Methods and data in English historical dialectology*. Bern: Peter Lang.

Fisiak, J. ed. 1995. *Medieval dialectology*. Berlin/New York: Mouton de Gruyter.

Fleischman, S. 2000. Methodologies and ideologies in historical linguistics: on working with older languages. In: Herring *et al.* 2000: 33–58.

Franzen, C. 1991. *The Tremulous Hand of Worcester: a study of Old English in the thirteenth century*. Oxford: Clarendon Press.

Gradon, P.O.E. 1979. *Ayenbite of Inwyt*, vol. 2, Introduction, Notes and Glossary. EETS OS 278. Oxford University Press.

Herring, S.C., van Reenen, P. and Schøsler, L., eds. 2000. *Textual parameters in older languages*. Amsterdam/Philadelphia: John Benjamins.

Hunt, Tony. 1991. *Teaching and Learning Latin in Thirteenth-Century England*, 3 vols. Cambridge: Brewer.

Ker, N.R. 1937. The date of the 'Tremulous' Worcester Hand. *Leeds Studies in English* 6.28–29.

Kitson, P. 1995. The nature of Old English dialect distributions, mainly as exhibited in charter boundaries. In: Fisiak 1995: 43-135.

Kytö, M., Rissanen, M. and Wright, S. eds. 1994. *Corpora across the centuries*. Amsterdam/Atlanta Ga: Rodopi.

Laing, M. ed. 1989. *Middle English Dialectology: essays on some principles and problems*. Aberdeen University Press.

Laing, M. 1994. The linguistic analysis of medieval vernacular texts: two projects at Edinburgh. In: Kytö *et al.* 1994: 121–141.

Laing, M. 2000a. *Never the twain shall meet*: early Middle English — the east-west divide. In: Taavitsainen et al. 2000: 97–124.

Laing, M. 2000b. The linguistic stratification of the Middle English texts in Oxford, Bodleian Library, MS Digby 86. *Neuphilolgische Mitteilungen* 101.523–569.

Laing, M. 2004. Multidimensionality: time and space and stratigraphy in historical dialectology. In: Dossena and Lass 2004: 49–96.

Laing, M. and Lass, R. 2006. Early Middle English dialectology: problems and prospects. In: A van Kemenade and B.L. Loos 2006: 417–51.

Laing, M. and Williamson, K. eds. 1994. *Speaking in our Tongues*. Cambridge: D.S. Brewer.

LALME = McIntosh, A. Samuels, M.L. and Benskin, M. eds. 1986.

Lass. R. 2000. Language periodization and the concept 'Middle'. In: Taavitsainen *et al.* 2000: 7–42.

McIntosh, A. 1956. The analysis of written Middle English. *Transactions of the Philosophical Society*, 26–55. Reprinted in: Laing 1989: 1–21.

McIntosh, A. 1973. Word geography in the lexicography of Middle Engish. *Annals of the New York Academy of Sciences* 211. 55–66. Reprinted in: Laing 1989: 86–97.

McIntosh, A. Samuels, M.L. and Benskin, M. eds. with the assistance of M. Laing and K. Williamson. 1986. *A Linguistic Atlas of Late Mediæval English*, 4 vols. Aberdeen University Press.

Malone, K. 1930. When did Middle English begin? In: J.T. Hatfield ed. *Curme Volume of Linguistic Studies*. (Philadelphia: Linguistic Society of America, 110–117.

Moffat, D. ed. 1987. *The soul's address to the body: the Worcester Fragments*. Lansing: Michigan State University Press.

Orme, N. 1973. *English Schools in the Middle Ages*. London: Methuen & Co.

Pelteret, D.A.E. 1990. *Catalogue of English post-conquest vernacular documents*. Woodbridge: The Boydell Press.

Prior, O.H. 1923. Remarques sur l'Anglo-Normand. *Romania* 49.161-85.

Richter, M. 1979. *Sprache und Gesellschaft im Mittelalter. Untersuchungen zur mündlichen Kommunikation in England von der Mitte des elften bis zum Beginn des vierzehnten Jahrhunderts*, Monographien zur Geschichte des Mittelalters 18. Stuttgart: Hiersemann.

Riddy, F. ed. 1991. *Regionalism in late medieval manuscripts and texts*. Cambridge: D.S. Brewer.

Rissanen, M., Ihalaien, O., Nevalainen, T., and Taavitsainen, I. eds. 1992. *History of Englishes: New Methods and Interpretations in Historical Linguistics*. Berlin and New York: Mouton de Gruyter.

Rothwell, W. 1968. The teaching of French in medieval England. *Modern Language Review* 63.37-46.

Rothwell, W. 1975-76. The rôle of French in thirteenth-century England. *Bulletin of the John Rylands Library* 58. 445-66.

Rothwell, W. 1978. A quelle époque a-t-on cessé de parler français en Angleterre? *In Mélanges de philologie romaine offerts á Charles Camproux*, 1075-89. Montpellier: CEO.

Rothwell, W. 1983. Language and government in medieval England. *Zeitschrift für französische Sprache und Literatur* 93.258-70.

Shelley, P. van Dyke. 1921. *English and French in England 1066–1100*. Philadelphia: University of Pennsylvania PhD thesis.

Short, I. 1980. On bilingualism in Anglo-Norman England. *Romance Philology* 33.467-79.

Short, I. 1992. Patrons and polyglots: French literature in twelfth-century England. *Anglo-Norman Studies* 14.229-49.

Short, I. 1996. *Tam Angli quam Franci*: self-definition in Anglo-Norman England. *Anglo-Norman Studies* 18.153-75.

Taavitsainen, I., Nevalainen, T, Pahta, P. and Rissanen, M. (eds.) 2000. *Placing Middle English in Context*. Berlin and New York: Mouton de Gruyter

van Kemenade, A. and Loos B.L. (eds.) 2006. *Handbook of the History of English*. Oxford: Blackwell.

Williamson, K. 1992/3. A computer-aided method for making a linguistic atlas of Older Scots. *Scottish Language* 11/12.138–157.

Wilson, R.M. 1943, English and French in England 1100–1300. *History* 23.37–60.

Woodbine, G.E. 1943. The language of English law. *Speculum* 18.395–436.