

Tracking the history of words: changing perspectives, changing research

Philip Durkin

Anna Morpurgo Davies Lecture, 6 May 2022

Abstract

Traditionally, most words have been seen as having simple linear histories, with the earliest known attestation regarded as the date at which a word ‘entered the language’. Changing perspectives, especially from historical sociolinguistics and from detailed research on language varieties, are bringing different questions into focus. Whose language does a particular word belong to? How is it used differently by different speakers? How has this changed over time? Additionally, renewed etymological interest in the origins of complex words has prompted questions about how frequently words show convergent lines of development, polygenesis as opposed to monogenesis. This article examines some of the challenges and opportunities presented by such issues for one of the oldest tools in historical linguistics, the historical dictionary.

Keywords

Etymology, lexicography, lexicology, lexical borrowing, loanwords, multilingualism, polysemy, polygenesis, basic vocabulary, word frequency.

Author

Philip Durkin is Deputy Chief Editor at the Oxford English Dictionary, where he has led the dictionary’s team of specialist etymology editors since the late 1990s. Among his publications are *The Oxford Guide to Etymology* (2009) and *Borrowed Words: A History of Loanwords in English* (2014). He is editor of *The Oxford Handbook of Lexicography* (2016) and is currently editing a handbook of etymology in the same series.

1. Introduction

This article is based on the 2022 Anna Morpurgo Davies lecture organised jointly by the British Academy and the Philological Society. Anna Morpurgo Davies was for very many years the staunchest, most perceptive, and sometimes also the frankest of friends to the project on which I work, the *Oxford English Dictionary*, and this article is very respectfully dedicated to her memory, and in particular to the indelible mark she left on the *OED* and its staff.¹

I approach the topic of this article with two hats on, as a historical lexicographer, and as someone who does research on the history of English lexis and the history of the English lexicon, i.e., a lexicologist. To play a little with an old metaphor, you might say that that makes me simultaneously part ‘gamekeeper’, as lexicographer, and part ‘poacher’, as lexicologist. I do not want to imply that that is particularly a virtue, but it is certainly what I am, so I will endeavour to make a virtue of it. And there is one other thing I must stress at this point: as a lexicographer, I am part of a large and world-leading team. In a dictionary like the *OED* where individual entries are not ‘signed’ (and where numerous hands have invariably been at work) the team and its members can seem somewhat invisible, but the dictionary is only as good as it is through many people with different strengths and specialisms working together.

As a historical lexicographer and as a historical lexicologist, I am fundamentally interested in the histories of words, both as individual lexical items, each with their own histories, and as units in the lexicon, hence with interrelationships with one another. I am also interested in how two other perspectives touch on this: how words are actually used by speakers (which is fundamental to a descriptive approach to lexicography), and which speakers use words and in which contexts (which necessarily brings in perspectives from sociolinguistics and from the study of language varieties). I have a particular interest in loanwords, which brings in also perspectives from studies of bilingualism and multilingualism.

I highlight all of these perspectives not as a shop window for my own Jack-of-all-trades dabbling in far too many areas, but because the main theme running through this article is what implications all of these issues have for how historical lexicographers put together dictionary entries – and also for how the wider community of people interested in the history of words, who have many different but often overlapping perspectives, can best make use of data from historical dictionaries, especially in combination with other sources of information.

2. A word in its dictionary context

Fundamental to any historical dictionary is the presentation of the histories of individual words. Figure 1 shows the beginning of the quite large *OED* entry for one high-frequency word, *general* (adjective and noun), a Middle English borrowing from French and Latin; the history of this particular word, and the contexts of its borrowing into English, will be discussed in more detail later in this article. This entry has been revised as part of the *OED*’s programme of comprehensive revision (the results of which are collectively known as *OED3*, to distinguish them from the first

¹ She was also, of course, a ground-breaking lexicographer, having early in her career produced the first ever dictionary of Mycenaean Greek (Morpurgo Davies 1963), and throughout her career she remained closely involved with many dictionary projects, not least the *OED*.

The screenshot shows the OED entry for 'general'. At the top, the OED logo and 'Oxford English Dictionary' are visible, along with a search bar and navigation links. The entry title 'general, *adj.* and *n.*' is prominently displayed. Below the title, there are options to view the entry as an outline or full entry, and controls for quotations and keywords. The pronunciation is given as /ˈdʒɛn(ə)rəl/ for British and /ˈdʒɛn(ə)rəl/ for US. The entry includes a section for 'Forms' (partially hidden), a 'Frequency' indicator, and an 'Origin' section stating it has multiple origins from French and Latin. The 'Etymology' section is also partially hidden. The main definition section is headed 'A. *adj.*' and includes a sub-section 'I. With reference to things, collective unities, etc.' followed by a numbered list of meanings. A 'Thesaurus' link is visible. On the right side, there are informational boxes: one stating the entry was updated in the OED Third Edition in September 2009, another showing the previous version (OED2, 1989), and a 'Thesaurus' section listing related terms like 'general adaptation syndrome' and 'general contractor'.

Figure 1. The beginning of the *OED* entry for *general*, as it appears in a default view on oed.com, May 2022.

edition of the dictionary, published 1884–1928, its Supplements of 1933 and 1972–1986, and its second edition of 1989). The *OED3* entry was published online in September 2009, with other smaller changes incorporated since. Figure 1 in fact shows the beginning of this entry with various options selected, in particular one that conceals the full listing of historical spelling forms, and another that conceals the full discussion of the etymology, in order to enable the earliest recorded meaning in English to be displayed on the first screen of data that a reader sees.² The material that is toggled away in this way is substantial. The listing of spelling forms for *general* lists twelve distinct forms, with information on the period in which each is attested, and labelling (for instance of geographical region) where appropriate, while the equivalent section for many older words can run to hundreds of spelling forms. The etymology section runs to over a thousand words of documentation and analysis (including detailed presentation of relevant spelling forms and meanings in the donor languages, French and Latin), prefaced by a short summary for rapid orientation. (The main part of the etymology section, detailing the word's French and Latin etymons, is quoted in full in section 7 below.) The presentation of the word's main meanings in English, including phrasal uses, takes up another 12,000 words of text (including illustrative examples), followed by a presentation of the word's compound uses (taking up another 9,000 words of text in the main entry, while some semantically complex compounds are presented as separate dictionary entries).

