French lexis in the Auchinleck Manuscript: A digital-philological approach

Rory G. Critten[1], Cyrille Gay-Crosier[1,2] and Davide Picca[1,2]

[1]Department of English, University of Lausanne, Switzerland
[2]Department of Language and Information Sciences, University of Lausanne, Switzerland

Abstract

The Auchinleck Manuscript was produced in the 1330s in London and is best known to scholars of Middle English literature on account of the romances that it transmits. Several of these texts treat the establishment and defense of England and it has been argued that their interest in English history is matched by the language in which almost all of the manuscript's texts are written: English. This article reconsiders the Englishness of the Auchinleck Manuscript via a quantitative analysis of its lexis. We show that a comparatively large proportion of the Auchinleck lexicon has connections to French and that, of these words with French connections, many do not appear to have been much used in English writing before the 1300s. Our statistics are derived via program scripts that match Auchinleck lexicon items to headword entries in the *Middle English Dictionary* and collect data pertaining to word etymology and earliest dates of citation from those entries. Where previous studies have emphasized the porous boundaries between English and French in 14th-century English contexts, we posit that some poets might aim to make creative capital out of the deliberate juxtaposition of the languages. The argument is supported by a series of visualizations; interactive versions of these visualizations and the data on which they are based are archived at our project website (https://solliryc.github.io/AuchinleckDataViz/).

Correspondence: Rory G. Critten, University of Lausanne, Lausanne, Switzerland. E-mail: rory.critten@unil.ch

1 Introduction

The Auchinleck Manuscript was produced in the 1330s in London; in its current, damaged state, it contains forty-four texts. Scholars of Middle English literature have been primarily interested in the romances that the book transmits, several of which treat the establishment and defense of England by legendary English heroes such as Guy of Warwick, Beves of Hampton, Arthur, and Horn. This concern with the story of England has been linked to the language of the Auchinleck Manuscript's texts: whereas contemporaneous manuscripts containing poetry in English transmit that poetry alongside texts in French and Latin, almost everything in the Auchinleck book is in English. Theme, language choice, and audience are often said to be in alignment here, where the history of England is told in English for the English. In an influential monograph, Turville-Petre termed Auchinleck a 'handbook of the nation' (1996, pp. 108–41, at p. 112) and, as such, the manuscript has achieved an almost totemic status in accounts of early English literature.[1]

This article reconsiders the second of the two claims that are typically made for the Auchinleck Manuscript. While it is undeniable that many of the book's texts take England as their locus, what do we mean when we say that these works are written *in English*? The 14th century marks a period of rapid change in the history of the language. In particular, it is in these years that so many French words that are now thoroughly naturalized make their first appearance in written English.[2] Much of this in-coming lexis will already have been current in England because French was widely used as a second vernacular by the literate there throughout the later Middle Ages. When French words make their first appearances in English texts, their effect on readers and audiences does not bear straightforward comparison with the jarring that still accompanies the introduction into a Present Day English sentence of phrases such as *joie de vivre*, *raison d'être*, or *bœuf bourguignon* (so Rothwell, 1979, 1980–81, 1991; Trotter, 1996, 2003b, 2010).

And yet the 14th century also sees clearer distinctions being drawn between English and French in the course of the prolonged period of conflict between England and France that we now call the Hundred Years War (1337–1453). These hostilities could express themselves in a desire for conflation as well as differentiation. The war proper began, arguably, with Edward III's formal claim to the French throne in 1340; if he has had his way, France would have become a part of England, not an inalienably distinct adversary lying across the sea.[3] But the conflict also sees a process whereby the populations and lands of England and France begin to separate out from each other. The status of French in this geopolitical context is especially labile. In the later Middle Ages, French is a language that divides as well as unites speakers inhabiting the territories on either side of the English Channel.[4]

Our primary aim in this article is to nuance accounts of the Auchinleck Manuscript's nationalist commitments via a quantitative analysis of the debts to French that the book's lexis owes. In so proceeding, we take advantage of two trail-blazing digital humanities projects produced by medievalists: the online facsimile edition of the Auchinleck Manuscript by Burnley and Wiggins published in 2003 (https://auchinleck.nls.uk), and the *Middle English Dictionary* (MED), whose compilation began in 1925 and which in 2018 was provided with the new online interface that has allowed us to gather the statistics that we present below (https://quod.lib.umich.edu/m/middle-english-dictionary/). These statistics use data from the MED to explore how many of the Auchinleck Manuscript's words have French connections and the extent to which these words are likely to have been naturalized by the time of the book's inscription: we are especially keen to reconsider how the Auchinleck Manuscript's French lexis might have been perceived by the book's first users. Our results are given in extracted form. All of the datasets and visualizations that we reference can be consulted at the project website accompanying this article (https://solliryc.github.io/AuchinleckDataViz/).[5]

We start by describing the programming scripts that we wrote in order to match the words used in the Auchinleck Manuscript to the MED headword entries from which we extracted our data.

2 Matching Auchinleck Manuscript words to MED headword entries

Working with the XML files of Burnley and Wiggins's website that are archived at the Oxford Text Archive at the Bodleian Library in Oxford (https://ota.bodleian.ox.ac.uk), our first programing script retrieved a list of 17,042 different words used in the Auchinleck Manuscript's texts and the frequency of those words' occurrence.[6] Our second script then put these words through the MED's search engine individually and recorded the results returned. For all the searches performed, we chose the MED's default search method, 'Headword (with alternate spellings)', and the default sorting method, 'by relevance'. When the search engine returned more than one result, the first result was selected as the MED match. This initial search allowed us to match 10,384 of the 17,042 different words in the Auchinleck Manuscript to a MED headword entry.

The new MED online search interface considers all the attested forms of words in its database rather than just its headwords; the dictionary now also allows for some differences between inputted words and its recorded forms that reflect the considerable diversity of Middle English orthography (see https://quod.lib.umich.edu/m/middle-english-dictionary/about). Nevertheless, the MED relevance search did not catch

all of our items. In order to optimize the number of results that could be obtained automatically, we wrote a third script that applied a series of alterations to the unmatched Auchinleck words and put them through the MED search engine again. The changes made were designed to lemmatize the unmatched items (i.e. bring them closer into line with forms that the MED would recognize).

For example, the genitive- and plural-forming suffixes attached to nouns were removed and finite verb endings were removed and replaced with the infinitive-forming suffix -*en*. These modifications yielded a further 3,587 automatic matches between Auchinleck lexicon items and MED headword entries.[7] A fourth script excluded from our search the proper nouns listed in the Auchinleck Manuscript lexicon.[8] This operation allowed us to remove 1,672 items from the list of Auchinleck words to be matched to the MED; we also excluded a further thirty-six words identified by the Auchinleck website editors as miscopyings by the addition of rounded brackets.

These exclusions left us with a total of 15,334 words requiring matches for which 13,971 matches had been found automatically; the remaining 1,363 Auchinleck lexicon words were treated manually. At this stage, we also removed unmatched words that were (1) compounds or contractions not recorded in the MED; (2) proper nouns not already excluded; or (3) words identified as citations from Latin or French by Burnley and Wiggins, who italicize such items (e.g. the Latin lines in Auchinleck's *Dauid þe King*). Finally, of the 17,042 Auchinleck lexicon words, we had matched 14,848 words to a MED entry and excluded 2,194 words

from our search. The outcomes of this matching process are summarized in Table 1.

The final results of the matching process are posted at our project website as Document 3.

3 Accuracy of the matches and limitations of MED data

Both manual completion and checking of the results were conducted via the search engine on Burnley and Wiggins's website, which shows where individual words are used in the Auchinleck Manuscript's texts (https://auchinleck.nls.uk/search). Manual checking of randomly generated lists giving 100 matched and excluded words from our final results revealed that our matches are not perfect. Our rate of clearly incorrect matches is around 20%; a sample check of 100 matched words showing eighteen patent errors is uploaded as Document 4 at our website. While this result is far from ideal, it is not disastrous for our project for the reasons that we outline below.

Often the bad results concern homonyms, that is, words that sound and are spelled the same but which have different meanings and, potentially, different etymologies. For example, our sample check of 100 matched words alerts us to the automatic mismatching of the word *biker*, which occurs ten times in the Auchinleck Manuscript, not to MED s.v. *biker* n. 2, 'a martial encounter', as per its usage throughout the book, but to MED s.v. *biker* n. 1, 'ornate covered bowl'. Other examples of mismatched homonyms in the sample include the words *godes*, used many times in the Auchinleck Manuscript with the meaning 'God's' or 'gods' (MED s.v. *God* n. 1) but matched by the relevance search with MED *gode* n., 'a pointed rod for driving oxen etc.'; *wif*, used many times in the Auchinleck Manuscript with the meaning 'a human biological female' (MED s.v. *wif*, n. 2) but matched by the relevance search to MED *wif*, n. 1, '? a bundle of harvested plants'; and *irad*, which occurs twice in the Auchinleck Manuscript, once as a part of the verb *ireden*, to advise, and once as part of the verb *reden*, to read, but which is matched by the relevance search with the noun *irad*, 'a condition, stipulation' (see MED s.vv. *ireden* v.; *reden* v. 1; *irad*, n.).

Table 1. Matches between Auchinleck lexicon words and MED headwords

Total number of different Auchinleck words	17,042
Automatic exclusions	1,708
Proper nouns	*1,672*
Other	*36*
Auchinleck words to be matched to MED entries	15,334
Automatic matches	13,971
Simple	*10,384*
Modified	*3,587*
Manually sorted words	1,363
Manual matches	*877*
Manual exclusions	*486*
Total number of matched Auchinleck words	14,848

These returns highlight a problem with the MED's relevance search, which allows obscure lexis to trump more common words. The relevance search is one of a series of features introduced as part of the most recent update of the MED's online interface (described in Schaffner, 2019). One proposition for further development of this tool would be to give greater weight in the calculation of results to the number of quotations that the MED has for a given item. This is not an infallible sign of a word's frequency in Middle English texts but it would help avoid many unlikely matches. For example, *wif* meaning '? a bundle of harvested plants' has just one quotation in the MED, whereas *wif* meaning 'a human biological female' has 173.

Some confusion over homonyms remains inevitable in our case because the statistical model that we have adopted does not allow for words like *irad* that might legitimately be linked to more than one MED headword. For our immediate purposes, it is important to recognize that, of the eighteen straightforward mismatches in the checked sample, only one affects a word with French connections: the word *geste* is used once in the Auchinleck Manuscript to refer to a poem or song about heroic deeds but is matched in the relevance search with MED *geste* n. 2, 'a race, family, kindred'. Errors do occur in the matching of French words, but they are generally less significant. For example, in the checked sample, the word *erbere*, 'arbour', is mismatched to MED *herbe*, n., 'any non-woody plant', instead of MED *herber* n. 1, but both words share a connection to the same French word, *erbe*.

Mismatches like that affecting *erbere* are less crucial for the etymological statistics that we present below than for the data pertaining to the first dates of citation. Etymology can usually be perceived in a word's form regardless of its semantic and syntactic functions but these functions are more crucial when attempts are made to determine degrees of naturalization. It is a general rule in comparative and historical linguistics that words borrowed from other languages are most often introduced as nouns and only later adapted to serve in other grammatical roles, for example (so Tadmor *et al.*, 2010, p. 231). Nevertheless, in what follows, we let incorrect matches stand and make manual alterations only in the case of two very common words to which the MED's relevance search

mistakenly assigns French connections: the indefinite article *a*, linked by the relevance search to MED *a* prep. 2 instead of MED *a* indef. art., and the pronoun *sche*, linked to MED *se* n. 2 instead of MED *she* pron.

The matches that we have collected represent a classification of the Auchinleck Manuscript's lexis according to the current parameters of the MED's relevance search: this is what currently happens when the MED 'reads' a Middle English text. Like the matches we have made, the data that the matches allow us to retrieve from the dictionary should be treated with caution. No dictionary's corpus is ever complete: below we present evidence allowing for the antedating of several MED headwords. The long period over which the MED has been compiled also leads to complications because advances in lexicography that have taken place over the last century are unevenly reflected in the dictionary's entries. The MED's editors are upfront about this limitation (see Lewis, 2007).

We are fortunate, however, in that one of the MED's flaws suits our purposes. The MED tends to draw its quotations from literary, monolingual texts and to ignore the frequent appearance of Middle English in nonliterary and multilingual environments. In this respect, it bears the mark of its earliest debts to wordlists designed to help readers of Chaucer (described in Blake, 2002). The disproportionate representation of monolingual, literary writing in the MED has attracted criticism from those interested in reconstructing England's multilingual past (e.g. Rothwell, 2000, 2002). It is a boon for us, however, because we are attempting to reconstruct the expectations of medieval readers encountering precisely this sort of text: the monolingual Middle English literary work.

4 Etymological connections of Auchinleck Manuscript words

Once our matches between the Auchinleck Manuscript lexicon and the MED had been established, we wrote programming scripts that scraped (i.e. collected) the etymological data recorded at the MED and that assigned this information to the matched lexicon words. Table 2 shows the

Table 2. Etymology of different Auchinleck manuscript words

Etymological connection to	Number of words	Percentage of matched lexicon
French	4,051	27
Latin	1,738	12
English	7,856	53
Scandinavian	1,243	8
Other	928	6
None given	2,856	19
Total connections	15,816	106
Total matched words	14,848	100

Table 3. Etymology of MED headwords

Etymological connection to	Number of words	Percentage of matched lexicon
French	13,386	25
Latin	11,594	21
English	11,585	21
Scandinavian	2,023	4
Other	3,290	6
None given	22,784	42
Total connections	41, 878	77
Total MED headwords	54,507	100

Table 4. Etymology of all Auchinleck manuscript words used, including repetitions

Etymological connection to	Number of occurrences	Percentage of total words occurring
French	37,362	11
Latin	28,434	9
English	233,903	70
Scandinavian	20,196	6
Other	20,103	6
None given	50,834	15
Total connections	339,998	102
Total matched occurrences	333,847	100

etymologies that we have found for the 14,848 different Auchinleck lexicon items matched to a MED headword entry. Tables 3 and 4 nuance this picture. Table 3 provides parallel statistics for all of the headwords in the MED and Table 4 considers all the words used in the Auchinleck Manuscript's texts, including repetitions of individual words. The data in Tables 2 and 4 can be called up in Chart 1 on our project website.

Notes on the tables:

- Etymological connections
  - 'Etymological connection to French' gives words listed in the MED as Anglo-French, Anglo-Norman, Continental French, French, Middle French, Modern French, Northern French, Old French, Old Northern French, Old Provencal, and Provencal[9];
  - 'Etymological connection to Latin' gives words listed in the MED as Anglo-Latin, Classical Latin, Latin, Latinate, Late Latin, Medieval Latin, Vulgar Latin, Neo-Latin, and pseudo-Latin;
  - 'Etymological connection to English' gives words listed in the MED as Anglian, East Midland, Kentish, Late Old English, Mercian, Middle English, Midland, North Midland, Northern, Northeast Midland, Northwest Midland, Northumbrian, Old English, Old Kentish, Southern, South Midland, Southeast Midland, Southeastern, Southwest Midland, Southwestern, Western, West Midland, and West Saxon;
  - 'Etymological connection to Scandinavian' gives words listed in the MED as Danish, Icelandic, Norwegian, Old Danish, Old Icelandic, Old Norse, Old Norwegian, Old Swedish, Scandinavian, and Swedish;
  - 'Etymological connection to Other' gives words provided with an etymology in the MED that is not otherwise covered in our statistics;
  - 'None given' gives words for which no etymology is provided in the MED. A high proportion of MED words are not given an etymology (42%). We assume that this reflects a hangover from the MED's presentation in print, where etymologies are attached to headwords but not made explicit under related items which, in the MED online, have been given their own webpages.
- Totals
  - In Tables 2 and 4, the numbers of connections listed exceed the number of words analyzed. This is because the MED sometimes lists two or more etymological connections for a given word. Where this happens, we count the word twice. For example, MED *gentil*, adj. is listed in

the MED as having both Old French and Latin connections. We count it once under 'French' and once under 'Latin'. This strikes us as the fairest way of dealing with the MED's inconsistent treatment of words having multiple derivations. The latest edition of the *Oxford English Dictionary* (OED; https://www.oed.com) offers a more fine-grained approach to this problem (see further Durkin, 2002).

Tables 2 and 3 show that the proportion of different words with French connections in the Auchinleck Manuscript (27%) is similar to that in the MED (25%). This is remarkable given that the MED aims to represent Middle English for the whole period c.1100–c.1500 and that the half century after the production of the Auchinleck Manuscript (1350–1400) has traditionally been identified as the busiest period for imports from French into English (e.g. Jesperson, 1905; Baugh, 1935; Mossé, 1943). The bulk of the Latin words in Middle English enters the language later in the period (see Durkin, 2014, p. 258) and the discrepancy between the percentages of words having Latin connections in the Auchinleck Manuscript lexicon (12%) and the MED (21%) is much clearer.

If the lexicon of the manuscript's texts might be especially rich in French words, Table 4 shows that, of all the words occurring in the manuscript, including repetitions of individual words, words with English connections make up the largest portion (70%); words with French connections occur less frequently (11%). This suggests that, while the usage of French lexis might differ from text to text, as we outline below, French words were less likely to constitute part of the core vocabulary of Auchinleck's poems. Where French words are deployed in the manuscript's texts, they might thus have retained an exotic charge. Burnley (1992, p. 432) reaches this conclusion when faced with a similar set of statistics.

Even with regard to the figures given in Table 4, however, the French aspect of the manuscript's lexis appears unusually dominant. An earlier study by Dekeyser (1986) counting the frequency of French words as a portion of all words used in a sample of five early Middle English texts finds that 5.6% of all words used have a romance connection (i.e. they might be linked to French or Latin; our 'French' and 'Latin' categories are combined in this analysis), whereas 94.4% of the vocabulary is native English. The proportion of romance words observed by Dekeyser rises to 12.5% when he considers late Middle English texts but the texts in the sample that show the most frequent use of romance lexis postdate the Auchinleck Manuscript by at least fifty years. Dekeyser's results for early Middle English are most likely skewed owing to his avoidance of samples, for example, from the Auchinleck Manuscript. Nevertheless, his statistics provide valuable information regarding the linguistic background against which the Auchinleck poets were writing.

5 Perceiving French lexis in the 1330s

The picture that is emerging is of a lexicon whose French connections might be considered unusually strong in the 1330s. Our statistics help to corroborate comments to this effect that the editors of individual Auchinleck texts have made in passing (e.g. Smithers, 1952–57, II. p. 56–57; Macrae-Gibson, 1973–79, II. p. 62; Fellows, 2008, p. 82). Numerous factors might be adduced to explain the preponderance of words with French connections in the Auchinleck book, including the observations that many of the manuscript's texts are translations from French and that the courtly and chivalric themes that they often treat are especially likely to require French terms. What interests us specifically is how the Auchinleck lexicon's debts to French might have been perceived by the book's users in the 1330s. Was the manuscript's French-infused English heard as English with a French accent, or French with an English accent, for example? And what literary effects might French lexis conjure among an early 14th-century audience?

Our attempts to answer these questions are frustrated by the paucity of data currently available regarding the first readership of the Auchinleck book. The dialect of its main scribe as well as the ambition manifested in its manufacture allow for its production to be located in London, and whoever owned the book must have been wealthy: even in its now damaged state, Auchinleck is an imposing codex, having 331 folios at 250 × 190 mm; most of its texts were

once introduced by expensive illuminated miniatures (many of these have since been removed from the book).[10] Scholars have argued that Auchinleck was bought by readers belonging either to the mercantile elite or the nobility in London, but even such general claims have been disputed.[11] It seems fair to presume, however, that whoever did own Auchinleck in the 1330s will have had a working knowledge of French. Accounts of book ownership in the 14th century demonstrate that those who bought books often preferred French material (see e.g. Boffey and Edwards, 2008); recent attention to the currency of French in late medieval England in particular by Ingham (2012, 2014) has post-dated the decline of the language there to c.1350; and medieval London has been found to be a location especially propitious for French–English language contact (see e.g. Rothwell, 1983; Burnley, 2003; Butterfield, 2009, pp. 203–9).