²The precise mode of display of entries on oed.com is currently subject to review, and likely to change considerably within the next several years.

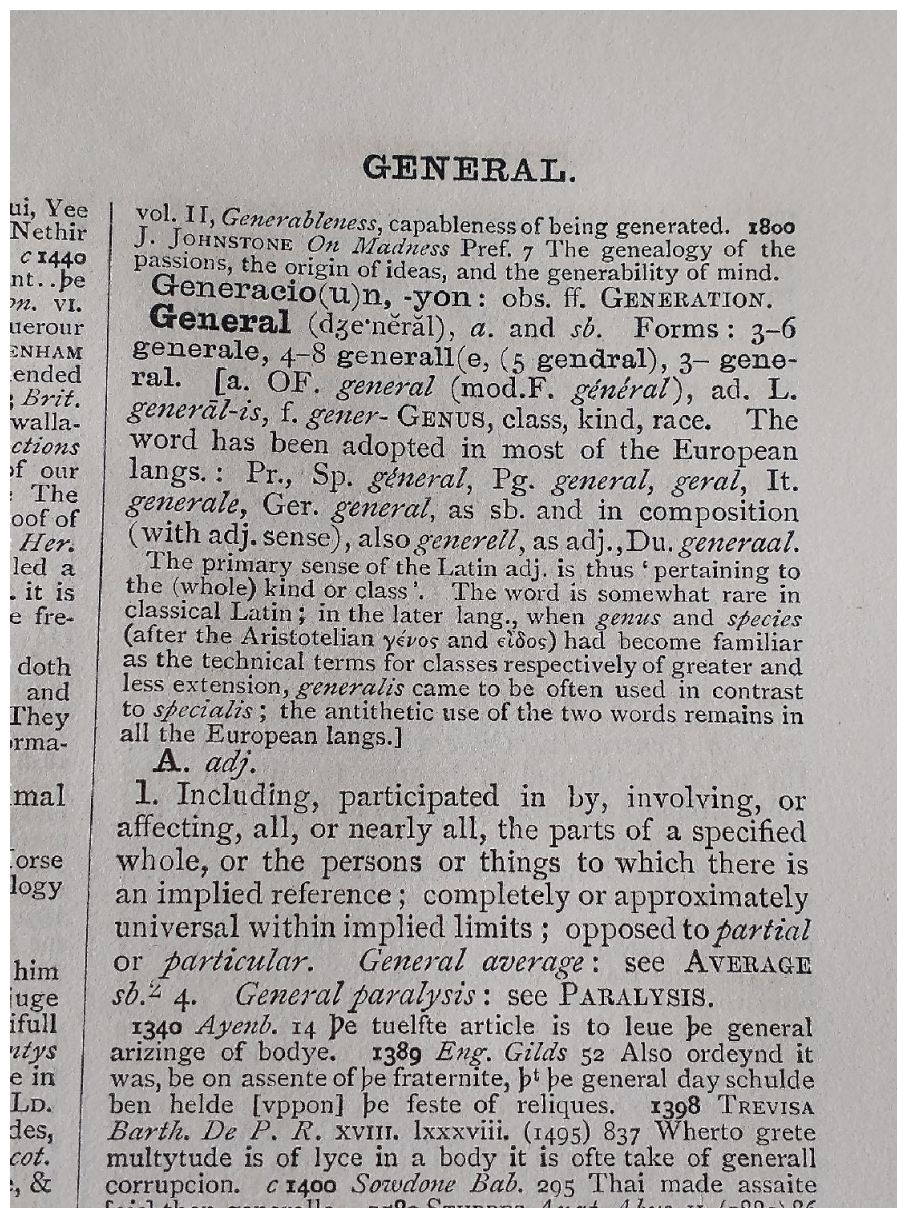


Figure 2. A detail of the page containing the beginning of the entry for *general* as it appeared in the first edition of the *OED*, in 1898.

In the digital environment, the scale of a historical dictionary entry can be difficult for readers to take in rapidly. Here the age of print did have some advantages. Figure 2 shows a detail of the page containing the beginning of the entry for *general* as it appeared in 1898 in the first edition of the *OED*, then still officially titled the *A New English Dictionary on Historical Principles*. There is no facility for toggling anything on or off here, the entry simply takes up approximately seven columns of space on pages 101 to 104 of volume IV, and the only way to find any of this content was to turn the physical pages to find it. In this format, space was always at a premium. A compressed style, with heavy use of abbreviations and other space-saving devices was often employed, which readers often find challenging today. Regardless of such space-saving devices, typically much less documentation, analysis, and explanation could be included in each entry: the complete

entry extended to a little over 9,000 words, including the treatment of compounds, which were (unlike in *OED3*) integrated among the main senses of the word, and with the etymology occupying only 125 words of highly abbreviated text. One of the virtues this limited mode of publication did bring was that readers were always confronted, physically, by the context: they could see that the space taken up by the entry *general* was dwarfed by the full size of the dictionary, but they could also see at a glance that it was much larger than most of the other dictionary entries nearby. Although the dictionary if printed today would take up many more pages than this, this bulk, and its contextualising value, is less immediately evident to many users. Some steps have already been taken to address this for the *OED*, by offering statistics on how each entry compares with others in the dictionary, in terms of factors such as average date of first attestation, average number of senses, or comparative frequency in contemporary English. Thus, in the case of *general* (adjective and noun), readers are currently informed that it is one of the earliest 4% of entries in the *OED*, that about 220 entries have first dates in the same decade (1230–39), and that *general* is among words that occur between 100 and 1000 times per million words in typical modern English usage, placing it in the second most frequent band of words, and among the most frequent 0.2% overall. Further innovations of this sort are likely in the future, but this general area remains an interesting challenge for future electronic iterations of the dictionary, and indeed for all electronic dictionaries.