Auchinleck's readership was thus at least theoretically well placed to spot moments where the vocabularies of French and English overlapped. This argument runs counter to the claims voiced in the well-known prologs to Auchinleck's romances of *King Richard* and *Of Arthour and of Merlin*: there we read that knowledge of French is on the wane. We propose a reconsideration of these famous passages in the light of our findings below. At this stage in our argument, it is worth briefly noting the evidence preserved in a range of slightly earlier and contemporaneous texts that suggests the cultivation of a new philological sensitivity. Where in its life of saint Kenelm the *Early South English Legendary* (1275–85) describes an Old English message sent from heaven to the pope at Rome as 'puyr on Englisch i-write' (l. 259: written purely in English), for example, the text's phrasing indicates the existence of another type of English that is mixed, perhaps the language in which the *Early South English Legendary* is itself composed (cited from Horstmann, 1887, pp. 345–55). And in the early 14th century, writers such as Robert Mannyng and Richard Rolle likewise suggest an awareness of the openness of English to external influences where they refer to a type of the language that might be 'straunge', a word whose possible meanings include 'foreign' (see MED s.v. *straunge*).[12]

The metalinguistic commentary inhering in each of these passages rewards more detailed reading than we can offer here. For our current purposes, it must suffice to note their implication that some English authors and their readers will have been aware of the effects of language contact on English long before the debates belonging to the so-called Inkhorn controversy of the 16th century, which provide us with some of the first secure evidence of perceived conceptual links between language usage and nationality (see further Barber, 1997, pp. 56–62). Contact between French and English may have been a particular topic of interest to the medieval English: Mannyng disserts at some length on the difficulty of rendering French verse forms in English, for example.

A corollary of these observations is that some Middle English writers might deliberately mobilize the developing lexicon of English in order to achieve a range of artistic effects. Here we follow arguments developed by researchers in Middle English literary studies via the close reading of particular texts (e.g. Lerer, 2008; Machan, 2016; Critten 2019); ground-breaking analysis of this sort with direct reference to the Auchinleck Manuscript has been conducted by Summerfield (2013), whose work signals a renewed interest in the multilingual contexts of the book.[13]

The data that we have assembled allow us to draw a series of wider-angle perspectives on the possible valences of the French lexis deployed in the Auchinleck Manuscript. The first is contrastive.

5.1 French lexis by manuscript text

Looking from text to text in the Auchinleck Manuscript, it becomes clear that some of the book's poems have more French lexis than others. Chart 1 on our project website can be manipulated to show the etymological debts of individual Auchinleck texts. At the bottom end of the scale, 7% of the matched words in *The Sayings of St Bernard* can be identified as having a French connection; the percentage rises slightly in the case of *The Thrush and the Nightingale* to 8%. At the other extreme, 20% of matched words in *Of Arthour and of Merlin* and in the surviving portions of *King Richard* and *Kyng Alisaunder* are identified as having French connections. This discrepancy is not surprising given that the different texts compiled in the Auchinleck Manuscript are now thought to have been composed at different times and locations prior to their inscription in the manuscript in London in the 1330s; that five or six different scribes participated in copying the book further enhances the likelihood of variation

across the linguistic profiles of the book's different texts.[14]

An early 14th-century reader who consulted several Auchinleck texts in sequence might perceive their relative novelty in their varying densities of French lexis. *The Sayings of St Bernard* and *The Thrush and the Nightingale* significantly predate the Auchinleck Manuscript and enjoyed circulation outside London: both these texts are also compiled in Oxford, Bodleian Library MS Digby 86, a manuscript thought to have been inscribed c. 1272–82 in Worcester (see further Fein, 2014). In contrast, newer productions compiled in the manuscript such as *Of Arthour and of Merlin*, *Kyng Alisaunder*, and *King Richard* manifest a willingness to advertise their connections to French-language culture via the inclusion of French tags in rhyme position, e.g. *saun fable*, *saun faille*, *par amure*, *tut entoure*.[15] Perhaps too works bearing greater debts to French were felt to advertise a more cosmopolitan outlook: this argument was developed some time ago by Clark (1966) in her comparative study of the lexicons of the Katherine Group hagiographies and the *Ancrene Wisse* (both c. 1225). These texts share an audience but differ in the proportions of French lexis that they deploy.

5.2 The novelty of the French lexis

The next perspective that our matching of the Auchinleck lexicon to MED headwords allows us to draw is chronological. We have executed programming scripts that scraped the MED's first dates of citation and reassigned this information to our matched words. As a result, we can now offer an account of how well established the French lexis used in the Auchinleck Manuscript is in Middle English writing in the 1330s, according to the MED.

The matter of providing first dates of citation for English words is notoriously thorny. The MED deals with the issue via their 'double-dating' system, which gives dates for the first recorded appearances of words in manuscript and, if the text cited is thought to be younger by 25 years or more than the manuscript that transmits it, its presumed composition date (as described in Lewis, 2007, p. 44). Since we are interested in tracing the earliest appearances in written English of the French lexis used in the Auchinleck texts, the dates of presumed composition are of interest to us, but we cannot rely on them. There is nothing

to say that a word occurring in a later manuscript copy of a text also occurred in an earlier redaction. The scope for change between versions of a text copied at different times seems particularly great in the case of loanwords. It might be anticipated that texts collected more French vocabulary as they passed through time in parallel with the expansion of the English lexicon. But even this general rule is not secure. In her edition of *The King of Tars*, Perryman (1980, p. 26) notes that the earlier version of this poem contained in the Auchinleck Manuscript has a higher proportion of French lexis than a later copy of the same text, which seems to have re-anglicized its vocabulary.

For these reasons, we record separate sets of statistics for the earliest dates of citation of words in Middle English by manuscript and by date of presumed composition. The manuscript dates have the benefit of being relatively secure; the dates of text composition give a sense of how much earlier the French lexis under discussion might already have been circulating in written English. We attempt to make full use of the dating information provided by the MED by establishing the ranges attributed to a manuscript's or a text's production and taking the mid-point of those ranges as our reference figure. This allows us to distinguish between the MED's practice of dating e.g. to 'c1300' (25 years either side of 1300) and 'a1300' (up to 25 years before 1300) (described in Lewis, 2007, p. 44).

Figures 1 and 2 below show the MED's earliest dates of manuscript citation for Auchinleck words having French connections. Figure 1 weights the appearance of all different words in the Auchinleck lexicon equally; Fig. 2 takes into account how often Auchinleck lexicon items are used. Chart 2 on our website can be used to call up these and other figures using the MED's earliest citation dates.

Figure 1 suggests that much of the Auchinleck lexis having French connections was not commonly used in English texts before the 1300s. Some of these words are attested for the first time in citations from Auchinleck texts, e.g. in MED entries s.vv. *accordement*, *affliccioun*, *afrounten*, *allegeaunce* n. 2, and *argument* (lists of all the words represented in the chart data can be downloaded by clicking on the charts at our project website). Figure 2 shows that French lexis newly current in English writing was used as frequently as more familiar words with French connections; the dip in the
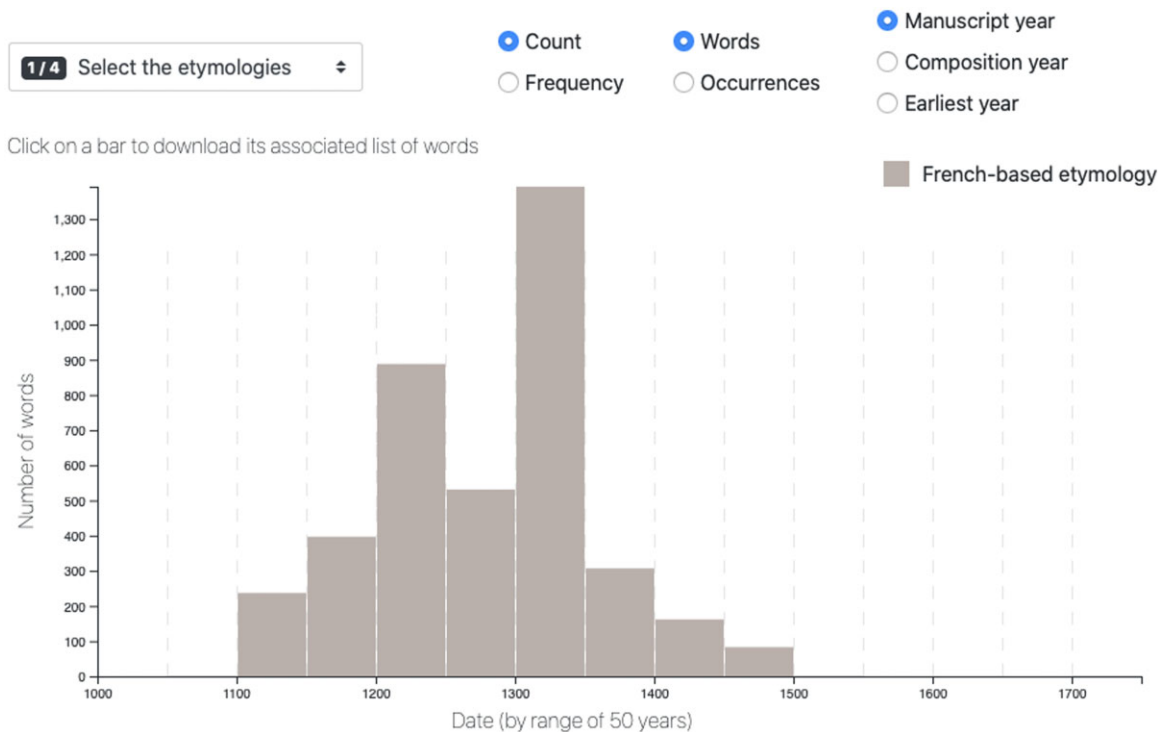
Fig. 1 Different Auchinleck Manuscript Words with French Connections by Date of First Manuscript Citation.

usage of words with French connections first cited in the period 1250–1300 might suggest a stylistic preference for newer French imports. This observation is of interest because only Chaucer among the Middle English poets is usually credited with having a sense of which words might be in or out of fashion.[16] Here, our statistics afford another perspective on the stylistic sensitivities of the Auchinleck Manuscript's poets, which have traditionally been downplayed.[17]

Finally, it is interesting to note that when the earliest MED citation date (manuscript or composition) is selected for all the words with French connections in the lexicon, some words remain whose first date of citation in the MED postdates the Auchinleck Manuscript's manufacture. Our lists of words whose earliest MED manuscript citation postdates the 1330s need to be treated with caution: errors in matching are especially frequent there. Nevertheless, it is possible to pick out several Auchinleck items that predate the MED's earliest manuscript citations. For example, including citations of the Auchinleck words *avengy*,

enviroun, *scomfitour*, *trappour*, and *turmentri* in the MED would permit the antedating of its entries s.vv. *avengen* (earliest citation date by manuscript is currently 'c1380'), *enviroun*, adv. (currently 'a1400'), *scomfitour* (currently 'c1440'), *trappour* (currently '? a1425'), and *tormentri* (currently 'c1415').