If we remain a little longer with our single dictionary entry, *general*, in both the 1898 fascicle and in its modern revision we find a structure common to almost all entries in the dictionary. A modern English spelling form stands at the top of the entry as headword. The main pronunciation information is synchronic and contemporary, nowadays taking note of variation between major world varieties.³ The detailed presentation of historical spellings is a necessary companion to both the etymology and the main presentation of the word's meaning history in English, and can conveniently be presented separately from both. As already noted, the account of the word's etymology in the revised entry attempts to provide both a brief 'headline' account and more detailed (if still succinct) commentary – necessarily, since for many English words there is essentially no other detailed etymological account elsewhere, hence the *OED* must endeavour to satisfy many different user communities simultaneously. The main torso of the dictionary entry (and normally, even if not always, its largest part) consists of a presentation of the word's main meaning divisions, as determined by lexicographers examining source data, arranged in a chronological sequence, with senses and their descendant subsenses grouped under overarching sense branches. Figure 3 shows a graphic of the main sense divisions of the *OED* entry for *general* as adjective, with dates of first attestation for each of these main senses, generated from *oed.com*. It should be stressed that this a mid-level view of the *OED*'s sense structure, showing the overarching sense branches (numbered I and II) and the main senses within these, numbered with arabic numerals; including subsenses, there are in fact 19 separate senses given for the adjective, each with its own set of illustrative quotations, the first quotation given for each sense constituting the earliest available example of use in that sense. The sense structure presented by the *OED* is hence a deeply branching one, with multiple lines of semantic development. I would stress that even this detailed branching one of senses and subsenses is intended to be a presentation only of the main course of

³ Important historical developments in the pronunciation of individual words, especially when these are non-predictable from the main regular sound changes in the history of English, are often detailed in the etymology section. This is, of course, especially the case where hypotheses concerning the phonological development of a word are key to explaining its etymology and historical development. Such an example is provided by the *OED* entry for *feud* n.¹

Timeline of sense development

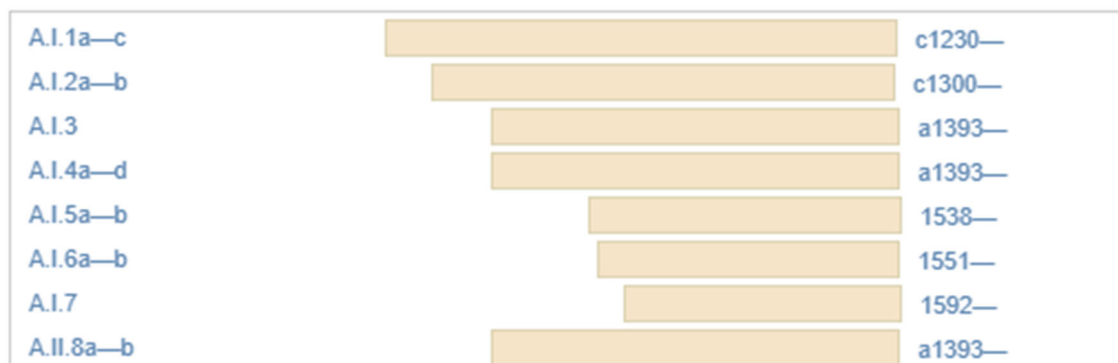


Figure 3. Graphic of the main sense divisions of *general* adj., with dates of attestation, generated from oed.com, May 2022.

the sense development: there are almost always cross-currents of influence between meanings that may be in different places in the structure, and the most important of these will often be dealt with explicitly by notes or other commentary. There is often also interplay with the meanings of etymologically or semantically related words within English, and also with the meanings of etymologically or semantically related words in other languages, a topic I will return to later in this article (see especially sections 7 and 10 below).

The narrative of each dictionary entry thus begins with etymology and the earliest attested uses and meanings, and moves towards the present day – in the case of a polysemous word, doing so in parallel for every sense that survives to the present. Nonetheless, there is an important sense in which the dictionary is anchored in the present day: the headword form standing at the head of each entry is a standard modern English spelling, and the wordlist is essentially that of contemporary English plus obsolete words. We might see the dictionary’s narrative as in a sense teleological, with the present day as the key destination for non-obsolete words.

3. The dictionary and lexical change: implications of lexical splits

For most dictionary entries this teleological perspective has relatively few practical implications, but for others it is far more significant. For instance, where lexical split has occurred, a dictionary anchored in a historical period prior to this split would have a different number of dictionary entries.⁴ The *OED* has separate entries for *metal* ‘metallic substance’ and *mettle* ‘spirit, character, courage’ because these are regarded as distinct words, with different standard spellings, in modern English, even though the one originated as a metaphorical extension of the other, probably in the early 16th century. We must then decide how to split the historical material, by spelling or by meaning: the *OED* places all instances meaning (crudely) ‘spirit, character, courage’ at the entry *mettle* even when those individual instances are spelt *metal*, which provides an efficient presentation, but there is really no right or wrong here, and a fair amount of commentary is required to explain the historical divergence in spelling forms to readers, as well as some agonising judgements

⁴On the process of lexical split see further Durkin (2006), Durkin (2009: 83–6); on the implications for historical dictionaries see Durkin (2016a: 247–51).

about where to place the many densely punning uses that occur in, for instance, Shakespeare. Almost certainly, a dictionary of Early Modern English (such as was once planned but never completed) would make different decisions here. Such phenomena also affect much less peripheral words: for instance, the *OED*, of course, has separate entries for the numeral *one* and the indefinite determiner *a*, *an*; so does the Michigan-based *Middle English Dictionary* (*MED*), but, like the *OED*, with copious commentary on its editorial decisions; the Toronto-based *Dictionary of Old English* (*DOE*), very properly, has just the one, very complicated, dictionary entry, since in the Old English period this individualising function was simply one of the functions of the undifferentiated word form *ān*.

4. Lexical polygenesis

The preceding section has identified one category of words that illustrate one of the main themes of this article, namely words whose histories in one way or another pose challenges for the linear narrative of a historical dictionary entry. The examples looked at in the preceding section showed diachronic split, with two distinct words resulting from the divergence of what was a single word at an earlier stage in the history of the same language. Much more common are cases where a single word results from the convergence of multiple inputs, whether these are (in linguistic terms) identical or distinct.

A term that may be found useful in investigating this theme further is polygenesis.⁵ I would contend that, rather than having a single historical point of origin from which all other uses emanate, it is quite normal and indeed usual for the identical word form to arise quite independently on different occasions as different language users satisfy the same or very similar communicational needs. This may seem like a lexicographer's hair-splitting, but it is relevant when we are thinking about, for instance, dictionary first dates. Except where we have deliberate coinage of a term, most frequently in scientific or technical discourse, what do we mean by a first use, and why should it be a particularly privileged piece of information? In many ways just as interesting, although even harder to ascertain, is when (if ever) a particular word form in a particular meaning first came to be common currency, known to the average language user. This becomes even more interesting, but also more complicated, if our interest in words and their histories has any sort of sociolinguistic or variationist slant, and we want to know who was using a particular word and in which contexts.

In a famous image on the first page of the 'General Explanations' issued with the first published parts of the *OED* in 1884 (see Figure 4), its illustrious early editor Sir James Murray sketched the lexicon as a common core surrounded by a periphery consisting of various language varieties (on some of which we may take different views today). If we imagine ourselves as standing at Murray's common core, I would suggest that most new words first come into view as flashes on the horizon, typically multiple flashes at slightly different points on the horizon. I often hear my *OED* colleagues use formulations like 'we know that at this date someone somewhere was using this word', which in my view is exactly right. Except in really exceptional circumstances, what we are observing will be only the tiniest glimpse of the extent of actual usage at this earlier point in time; additionally,

⁵ See further Durkin (2009: 68–73, 228), and, for the work on historical semantics by which my use of this term is inspired, Geeraerts (1997: 62–8).

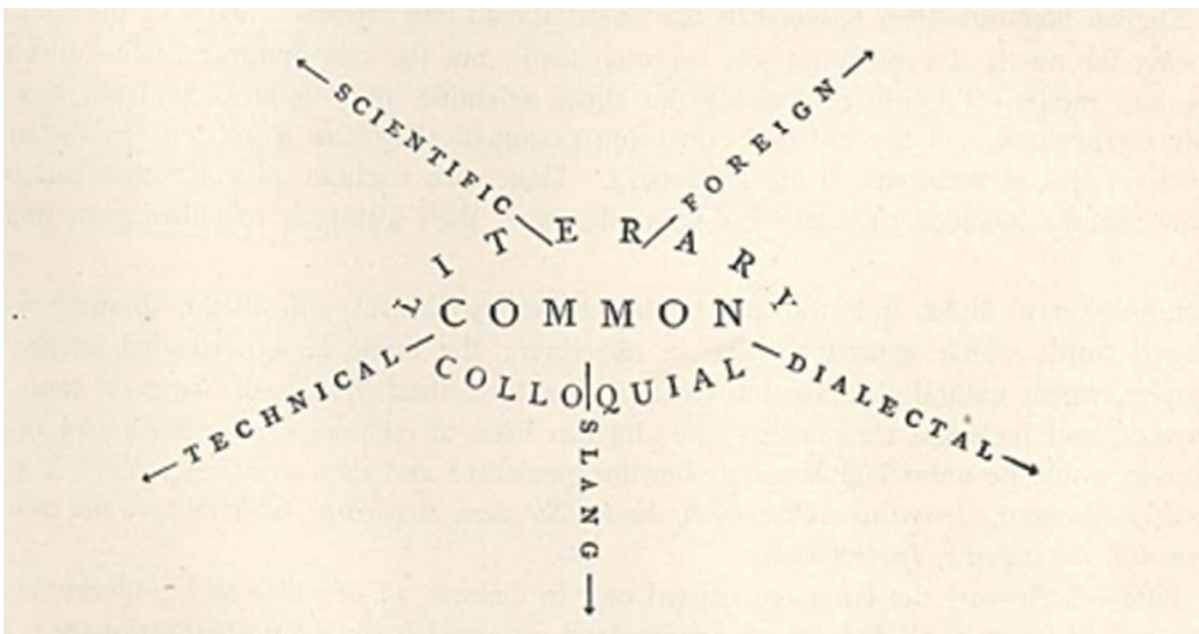


Figure 4. Diagram by Sir James Murray depicting ‘the circle of the English language’, reproduced from the first page of the General Explanations issue with the first published parts of the *OED* in 1884.

there is often no good reason for assuming that all of the uses we are aware of have any direct connection with one another, except in so far as all have arisen in more or less the same way, in response to the same or similar communicational needs. When some words begin to spread to the active or passive vocabulary of more and more speakers, which is the case at some point in time for nearly all of the words that have made the threshold for inclusion in the dictionary, it may well be the case that (to switch metaphors abruptly) a number of trickles of early usage converge into tributaries and even later great rivers of usage, but it is also highly probable that others do not, and we simply have dead ends.