5.3 French lexis in action

The two previous subsections of this article point out that the degree to which words with French connections are used in the Auchinleck Manuscript changes from text to text and that a good proportion of the book's French lexis was not often found in English in written contexts prior to 1300. We suggest that French lexis is more likely to have been perceived as different from English vocabulary where it appears in contrast to the English wordstock and where its use in written English is more novel. The final visualizing tool that we have developed for gaging the perception of French lexis in the Auchinleck Manuscript allows these claims to be developed with reference to the manuscript's individual texts.
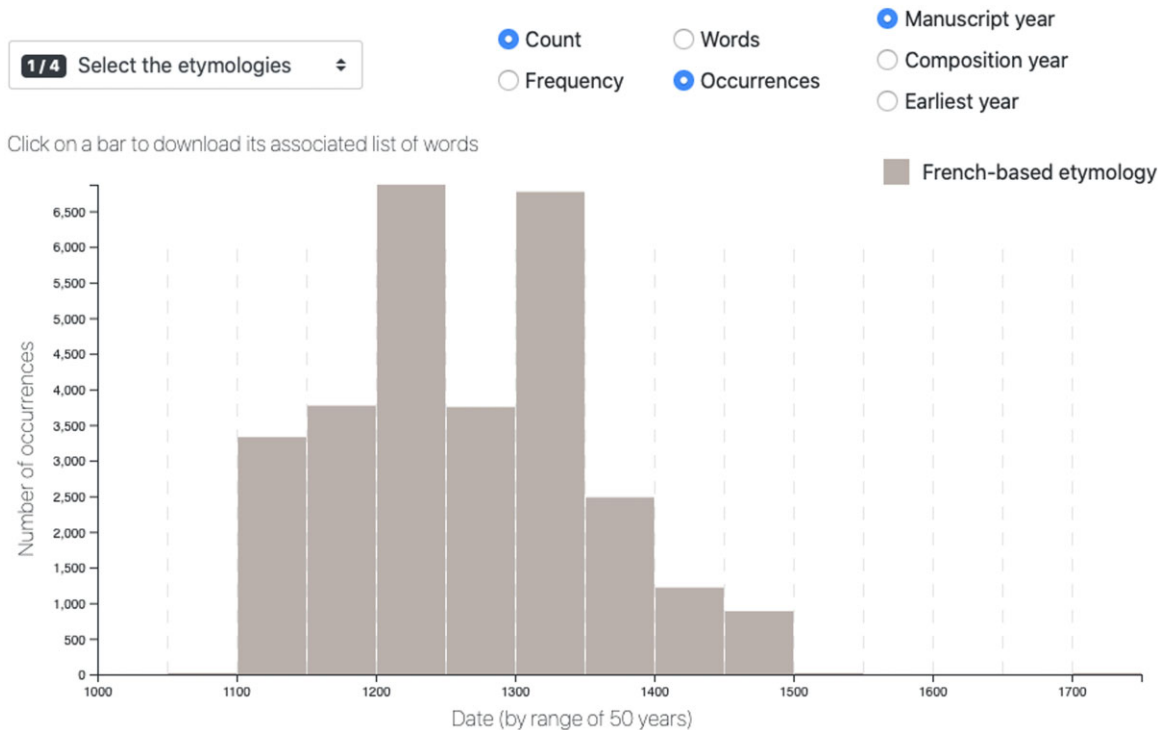
Fig. 2 All Auchinleck Manuscript Words with French Connections by Date of First Manuscript Citation.

Chart 3 on our project website is a scatterplot showing where words identified by the MED as having French connections occur in the individual Auchinleck texts and, again according to the MED, how recent their use in English writing is. The horizontal axis has a position for every word in a given text in the order in which the words appear there; the vertical axis shows the dates at which a word identified as having French connections is first cited in the MED. Where a French word is registered, a red dash is added above the position of that word in the text. The length of the dash corresponds to the duration of the date of first citation—either by manuscript or composition date—that is given in the MED. Users of the chart on the website can see which words are recorded as well as the passages in the poem where they are to be found by letting the cursor hover over the red lines and, if necessary, zooming in. These charts show us where words having French connections cluster as well as where newer French lexis is to be found.

Separate charts can be made for each of the manuscript's texts on the project website.

Here we demonstrate the potential of our Chart 3 with reference to three Auchinleck texts showing different kinds of French lexis in action.

The Thrush and the Nightingale is a debate poem in which a narrator reports a conversation that he claims to have overheard between the eponymous birds one spring. The topic of their discussion is women: the thrush accuses women of inconstancy and falseness and, in response, the nightingale defends them. The three clusters of French lexis that we highlight in Fig. 3 correspond to the nightingale's three speeches; the third speech is incomplete because the text breaks off here (five leaves are missing in the manuscript after the opening of *The Thrush and the Nightingale*). All the other French words in the poem can be attributed either to the narrator, who establishes the context for the debate in the poem's first twenty-four lines, or to the paratext that
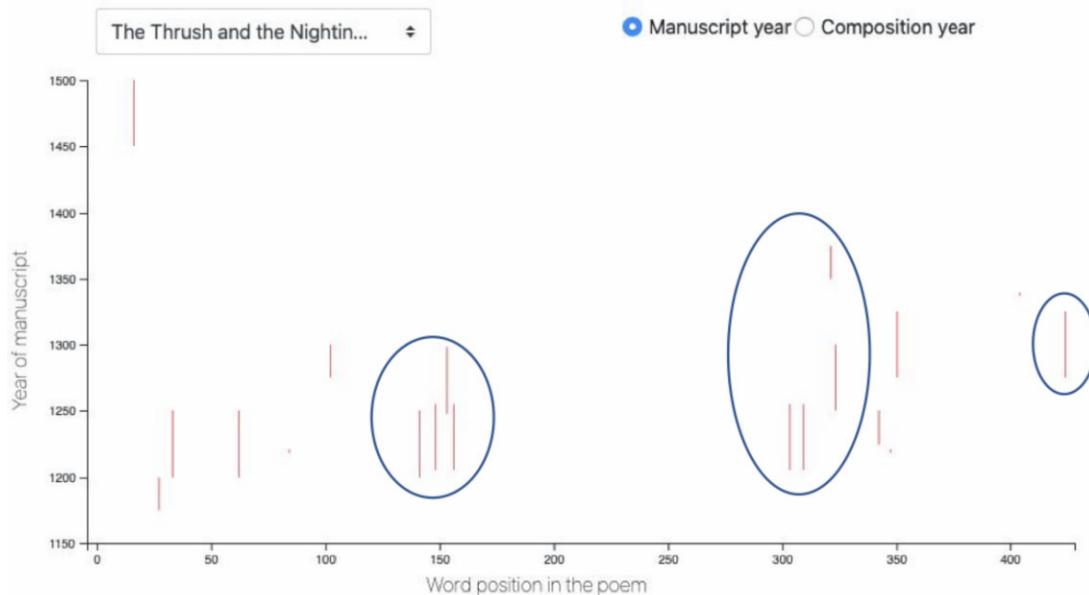
Fig. 3 Distribution and First Citation Dates of Words Having French Connections in *The Thrush and the Nightingale*.

announces the text's changes in speaker. No words having French connections are spoken by the thrush.

As a sample of the nightingale's language, we cite from her second speech in defense of women. Throughout this section, when quoting from the manuscript's texts, we give words with French connections in bold followed in square brackets by their earliest dates of manuscript citation as listed in the MED:

> 'Þrustelkok þou art wode
> Or þou canst to litel gode
> Wimen for to schende.
> It is þe best **drurie** [c1230]
> & mest þai cun of **curteisie** [c1230],
> Nis noþing also hende.
> ¶ Her loue is swetter, ywis,
> Þan þe **braunche** [c1300] of **licoris** [c1275];
> Lofsum þai ben & hende (ll. 49–57).

Translation: Thrush you are mad/or you have too little sense of what is good/if you blame women./Theirs is the best love-making/and they are masters of courtesy,/there is nothing so refined./Their love is sweeter, indeed,/than

the branch of liquorice;/they are beautiful and courteous.[18]

That the nightingale's speech is peppered with words having French connections is unsurprising insofar as courtesy, the virtue that she defends, achieves its fullest expression in the French-language literature of the Middle Ages, an important part of which was written in England. Many of her words—e.g. *drurie*, *curteisie*, *licoris*—were well established in English written contexts by the 1330s. What is interesting for our purposes is that, within the fictional universe of the poem, the contrasting personalities of the thrush and the nightingale are given words having different etymological backgrounds: there were more than enough Middle English insults with French connections that the thrush might have been given (see e.g. Rothwell, 1996). What the clusters of citations in Fig. 3 allow us to see is that the use of French lexis could sometimes be a matter of characterization.

Our second scatterplot shows a density of newer French lexis first cited in the MED c. 1300 clustering between words 2,500 and 3,000 in Auchinleck's
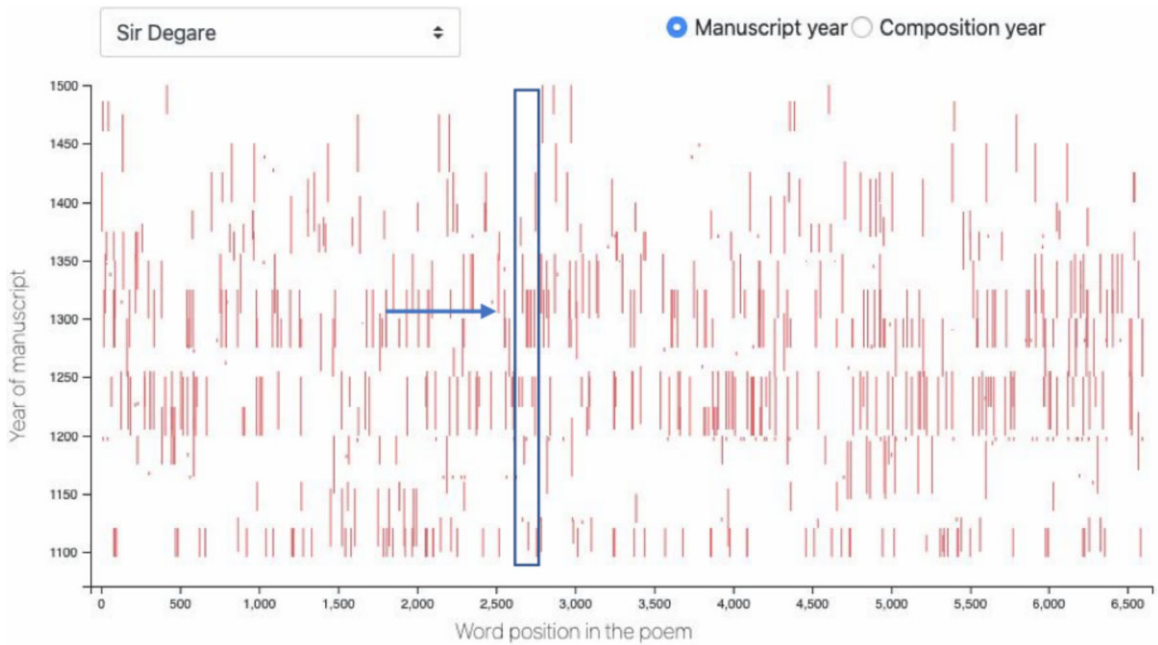
Fig. 4 Distribution and First Citation Dates of Words Having French Connections in *Sir Degaré*.

romance of *Sir Degaré* (Fig. 4). This portion of the poem comes in the course of a meeting on the road between the hero and an impressive group of earls and barons. From a sergeant attending these men, Degaré learns that the earls and barons have just left the court of a king who has determined to give his daughter in marriage only to a man who can unhorse him in a joust (Degaré will not learn until later that the daughter in question is his own long-lost mother). Here, we give the opening of the sergeant's speech:

> 'Sire' [c1225] he seide 'verraiment [c1300],
> We come framward a parlement [c1300].
> Þe king a gret counseil [a1126] þer made
> For nedes þat he to don hade.
> Whan þe parlement [c1300] was plener [a1325],
> He lette crie [c1225] fer and ner,
> ȝif ani man were of armes [a1250] so bold
> Þat wiȝ þe kinge iusti [c1330] wold,
> He sscholde haue in mariage [c1300]
> His dowter and his heritage [c1225] (ll. 433-42).

Translation: 'Sire,' he said, 'truly,/we come from a parliament./The king made a great council there/to suit his pending requirements./When the parliament was assembled,/he had it proclaimed far and near,/that if any man were so bold in arms/as to joust with the king,/he should have in marriage/his daughter and his inheritance.[19]

Here too French lexis contributes to the depiction of character: the vocabulary given to the sergeant helps to establish both his identity and the identity of the court from which he has come. In the passage cited above, the sergeant's first two words—*sire* and *verraiment*—could be used in either French or English, although the overlapping of *verraiment* into English appears to be more recent. The remainder of his reply to Degaré has several content words and phrases in French that reflect the idiom and preoccupations of 14th-century English courts: *parlement* and *plener parlement, counseil, letten crien, armes, jousten, mariage, heritage*. Since the

sergeant is reporting what he has heard, it seems that, like contemporary English courts, the language of Degaré's grandfather's court is French. This impression will be confirmed subsequently when the king begins his first address to *Degaré* with the words 'De par deus [...] he is welcome' (l. 478: by God he is welcome).[20]

As was the case with the French lexis used in the *Thrush and the Nightingale*, the French words highlighted in the quotation from *Sir Degaré* are shared with continental French: entries for all of these words can be consulted in the *Dictionnaire du Moyen Français* (*DMF*; http://zeus.atilf.fr/dmf/).[21] The particular uses of the sergeant's words within England will have reinforced their legal and administrative connotations there by the time they are taken over into English: especially in writing, French was a language of legal and bureaucratic action and record (see further Rothwell, 1992, 1998, 2006). A rich selection of specialist definitions reflecting these patterns of usage is recorded s.vv. *parlement, plener, conseil, heritage* at the *Anglo-Norman Dictionary*, which documents French lexis occurring in texts written in England (AND; https://www.anglo-norman.net).

What this extract in *Sir Degaré* allows us to see is the appropriation of specialist French lexis from the fields of the law and the royal administration for the purposes of Middle English poetry. Recognition of the particular history of the French lexis in this passage helps us as 21st-century readers to round out our appreciation of the effects of this sergeant's words on an early 14th-century audience. The *Degaré* poet's word choices establish his character's official credentials and contribute to the impressive effect created by the introduction of the band of travellers to whom he belongs. Elsewhere in the text, these men are called 'gret folk' (l. 428: grand people).

Our last scatterplot highlights a further use of French lexis in the damaged text of the Auchinleck *Kyng Alisaunder* (Fig. 5). As has already been pointed out, the density of French words in these fragments is comparatively high. We highlight a cluster of newer French words first cited in the MED c. 1300 occurring around the poem's word 490. This passage comes toward the beginning of the surviving portion of the text, corresponding to lines 6,780–89 in Smithers's 1952–57 edition of the fuller copy in Oxford, Bodleian Library MS Laud Misc. 622. The lines in question give the introductory description of the

marvelous trees of the sun and the moon in India, which Alexander visits towards the end of his story and from which he learns of his impending death by poisoning:

> Alisaunder so rideþ & wendeþ
> What he comeþ to þe trewes hende;
> Now þe **muge** [a1350] & þe **cetewal** [c1230]
> On hem smelleþ & þe **galingale** [? c1335].
> Þe **canel** [c1275] & þe **licoriis** [c1275]
> Swete **odour** [c1330] ȝiueþ, ywis,
> **Gelofre** [? c1335], **quibibbe** [? c1335] & þe **mace** [? c1335]
> **Gingeuer** [a1200], **comin** [? a1200] ȝeueþ **odour** [c1330] of **gras** [c1225].
> & vnder sonne of alle **spice** [c1230]
> He ȝaue **odour** [c1330] wiþ **delice** [c1230] (ll. 6780-89)[22]

Translation: Alexander rides so and makes his way/that he arrives at the noble trees;/now the musk and the setwall/he smells and the galangal./Cinnamon and liquorice/give off a sweet smell indeed./Cloves, peppercorns, and nutmeg,/ginger, and cumin give off a pleasant smell./And of all the spices under the sun/he smelled with delight.

In *Kyng Alisaunder*, the atmosphere of strangeness, exoticism, and tense anticipation which attends Alexander's visit to the trees of the sun and the moon is enhanced by the introduction of lexis belonging to the international French of the spice trade, whose use extended beyond Europe into the Middle East and Africa (see further Rothwell, 1999). Several of these words look to have been unfamiliar in English literary contexts c. 1330: *muge, galingale, odour, gelofre, quibibbe,* and *mace*. Differences between the extant texts of *Kyng Alisaunder* substantiate the MED's data regarding the novelty of one of these words: *odour*.

Auchinleck is the earliest extant witness to *Kyng Alisaunder*; further copies survive in the previously mentioned Laud MS and in London, Lincoln's Inn MS 150, both of which date to the later fourteenth century (so Smithers, 1952–1957, II. pp. 1–8). Where the Auchinleck text has the newer French word *odour* in three places (l. 6785, l. 6787, l. 6789),
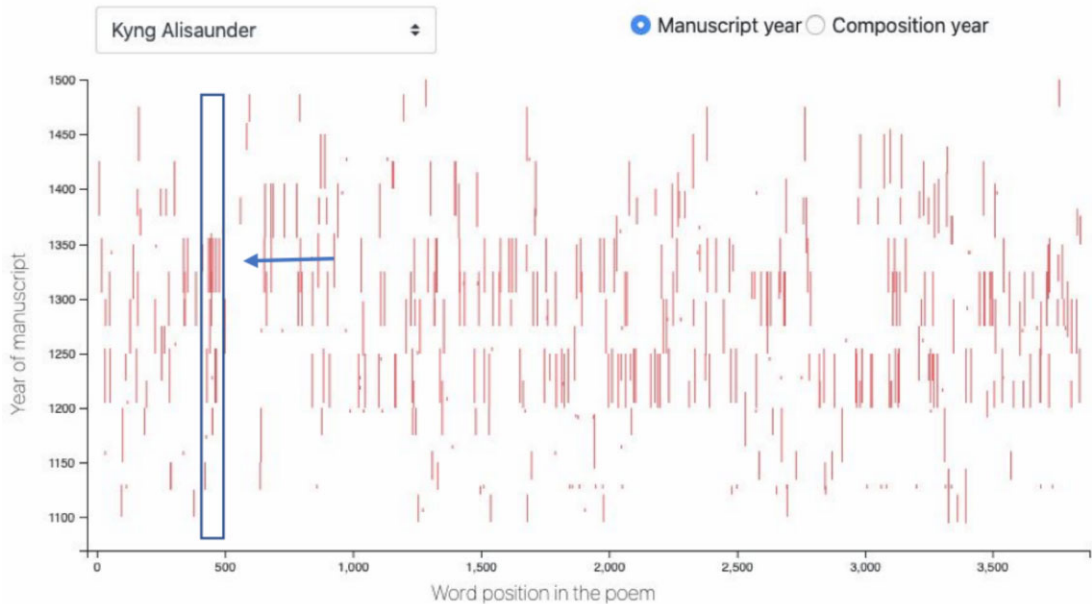
Fig. 5 Distribution and first citation dates of words having French Connections in *Kyng Alisaunder*.

the London text gives *sauour* (l. 5519), *odour* (l. 5521), and *sauour* (l. 5523), and the Oxford text gives *flauour* (l. 6785), *odour* (l. 6787), and *sauoure* (l. 6789) (cited from Smithers, 1952–1957, I. pp. 360–61). This variance suggests some uncertainty about the currency of *odour* in 14th-century Middle English: on three occasions the later scribes opt for *sauour*, a word with a long history of use in English that the MED can date back to a manuscript produced in the early thirteenth century (see MED s.v. *sauour*).