This, then, is one type of what I would term polygenesis, i.e., the same word form arising in the same or closely similar meaning more than one time, with at least some of these independent ‘coinages’ starting up little chain-reactions of usage which eventually accrete to give a word in common use, indisputably part of the lexicon, part of the language. This type of polygenesis is typically difficult to prove outright (compare discussion of *scientifically* below); some of the strongest evidence comes from what are probably extreme outlier cases, where we find a huge discontinuity in the historical record. Thus, the *OED*’s best efforts have been unable to find (to draw a few examples from among many similar cases) examples of *appley* ‘resembling apples’ between the early 15th century and 1854, of *appliedly* ‘by or in practical application’ between the 17th century and 1901, of *piquantness* between 1733 and 1918, or of *effeminizing* between 1639 and 1898. Further intensive searching might perhaps close up these particular gaps in the record (particularly as more and more historical material becomes searchable online), but the fact that words of this sort can remain either entirely unattested or vanishingly rare for long periods for which very substantial written records survive suggests strongly that we are seeing polygenesis of the same word form (in each of these instances arising from highly productive word-formation processes) in broadly the same meaning, rather than a continuity that has simply eluded extensive searching of the written record.

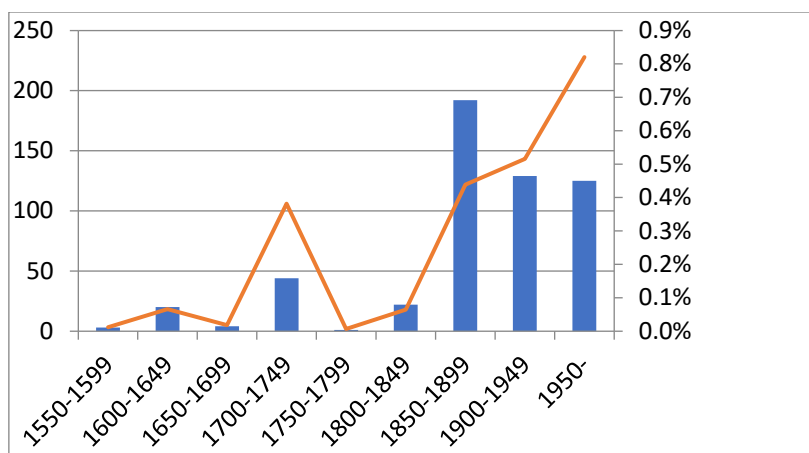
A second type of polygenesis is where two etymologically distinct inputs converge to give a single lexical item. One rather special type of this is lexical merger, where what were at an earlier stage of language history two distinct words merge to give a single word. For instance, the modern English verb *melt* results from the merger of two originally distinct verbs in Old English, a strong verb *meltan* (Class III, with the principal parts: past tense 1st and 3rd person singular *mealt*, past tense plural *multon*, past participle *gemolten*), and a weak verb *meltan* (with past tense and past participle formed with a dental suffix). The strong verb was originally intransitive, while the weak verb was originally transitive, ‘to melt (something)’, being in origin a causative formation, formed by means of a derivative suffix (causing *i*-mutation, a vowel harmony process, in the stem vowel) affixed to the same stem as is shown by the past tense singular of the strong verb. There was probably some confusion of the two verbs already in Old English, and in the Middle English period the two sets of forms are used more or less indiscriminately. In modern-day English there is one verb, *melt*, with regular past tense and past participle *melting* (showing the regular inflection developed ultimately from the weak verbs of Old English), in both transitive and intransitive uses, although a descendant of the original strong past participle form, *molten*, retains some currency, especially in adjectival use, designating metals or other substances liquefied by the application of heat.

In an instance such as this, we can observe a loss of formal distinctions over time, with two distinct words merging in one. Rather more commonly, we find that a single word form appears to reflect more than one input from the earliest stage of its history. Thus, *scientifically* probably arose multiple times, as people felt the need to express adverbially concepts corresponding to those expressed by the adjectives *scientific* and *scientific*; very probably when this happened some individuals had the adjective *scientific* more in mind, others *scientific*. Typically, if we look closely enough at the early evidence for such a formation, we will find it occurring sometimes paired with the adjective in *-ic*, sometimes with the adjective in *-ical*, and we can also compare changing patterns of frequency of use in these particular adjectives, and changing fashions in use of formations in *-ic* and *-ical* more widely. In a further complexity, not uncommon in derivative formations of this sort, *scientifically* shows two distinct meanings in its early history: the meaning ‘in a scientific manner; according to scientific methods’ that remains usual today, and a different meaning (first attested two decades earlier, in 1624) defined by the *OED* as ‘by means of demonstrative reasoning; so as to produce or provide axiomatic or certain knowledge; conclusively’, which is now rare (and used chiefly in discussions of Aristotle). These two meanings very possibly came into use independently of one another, and for both input from both *scientific* and *scientific* seems very likely.

It is cases like this, where both multiple instances of the identical coinage and convergence of two distinct formations seem both very likely to be at play, that seem to me to justify stretching the term polygenesis to cover both. When applied in this way to productive word-formation processes, this may seem to some simply a lexicographer’s rather perverse slant on processes for which morphologists have well worked-out frameworks, but I think that its virtue comes into play rather more when we turn to analogous situations in lexical borrowing, which is what I would like to do for the remainder of this article.