These data reinforce the impression given in the MED of the novelty of *odour* in English contexts.[23] Variance of this kind does not affect the passages from *The Thrush and the Nightingale* and *Sir Degaré* discussed above. Where the passages from these poems that we have analyzed can be compared with parallel passages extant in other manuscripts, all the extant copies agree in their use of French lexis.[24]

The passages analyzed above show French lexis being used variously to embellish the courtly arguments of a nightingale, to confirm the official credentials of a sergeant, and to enhance the atmosphere of wonder and strangeness attending a moment of romance magic. In each of the instances that we have

examined, poets appear to have drawn on the associations having accrued to the French words that they deploy owing to their uses in England and elsewhere in the pursuit of courtly love, royal administration, and international trade. Our other scatterplot visualizations highlight moments in the Auchinleck texts that parallel those just discussed. For example, the use of newer French lexis to create an atmosphere of surprise and wonder, as in *Kyng Alisaunder*, would appear to have been a reasonably regular procedure and has been commented upon previously by critics of individual Auchinleck texts (see e.g. Lerer, 1985; Easting, 1988).

The poetic procedures that we have highlighted suggest a growing sensitivity to the registers of expression newly available within English thanks to its uptake of French lexis. In this regard, our findings agree with recent work by Ingham (2017, 2018) considering the significance of the French lexis deployed in earlier Middle English works whose applications may have included preaching. But where Ingham argues that the appearance of French words in his texts demonstrates their thorough naturalization, the information that we have extracted from the MED suggests to us that many

of the lexical items that we have highlighted will have been more salient to readers of Middle English literature in the 1330s. From our perspective, it is possible to see the Auchinleck poets participating in the creation of new English registers (e.g. administrative English) as well as exploiting developments in the lexicon that were substantially complete by the early 14th century (e.g. courtly English).[25]

In none of the texts that we have looked at is French presented as the specific property of France. For the development of claims along these lines, we must turn to the prologs of the two Auchinleck romances that were mentioned above. Our penultimate section addresses these well-known texts and returns to our opening claims about the possible links between languages and nation in the 1330s.

6 Metalinguistic commentary in *King Richard* and *Of Arthour and of Merlin*

The first of the passages that concerns us opens *King Richard*, Auchinleck's romance giving the deeds of England's Richard I on the Third Crusade (1189–1192). After an opening address to Christ recalling the adventures and victories sent to Richard, the poet goes on to mention that clerks write books in Latin that are read by the Germans and the Picards and that 'folk of Fraunce' (l. 10) write romances treating the stories of Roland, Oliver, and the Twelve Peers; Alexander and Charlemagne; Hector; Ogier the Dane; and Arthur and Gawain. The language in which these romances are written poses a problem that the poet proposes to remedy:

> As þis romaunce of Freyns wrouȝt,
> Þat mani lewed no knowe nouȝt,
> In gest as-so we seyn.
> Þis lewed no can Freyns non—
> Among an hundred vnneþe on,
> In lede is nouȝt to leyn.
> Noþeles, wiþ gode chere
> Fele of hem wald yhere
> Noble gestes, ich vnderstond,
> Of douȝti kniȝtes of Inglond.

> Þerfore now ichil ȝou rede
> Of a king douhti of dede:
> King Richard, þe werrour best
> Þat men findeþ in ani gest (ll. 19–32).

Translation: These romances are made in French,/which many untaught people don't know at all,/as the story goes./These untaught people don't know any French—/hardly one among a hundred of them,/to tell the truth before you./Nevertheless, gladly/would many of them hear/noble tales, I've heard,/of doughty knights of England./Therefore I will now tell you/about a king doughty in deeds:/king Richard, the best warrior/that may be found in any story.

These lines imagine an audience for *King Richard* that knows what it wants and what it is missing: romances of English heroes. Because they do not know French, they will need such narratives telling to them in English, and this is where our poet steps in.

Passages like this one are popular with historians of the English language, who typically describe the early 14th century as a period heralding the return to prominence of English at the expense of French. An extract from Auchinleck's *Of Arthour and of Merlin* that echoes some of the points made in the prolog to *King Richard* is especially popular in the standard manuals: Baugh and Cable (2013, pp. 138–43) present it alongside extracts from texts including the *Cursor mundi* (c.1300), the *Northern Homily Cycle* (c.1315), and William of Nassington's *Speculum Vitae* (c. 1350), whose authors likewise discuss their reasons for writing in English.[26]

In *Of Arthour and of Merlin*, the poet appears to address a specifically noble English audience that is ignorant of French:

> Riȝt is þat Jnglische vnderstond
> Þat was born in Jnglond.
> Freynsche vse þis gentil man
> Ac euerich Jnglische Jnglische can.
> Mani noble ich haue yseiȝe
> Þat no Freynsche couþe seye.
> Biginne ichil for her loue—
> Bi Ihesus leue þat sitt aboue—

On Inglische tel mi tale:
God ous sende soule hale! (ll. 21–30).

Translation: Its right that a person understands English/who was born in England./Gentry men use French/but every English person knows English./I've seen many nobles/who couldn't speak any French./I'll begin for their sake/—with Jesus's leave, who sits above—/to tell my tale in English:/God send us salvation!

To the 21st-century reader, the idea that the English will want to read in English may seem so self-evident as to require no further commentary. What the foregoing analysis shows, however, is the tendentiousness of this argument in the 1330s, when the interpenetration of French and English is peaking and the boundaries between the two languages are far from clear. This is why we have held over these passages until such a late stage in our argument: it allows us to demonstrate that the familiar claims that *King Richard* is making would better be viewed as early attempts to yoke language to nation than as accurate statements of sociolinguistic reality.[27] While we have highlighted a few moments where readers might have detected the influence of French in Middle English literary contexts, the use of French cannot be assigned exclusively to the French at this stage in the history of England's two vernaculars. The unreliability of the prologs to both *King Richard* and *Of Arthour and of Merlin* as sociolinguistic documents is further underlined by Field (2010), whose study of romance manuscripts circulating in 14th-century England shows that, in most cases, French- and English-language versions of the same text were in use simultaneously during this period.

We are also keen to nuance the traditional interpretation of the opening to *Of Arthour and of Merlin*. Less frequently cited than the lines given above is this commentary, which directly precedes our last citation from the poem. Here, the continuing desirability of French and Latin among the English is stressed:

Childer þat ben to boke ysett
In age hem is miche þe bett
For þai mo witen & se
Miche of Godes priuete
Hem to kepe & to ware

Fram sinne & fram warldes care,
& wele ysen, ȝif þai willen,
Þat hem no þarf neuer spillen.
Auauntages þai hauen þare
Freynsch & Latin eueraywhare (ll. 9–19).

Translation: Children that are set to book (i.e. put to learning)/are much the better for it when they reach maturity/because they are able to know and to see/much of God's divine secret/and thus to keep and to guard themselves/against sin and against the hardships of the world,/and to see well, if they want to,/that they need never be damned./Advantages in all of this,/everywhere, bring French and Latin.

The attitudes manifested toward French in the prolog to *Of Arthour and of Merlin* are thus especially fraught. French is said to be something formerly possessed but now lost, but French is also found to be worth teaching for the salvation of learners, and future uses are anticipated for it. This indeterminacy leaves open a broad range of possible interpretations for the passage's own French lexis. In particular, what is the significance of *auauntages* in the penultimate line of the last citation? MED s.v. *avauntage* does not cite the word in English writing before c.1300 and attention is drawn to it in the prolog by the unusual decision to put *auauntages*, the grammatical object of the clause, in first position. If the word is processed as a citation from French, what is the perceived tenor of that citation: benign approval of the argument just expressed? Pedantic insistence on the enduring value of French and Latin? Mocking resistance to the continuing claims regarding French and Latin influence asserted by the English sponsors of these languages? Understandings of this polyvalent word in the 1330s will probably have changed from moment to moment and from person to person.

7 Conclusions

Our last example shows that the close reading of isolated moments in individual texts will remain necessary and rewarding for those interested in addressing the status of French in 14th-century England as well as the debts that Middle English owes to French. In

particular, metalinguistic commentary of the kind discussed in our previous section will always require care: here we have argued that claims in *King Richard* that French belongs elsewhere should be approached cynically and that a fuller reading of the prolog to *Of Arthour and of Merlin* reveals a profound ambivalence toward the status of French in English contexts. An important goal of this article has been to detach considerations of language contact between English and French from questions of national allegiance, which, in the 1330s, have yet to come fully into focus.

Our data and our visualizations provide a larger context in which future local analyses of the use of French lexis in the Auchinleck Manuscript's texts might be conducted. In so doing, they establish new parameters within which the skill of the Auchinleck poets can be appreciated. The large-scale searches that we have conducted confirm anecdotal reports of the deep French debts of individual Auchinleck texts that accompanied editorial work throughout the 20th century; they help us to see the potential novelty of that lexis in Middle English writing; and they show us where French lexis clusters in the book's texts. At the same time, flaws in our results have indicated at least one way in which the MED's new online interface might be developed. We thus offer this article as a fresh example of the ways in which the application of new digital resources can not only enrich the pursuit of more traditional philological methods but also remind us of the sheer complexity of the objects of philological enquiry.

Acknowledgements

References

AND. *Anglo-Norman Dictionary*. https://anglo-norman.net (accessed 11 September 2020).

Barber, C. (1997). *Early Modern English*. 2nd edn. Edinburgh: Edinburgh University Press.

Battles, D. (2010). *Sir Orfeo* and English identity. *Studies in Philology* **107**: 179–211.

Baugh, A. C. (1935). The chronology of French loan-words in English. *Modern Language Notes*, **50**: 90–93.

Baugh, A. C. and Cable, T. (2013). *A History of the English Language*. 6th edn. London: Routledge.

Blake, N. F. (2002). The early history of, and its impact upon, the *Middle English Dictionary*. *Dictionaries*, **23**: 48–75.

Boffey, J. and Edwards, A. S. G. (2008). Middle English literary writings, 1150-1400. In Morgan, N. J. and Thomson, R. M. (eds), *The Cambridge History of the Book in Britain, Vol. 2: 1100-1400*. Cambridge: Cambridge University Press, pp. 380–90.