5. Loanwords from Japanese, and their implications for dictionary first dates

I would like to look briefly first at lexical borrowing from Japanese into English, as reflected by the *OED*. This is an interesting example because the main contours of the historical contact situation are relatively clear-cut, and enable us to see unusually clearly both the potential power of dictionary data and some of the potential pitfalls. In very crude summary, Japanese lexical borrowing into English begins in the early period of European missionary and trading contacts with Japan in the late 16th and early 17th centuries, is heavily impacted by the lengthy period of Japan's official policy of near-total closure to external contacts from the 1630s until the mid 19th century, and then shows an explosion of new borrowings from the 1850s onwards, when the forcible opening up of Japan to external influences leads to a Japan-mania in the anglophone world and elsewhere in the West.⁶ The bars in the graph in Figure 5 show the totals of Japanese loanwords recorded in the *OED*, arranged into fifty-year subperiods from the late 16th century (when the first loanwords are recorded) to the present day (with the last subperiod, from 1950 onwards, being somewhat longer than the others). A similar bar chart can be produced automatically from tools available on *oed.com*, and such graphics are frequently employed in analyses of the history of the English lexicon. It will be seen that this graph maps quite well to the crude historical summary I have given, although with a couple of blips, which I would like to probe a little more. Firstly, in my own work with what are sometimes termed 'lexicostatistics' of this sort, I find a useful first step is to look not only at the crude totals of loanwords per sub-period, but to look also at the percentage this constitutes of all new words recorded by the dictionary in the same period, as shown by the red line in Figure 5. As will be investigated later in this article, this is particularly important when we are looking at loanword stories that extend back to Middle English or Old English, but doing this for Japanese



- The bars show totals of Japanese loanwords in the *OED*, arranged by date of first attestation, in 50-year subperiods.
- The line shows these totals as a percentage of all new words recorded in the *OED* in each 50-year subperiod.

Figure 5. Totals of Japanese loanwords in the *OED*.

⁶For the main historical outlines see Hane (2000). For studies of Japanese loanwords in English see Cannon (1996), Evans (1997), Durkin (2014: 395–8), and especially Hayakawa (2014).

immediately reveals something interesting: the proportion of all new words actually goes up, not down like the crude total, in the period after 1900. I will not probe this further here, since doing so involves some careful thinking both about the *OED* as data source and about the recent history of Japanese influence on the English lexicon to which proper justice could not be done in the confines of this article, but I hope that it nonetheless highlights how different perspectives on the same dictionary data can prompt new questions.

In both graphs, there is a bit of an elephant in the room, in the form of a bulge in the sub-period 1700–1750. Investigating this a little further helps illustrate a couple of methodological points, as well as being interesting in and of itself. There are 46 words in this bulge, among them *adzuki*, *katakana*, *koi*, and *samurai*. All 46 of these words in fact have the same date of first attestation, 1727, and all first appear in the same source, John Gaspar Scheuchzer's *History of Japan*. So this in itself tells us something about the dangers lurking in sub-periodising our results if we are not diligent in how we then examine them. Scheuchzer's work is interesting. It is a translation of a then-unpublished German work by Engelbert Kaempfer relating his travels in Japan in 1690–2. Kaempfer's German work uses Japanese words to denote various Japanese cultural items and concepts, which Scheuchzer has rendered into English, retaining the same Japanese words. (This also raises the question of when we have a loanword and when we simply have the citation of a foreign-language word, a tricky area where the lines are somewhat fuzzy and distinctions less than categorical.)⁷ The contact between Japanese and a German-speaker (which may well have been via a third language) occurred in the late 17th century, a fascinating exception to Japan's general impenetrability to Europeans at this time, but these words were not relayed into English until over thirty years later, via Kaempfer's German text. For most of these 46 words, the next illustrative example that the *OED* offers comes from after the Meiji Restoration and the forcible opening up of Japan in the mid 19th century, and in all cases frequency of attestation increases markedly after the 1850s, and it is this later frequency of use that has earned these words a place in the dictionary's wordlist. Nonetheless, Scheuchzer's work was published, and read, and it is not impossible that there is some thread connecting these uses with later ones, albeit only via the written medium.

This example has therefore brought us on to some of the more general issues surrounding first dates of attestation, for many people clearly one of the great draws of the *OED*'s documentation, but always needing to be approached with due caution. A further Japanese example may help illustrate this. Eight *OED* entries for loanwords from Japanese, including some now reasonably familiar words such as *shogun*, *furo*, or *bento*, have first examples from entries from between 1614 and 1617 in the diary of Richard Cocks, a merchant and employee of the East India Company, who was in Japan for more than ten years, taking the leading role in the East India Company's ventures in Japan for most of that period.⁸ Cocks definitely had some direct exposure to the language, although we cannot be certain which language was employed in his interactions with Japanese people, including his Japanese mistress: some Japanese, some English, or a lingua franca, likeliest Portuguese, all seem possibilities. Few of these words with first examples from Cocks's diary are attested again at all frequently until the second half of the 19th century – for several the *OED* in fact offers a second example from Scheuchzer's translation of Kaempfer – and Cocks's influence

⁷For an introductory discussion of some of the practical difficulties, and theoretical issues with which they may intersect, especially with regard to the relationship between lexical borrowing and code-switching, see Durkin (2009: 173–77). Compare also note 10 below.

⁸See Farrington (2008) and also Farrington (1991).

on later use can essentially be ruled out, as his diary remained unedited and unpublished until 1883.⁹

I hope these examples highlight a few of the issues that can arise from uncritical use of dictionary first dates. Inspection of the sources in which first attestations occur, and the contexts of those sources, leads us to fascinating insights into early language contacts, but it is almost in the nature of first recorded attestations to be atypical. This is all the truer today, as more and more resources become available for electronic search by lexicographers and others, and we may get a little closer to actual very early uses, but, paradoxically, at the same time these earliest known uses often become further removed from the typical, and from the main currents of widespread adoption in the lexicon.