Bridges, V. (2016). Absent presence: Auchinleck and Kyng Alisaunder. In Fein, S. (ed.), *The Auchinleck Manuscript: New Perspectives*. York: York Medieval Press, pp. 88–107.

Burnley, D. (1983). *A Guide to Chaucer's Language*. Norman, OK: University of Oklahoma Press.

Burnley, D. (1992). Lexis and semantics. In Blake N. (ed), *The Cambridge History of the English Language, Vol. 2: 1066-1476*. Cambridge: Cambridge University Press, pp. 409–99.

Burnley, D. (2003). French and Frenches in fourteenth-century London. In Kastovsky D. and Mettinger, A. (eds), *Language Contact in the History of English*, 2nd edn. Frankfurt: Peter Lang, pp. 17–34.

Burnley, D. and Wiggins, A. (2003). *The Auchinleck Manuscript*. https://auchinleck.nls.uk (accessed 11 September 2020).

Butler, P. (2016). A failure to communicate: multilingualism in the prologue to *Of Arthour and of Merlin*. In Fein, S. (ed.), *The Auchinleck Manuscript: New Perspectives*. York: York Medieval Press, pp. 52–66.

Butterfield, A. (2009). *The Familiar Enemy: Chaucer, Language, and Nation in the Hundred Years War*. Oxford: Oxford University Press.

Butterfield, A. (2013). Fuzziness and perceptions of language in the Middle Ages: part 3: translating fuzziness: countertexts. *Common Knowledge*, **19**: 446–73.

Calkin, S. B. (2005). *Saracens and the Making of English Identity: The Auchinleck Manuscript*. London: Routledge.

Clark, C. (1966). *Ancrene Wisse* and *Katherine Group*: a lexical divergence. *Neophilologus*, **50**: 117–24.

Clifton, N. (2003). *Of Arthour and of Merlin* as medieval children's literature. *Arthuriana* **13**: 9–22.

Clifton, N. (2005). *The Seven Sages of Rome*, children's literature, and the Auchinleck Manuscript. In Classen, A. (ed), *Childhood in the Middle Ages and the Renaissance*. Berlin: De Gruyter, pp. 185–201.

Conlee, J. W. (ed) (1991). *Middle English Debate Poetry: A Critical Anthology*. East Lansing, MI: Colleagues Press.

Cannon, C. (1998). *The Making of Chaucer's English: A Study of Words*. Cambridge: Cambridge University Press.

Critten, R. G. (2019). The multilingual English household in a European perspective: London, British Library MS Harley 2253 and the traffic of texts. In Burger, G. D. and Critten, R. G. (eds), *Household Knowledges in Late-Medieval England and France.* Manchester: Manchester University Press, pp. 219–43.

Crofts, T. H. and Rouse, R. A. (2009). Middle English popular romance and national identity. In Radulescu, R. L. and Rushton, C. J. (eds), *A Companion to Medieval Popular Romance*. Cambridge: D. S. Brewer, pp. 79–95.

Curry, A. (2002). *The Hundred Years War, 1337-1453.* Oxford: Osprey.

Dekeyser, X. (1986). Romance loans in Middle English: a re-assessment. In Kastovsky, D. and Szwedek, A. (eds), *Linguistics Across Historical and Geographical Boundaries, 2 vols.* Berlin: Mouton De Gruyter, I. pp. 253–65.

DMF. *Dictionnaire du Moyen Français*: http://zeus.atilf.fr/dmf/.

Donaldson, E. T. (1970). Idiom of popular poetry in the Miller's Tale. In Donaldson, E. T., *Speaking of Chaucer*. London: Athlone Press, pp. 13–29.

Durkin, P. (2002). "Mixed" etymologies of Middle English items in *OED3*: some questions of methodology and policy. *Dictionaries* **23**: 142–55.

Durkin, P. (2014). *Borrowed Words: A History of Loanwords in English*. Oxford: Oxford University Press.

Easting, R. (1988). Some antedatings and early usages from the Auchinleck *Owayne Miles*. In Reeks, D. M. (ed), *Sentences for Alan Ward: Essays Presented to Alan Ward on the Occasion of his Retirement.* Southampton: Bosphoros Books, pp. 167–74.

Fein, S. (2014). The fillers of the Auchinleck manuscript and the literary culture of the West Midlands. In Meale, C. M. and Pearsall, D. (eds), *Makers and Users of Medieval Books: Essays in Honour of A. S. G. Edwards*. Cambridge: D. S. Brewer, pp. 60–77.

Fein, S. (ed) (2016). *The Auchinleck Manuscript: New Perspectives*. Cambridge: D. S. Brewer.

Fellows, J. (2008). The Middle English and renaissance *Bevis*: a textual survey. In Fellows, J. and Djordjević, I. (eds), *Sir Bevis of Hampton in Literary Tradition*. Cambridge: D. S. Brewer, pp. 80–113.

Field, R. (1999). Romance in England, 1066-1400. In Wallace, D. (ed), *The Cambridge History of Medieval English Literature*. Cambridge: Cambridge University Press, pp. 152–76.

Field, R. (2008). Romance. In Ellis, R. (ed), *The Oxford History of Literary Translation in English, Vol 1: to 1550.* Oxford: Oxford University Press, pp. 296–331.

Field, R. (2010). Patterns of availability and demand in Middle English translations *de romanz*. In Ashe, L., Djordjević, I., and Weiss, J. (eds), *The Exploitations of Medieval Romance*. Cambridge: Brewer, pp. 73–89.

Görlach M. (1981). The Auchinleck *Katerine*. In Benskin M. and Samuels, M. L. (eds), *So Meny People, Longages and Tonges: Philological Essays in Scots and Mediaeval English presented to Angus McIntosh.* Edinburgh: Middle English Dialect Project, pp. 211–27.

Hanna, R. (2016). Auchinleck 'scribe 6' and some corollary issues. In Fein, S. (ed), *The Auchinleck Manuscript: New Perspectives*. York: York Medieval Press, pp. 209–221.

Higgins, A. (2016). Sir Tristrem, a few fragments, and the northern identity of the Auchinleck Manuscript. In Fein, S. (ed), *The Auchinleck Manuscript: New Perspectives*. York: York Medieval Press, pp. 108–26.

Holford, M. (2006). History and politics in *Horn Child and Maiden Rimnild*. *Review of English Studies* **57**: 149–68.

Horstmann, C. (ed) (1887). The Early South-English Legendary. *Early English Text Society o.s.* **87**. London: N. Trübner.

Ingham, R. (2009). The persistence of Anglo-Norman 1230-1362: a linguistic perspective. In Wogan-Browne, J., Collette, C., Kowaleski, M., et al. (eds), *Language and Culture in Medieval Britain: The French of England c.1100-c.1500*. Cambridge: D. S. Brewer, pp. 44–54.

Ingham, R. (2012). *The Transmission of Anglo-Norman: Language History and Language Acquisition*. Amsterdam: John Benjamins.

Ingham, R. (2014). The maintenance of French in later medieval England. *Neuphilologische Mitteilungen*, **115**: 425–48.

Ingham, R. (2017). Middle English borrowing from French: nouns and verbs of interpersonal cognition in the *Early South English Legendary*. In Fenster, T. and Collette, C. P. (eds), *The French of Medieval England: Essays in Honour of Jocelyn Wogan-Browne*. Cambridge: D. S. Brewer, pp. 128–39.

Ingham, R. (2018). The diffusion of higher-status lexis in medieval England: the role of the clergy. *English Language and Linguistics*, **22**: 207–24.

Jesperson, O. (1905). *Growth and Structure of the English Language.* Leipzig: B. G. Teubner.

Lerer, S. (1985). Artifice and artistry in *Sir Orfeo. Speculum*, **60**: 92–109.

Lerer, S. (2008). 'Dum ludis floribus': language and text in the medieval English lyric. *Philological Quarterly*, **87**: 237–55

Lewis, R. E. (2007). *Middle English Dictionary: Plan and Bibliography.* 2nd edn. Ann Arbor: University of Michigan Press.

Libbon, M. (2016). The invention of *King Richard*. In Fein, S. (ed), *The Auchinleck Manuscript: New Perspectives.* York: York Medieval Press, pp. 127–38.

Lusignan, S. (2004). *La Langue des rois au Moyen Âge: le français en France et en Angleterre.* Paris: Presses Universitaires de France.

Machan, T. W. (2016). The individuality of English in the multilingual Middle Ages. In Kytö, M. and Pahta, P. (eds), *The Cambridge Handbook of English Historical Linguistics.* Cambridge: Cambridge University Press, pp. 407–23.

Macrae-Gibson, O. D. (ed) (1973–79). *Of Arthour and of Merlin*, 2 vols. Early English Text Society o.s. 268, 279. Oxford: Oxford University Press.

MED. *Middle English Dictionary*: https://quod.lib.umich.edu/m/middle-english-dictionary/dictionary.

Mossé, F. (1943). On the chronology of French loan-words in English. *English Studies* **25**: 33–40.

Nielsen, H. F. and Schøsler, L. (eds) (1996). *The Origins and Development of Emigrant Languages.* Odense: Odense University Press.

OED. *Oxford English Dictionary*: https://www.oed.com.

Olson, L. (2012). Romancing the book: manuscripts for "euerich Inglische." In Kerby-Fulton, K., Hilmo, M., and Olson, L. (eds), *Opening Up Middle English Manuscripts: Literary and Visual Approaches.* Ithaca: Cornell University Press, pp. 95–151.

Perryman, J. (ed) (1980). *The King of Tars.* Heidelberg: Carl Winter Universitätsverlag.

Putter, A., Jefferson, J., and Minkova, D. (2014). Dialect, rhyme, and emendation in *Sir Tristrem. JEGP*, **113**: 73–92.

Rothwell, W. (1979). Anglo-French lexical contacts, old and new. *Modern Language Review*, **74**: 287–96.

Rothwell, W. (1980–81). Lexical borrowing in a medieval context. *Bulletin of the John Rylands Library*, **63**: 118–43.

Rothwell, W. (1983). Language and government in medieval England. *Zeitschrift für französische Sprache und Literatur*, **93**: 258–70.

Rothwell, W. (1985). Stratford Atte Bowe and Paris. *Modern Language Review*, **80**: 39–54.

Rothwell, W. (1991). The missing link in English etymology: Anglo-French. *Medium Aevum*, **60**: 173–96.

Rothwell, W. (1992). Chaucer and Stratford Atte Bowe. *Bulletin of the John Rylands Library*, **74**: 3–28.

Rothwell, W. (1996). Adding insult to injury: the English who curse in borrowed French. In Nielsen, H.-F. and Schøsler, L. (eds), *The Origins and Development of Emigrant Languages.* Odense: Odense University Press, pp. 41–54.

Rothwell, W. (1998). Arrivals and departures: the adoption of French terminology into Middle English. *English Studies*, **79**: 144–65.

Rothwell, W. (1999). Sugar and spice and all things nice: from oriental bazar to English cloister in Anglo-French. *Modern Language Review*, **94**: 647–59.

Rothwell, W. (2000). Anglo-French and Middle English vocabulary in *Femina Nova. Medium Aevum*, **69**: 34–58.

Rothwell, W. (2002). *OED*, *MED*, *AND*: the making of a new dictionary of English. *Anglia*, **119**: 527–53.

Rothwell, W. (2006). Anglo-French and English society in Chaucer's *The Reeve's Tale. English Studies*, **87**: 511–38.

Schaffner, P. (2019). The *Middle English Dictionary* revenant. *Dictionaries*, **40**: 201–19.

Schleich, G. (ed) (1929). *Sire Degarre.* Heidelberg: Carl Winter.

Smithers, G. V. (ed) (1952–57). *Kyng Alisaunder*, 2 vols. Early English Text Society o.s 227, **237**. Oxford: Oxford University Press.

Summerfield, T. (2013). 'And she answered in hir language': aspects of multilingualism in the Auchinleck Manuscript. In Jefferson, J. A. and Putter, A. (eds), *Multilingualism in Medieval Britain (c.1066-1520).* Turnhout: Brepols, pp. 241–58.

Tadmor, U., Haspelmath, M., and Taylor, B. (2010). Borrowability and the notion of basic vocabulary. *Diachronica*, **27**: 226–46.

Thaisen, J. (2020). Standardisation, exemplars, and the Auchinleck Manuscript. In Wright, L. (ed), *The Multilingual Origins of Standard English.* Berlin: De Gruyter, pp. 165–90.

Trotter, D. (1996). Language contact and lexicography: the case of Anglo-Norman. In Nielsen, H. F. and Schøsler, L. (eds)*, The Origins and Development of Emigrant Languages.* Odense: Odense University Press, pp. 21–39.

Trotter, D. (2003a). Not as eccentric as it looks: Anglo-French and French French. *Forum for Modern Language Studies*, **39**: 427–38.

Trotter, D. (2003b). The Anglo-French lexis of *Ancrene Wisse*: a re-evaluation. In Wada Yoko (ed), *A Companion to Ancrene Wisse.* Cambridge: D. S. Brewer, pp. 83–101.

Trotter, D. (2010). Language labels, language change, and lexis. In Kleinhenz, C. and Busby, K. (eds), *Medieval Multilingualism: The Francophone World and its Neighbours.* Turnhout: Brepols, pp. 43–61.

Trotter, D. (2013). *Deinz certeins boundes*: where does Anglo-Norman begin and end? *Romance Philology*, **67**: 139–77.

Turville-Petre, T. (1996). *England the Nation: Language, Literature, and National Identity*, 1290–1340. Oxford: Clarendon Press.

Vale, M. (2001). *The Princely Court: Medieval Courts and Culture in North-West Europe*, 1270–1380. Oxford: Oxford University Press.

Wiggins, A. (2005). Imagining the compiler: *Guy of Warwick* and the compilation of the Auchinleck Manuscript. In Kelly, S. and Thompson, J. J. (eds), *Imagining the Book. Turnhout: Brepols*, pp. 61–73.

Wogan-Browne, J., Watson, N., Taylor, A., and Evans, R. (eds) (1999). *The Idea of the Vernacular: An Anthology of Middle English Literary Theory, 1280-1520.* University Park: Pennsylvania State University Press.

Notes

1 The influence of Turville-Petre's work shapes the only book-length study written on Auchinleck, which extends Turville-Petre's argument to consider ways in which Auchinleck texts that do not ostensibly treat England might still manifest English concerns (Calkin, 2005). Turville-Petre's study also informs a host of shorter contributions addressing individual Auchinleck texts (e.g. Holford, 2006; Crofts and Rouse, 2009; Battles, 2010).

2 The scholarship on this topic is vast. Individual studies are referred to as the argument develops. For an overview, see Durkin (2014, pp. 223–79).

3 For a summary account of the opening phases of the Hundred Years War, see Curry (2002, pp. 11–36).

4 The major pioneering work on this topic is Butterfield (2009). Our article expands on that study where it considers the Auchinleck Manuscript, which Butterfield leaves untouched.

5 Script data for this visualization application are published with the website; remaining project scripts will shortly be published via the University of Lausanne's Digital Humanities Tool Kit (https://dhtk.unil.ch/).

6 For our working lexicon, see Document 1 at the project website. Burnley and Wiggins's lexicon of the manuscript lists 16,672 items (https://auchinleck.nls.uk/editorial/lexicon_ad.html). Our longer list includes items that Burnley and Wiggins reject, such as partial words or reconstructed words, as well as a few other items that the Burnley and Wiggins's lexicon does not register for reasons unclear to us. Working with our own lexicon allows us to produce the text-by-text statistics given below.

7 For a full account of the modifications applied, see Document 2 at the project website.

8 Proper nouns were defined as words that were capitalized in the Auchinleck Manuscript's texts but not positioned at the beginning of a line and not preceded by the punctuation marks <!>, <?>, <.>, <:>, <: '>, <: '> or <'>. If a word occurred at least once in the manuscript without being capitalized it was also left off the list of proper nouns to be excluded from our corpus.

9 We do not attempt to distinguish between Anglo-French or Anglo-Norman and Continental French because the MED is inconsistent in its approach to this matter (see Lewis, 2007, pp. 12–14). Consultation of the *Anglo-Norman Dictionary* (*AND*) might allow for refinement on some points but in the final analysis, distinctions between the various regional forms of French are difficult to make. Our procedure is in line with a trend in historical linguistics that stresses the similarities between the Frenches used in Britain and on the continent (e.g. Rothwell, 1985; Trotter, 2003a, 2013; Ingham, 2009). Below we model how *AND*'s recording of legal and administrative terms might be used to enrich readings of Middle English literature.

10 For a physical description of the manuscript and an account of the language of its scribes, see Burnley and Wiggins (2003).

11 For a recent summary account of the provenance of the Auchinleck Manuscript arguing for its early ownership by a noble family, see Olson (2012). For the argument that Auchinleck was taken northwards much earlier than has traditionally been thought (the book has been in the Advocates' Library in Edinburgh since the eighteenth century), see Higgins (2016).

12 For commented editions of the relevant texts by Mannyng and Rolle, see Wogan-Browne *et al.* (1999).

13 Among other recent publications attending to the Auchinleck Manuscript and multilingualism, see Bridges (2016), Butler (2016), and Libbon (2016).

14 Work on the language of individual Auchinleck texts has gradually undone the early argument according to which the manuscript's contents were not only inscribed at a London location, but also composed there for inclusion in the book (see e.g. Görlach, 1981; Wiggins, 2005; Putter *et al.*, 2014). The most recent work on the book's language suggests that its main scribe worked with exemplars written in no more than four different hands, however (see Thaisen, 2020). On the manuscript's scribes, most recently, see Hanna (2016).

15 For these romances, a French source is either known to exist or posited. Even where French versions exist, however, translators' precise source texts cannot be determined, so it remains unclear whether French terms such as these in rhyme position have been taken over from an original or introduced by the translator to complete a line (both procedures were probably in operation at different moments). On the French sources of Middle English romance and their translators' procedures, see further Field (1999, 2008).

16 See, for example, the commentary in Donaldson (1970) on Chaucer's use of familiar French lexis to parodic effect in his Miller's Tale. More recently, see too Cannon (1998) for the argument that Chaucer repeatedly renovated his lexicon over the course of his career.

17 For example, Burnley (1983, p. 83, pp. 129–31) uses Auchinleck poems as a foil against which the skillfulness of Chaucer's use of English might be displayed. More recently, disparagement of the Auchinleck Manuscript's poems has been transmuted into the claim that they

might have made good children's literature (see e.g. Clifton, 2003, 2005).

18 We cite the Auchinleck Manuscript's texts from Burnley and Wiggins's (2003) online facsimile-edition. The translations are ours. Data for the highlighted words are drawn from MED s.vv. *druerie, courteisie, braunch, licoris*.

19 Data for the highlighted words are drawn from MED s.vv. *sire, verreiment, parlement, counseil, plener, crien, armes, justen, marriage, heritage*.

20 On the use of French by the English royal court, see Lusignan (2004, esp. pp. 155–217). On the use of French in English aristocratic courts, which were predominantly francophone until the end of the 14th century, see too Vale (2001, pp. 292–94).

21 See *DMF* svv. *parlement, plénier, conseil, crier, arme, jouter, mariage, héritage*. For the French words highlighted in the passage from the *Thrush and the Nightingale*, see the entries s.vv. *druerie, courtoisie, branche*.

22 Data for the highlighted words are drawn from MED s.vv. *muge, setewale, galingale, canel* n. 1, *licoris, odour, gilofre, quibibe, macis, gingivere, comin, grace, spice* n. 1, *delice*.

23 It should be pointed out too that the latest edition of the OED entry s.v. *odour* (December 2020) cannot date the word to a manuscript earlier than ?c1335.

24 Compare ll. 49–54 of the Auchinleck text of *The Thrush and the Nightingale* with ll. 73–78 of the edition of the poem from MS Digby 86 in Conlee (1991, pp. 237–48); and see the textual apparatus for ll. 433–42 of *Sir Degaré* in Schleich (1929, p. 85).

25 For a consonant argument according to which an early 14th-century English poet might participate in the production of language registers as well as reflecting their development, see Butterfield (2013: 452–54).

26 These extracts are edited with commentary in Wogan-Browne *et al.* (1999).

27 This argument might also be applied to a few passages of Middle English poetry where French speakers are apparently made to speak English with French accents. See further Summerfield (2013, p. 250–54) and Critten (2019, pp. 228–29).