6. Middle English borrowing from Latin and French: contact situation and general outline

In the remainder of this article I look briefly at some aspects of lexical borrowing into English from French and Latin, particularly in the Middle English period, and especially in the light of factors examined above. In this instance the scale of borrowing is vast, and analysis of our copious but nonetheless imperfect data is challenging. In the Middle English period (roughly 1150 to 1500) there was intense contact between Anglo-Norman, British medieval Latin, and Middle English, in a society in which patterns of multilingualism were complex – and about which our knowledge is still imperfect. An important, and reasonably safe, generalisation is that for that segment of the (chiefly male) population whose daily life involved the use of literacy in the transaction of business, some degree of functional trilingualism was essential. In a typical pattern, while English was the mother tongue, a considerable, often native-like, proficiency in Anglo-Norman was also acquired, while Latin was also essential for the formal recording of legal business as well as in the Church and all areas of higher learning. Additionally, business records were often made in a mixed variety, in which Latin provided the matrix language, but most of the content vocabulary is typically English or Anglo-Norman, without Latin inflectional endings, and, significantly, without any clear demarcation between these two vernaculars. Simplifying crudely again, from the late 14th century onwards there are significant shifts in the patterns of language use, especially in writing, and particularly in terms of Anglo-Norman and mixed-language record keeping giving ground to written use of English.¹⁰

If we turn now to dictionary data, we can again look at the raw totals of loans per fifty-year subperiod, which gives the overall picture presented in Figure 6. Here we find that French loans show their highest peak in the second half of the 14th century, but the peaks and troughs are not enormously dramatic from the beginning of the 14th century through to the end of the 19th. By contrast, for Latin we appear to have the most dramatic occurrences in the Early Modern period, particularly in the earlier part of the Early Modern period, 1500 to 1650, and then again in the 19th

⁹Thompson (1883); now superseded by the edition issued by the Institute of Historical Research, University of Tokyo (1978–81).

¹⁰See the invaluable collections of studies Trotter (2000), Wogan-Browne *et al.* (2009), Ingham (2010), and Wright (2020). On the acquisition of Anglo-Norman, see especially Ingham (2012). On lone foreign-language items occurring in medieval texts, and the issues raised by mixed-language texts, in addition to numerous contributions in the collections cited, see also Ingham *et al.* (2020).

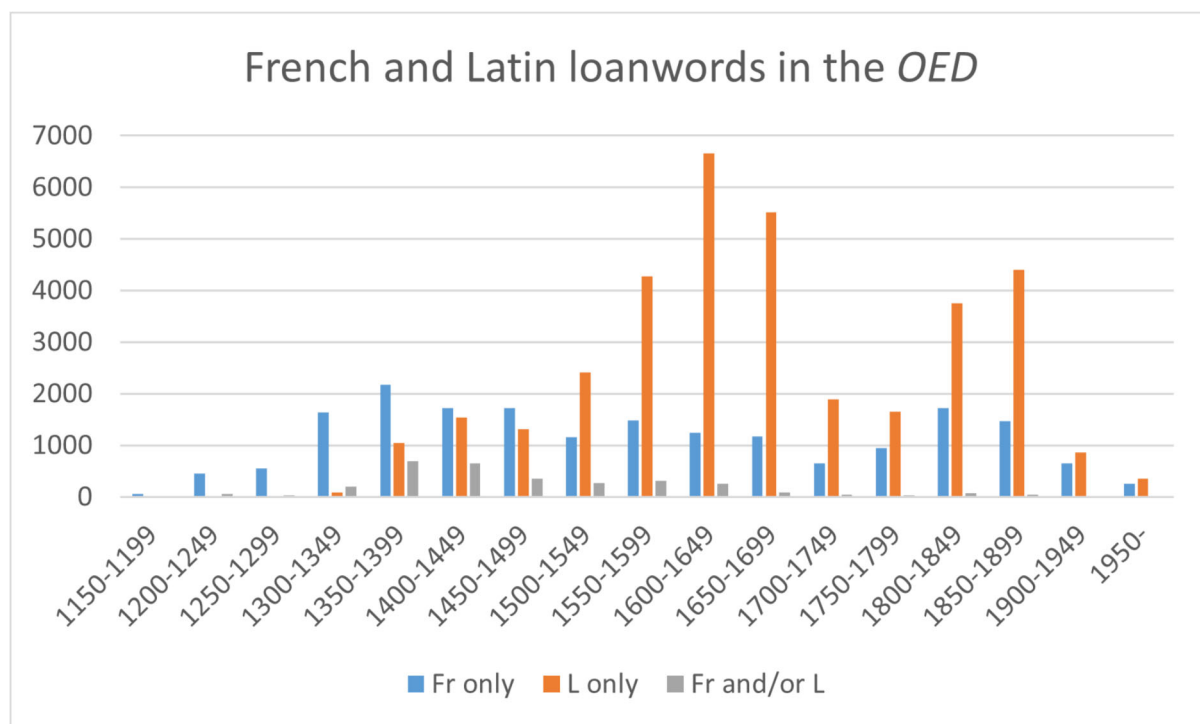


Figure 6. Loanwords from French, from Latin, and from French and/or Latin, as recorded by the *OED*.

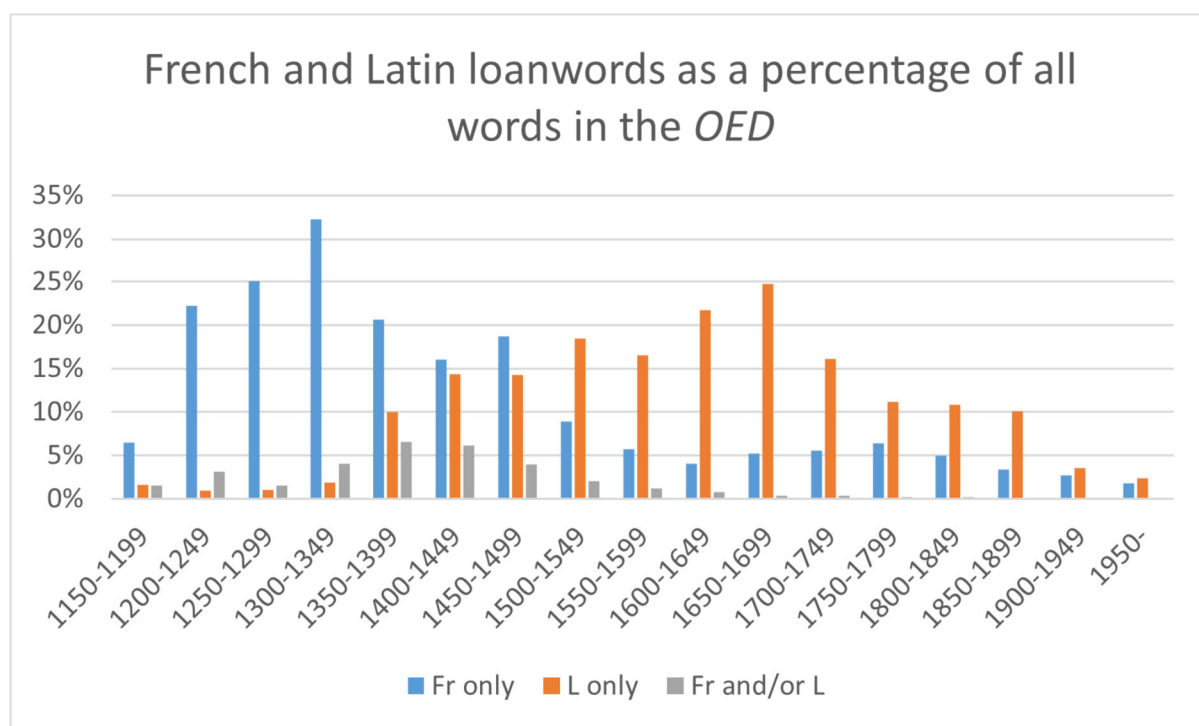


Figure 7. Loanwords from French, from Latin, and from French and/or Latin, as a proportion of all new words recorded in the *OED* for each 50-year subperiod.

century; there is little about the Middle English picture for Latin to immediately draw our eye. If now we look at the proportion of all new words in each subperiod that are borrowed from each language (Figure 7), just as was done for Japanese loans in section 5, we find a radically different picture: the ‘story’ for French is more focused on the Middle English period, with the 13th century now appearing much more prominent; the ‘story’ for Latin also looks rather different, with the 19th century no longer looking particularly significant, and the bars being rather more level from the late 14th century through to the late 18th century, and only really falling away in the 20th. Neither view is more ‘right’ than the other, they are complementary, but this second view reminds us rather more that we are looking at linguistic history as filtered through dictionary data.

If we are interested in the Middle English period, looking in this way at proportions of the total (dictionary) vocabulary also lets us more easily make a useful cross-check against the *Middle English Dictionary*, which has a somewhat fuller wordlist and more generous inclusion policy than the *OED*. If we do so, we find somewhat reassuringly that the *MED*’s etymologies and the *MED*’s wordlist give a broadly similar overall picture to the *OED*,¹¹ which is a helpful indication that the data reflect phenomena in the lexicon rather than purely facts about the dictionary.

7. Words showing input from both French and Latin

Besides loans from French and loans from Latin, both Figure 6 and Figure 7 also include a third category, words for which the *OED* records an origin from either French or Latin or (more typically) partly from each. In both charts, the total of such loans peaks just either side of 1400, at a key point in the shift of patterns of language use in various functions (such as written business records, or the transaction of legal matters). While they are outnumbered by loans identified as being from just one of these languages, these multiple-input loans are interesting to examine in a little detail, especially since, as will be shown below, they have had a very large impact on the core vocabulary of modern English, especially formal written English.

To return to an example touched on at the beginning of this article, for *general* the *OED* assumes that there is some input from Anglo-Norman (possibly also from continental Middle French) and some from Latin, giving rise to what is essentially a composite of multiple inputs. This analysis is supported by extensive documentation in the etymology section of the entry:

OED, *general* adj. and n., beginning of etymology section:
 < (i) Anglo-Norman *general*, *generall* and Old French, Middle French *general* (French *général*) (adjective) common or applicable to all or nearly all people (early 12th cent. in Anglo-Norman in *pechié general* original sin), concerned with or established for the whole of a territory or organization (end of the 12th cent.; chiefly with reference to deliberative bodies, e.g. in *assemblée generale*, *concile general*), (of a rule or principle) applicable to a variety of cases (1212 or earlier in Anglo-Norman), total, complete (second half of the 13th cent. or earlier in Anglo-Norman), not specifically limited or determined in application, relating or applicable to all members of a set (*c*1265), (of high-ranking officials) having overall responsibility (late 13th cent. or earlier in Anglo-Norman, frequently as postmodifier in titles, e.g. *coroner general*, *ministre general*, *vicaire general*, *tresorier general*), vague, imprecise (*c*1300 or earlier in Anglo-Norman), widespread, not localized (14th cent. or earlier in Anglo-Norman), (of a word, name, term) applicable to each of the individuals constituting a group or sharing a characteristic (second half of the 14th cent. or earlier), (of a

¹¹ See analysis in Durkin (2014: 255–63).

proclamation, charter, or other official document) covering all points (late 14th cent. or earlier in Anglo-Norman), (noun) general principle, universally valid fact or statement (*c*1265), chief administrative officer responsible for a particular department (late 14th cent., earliest in sense ‘chief public prosecutor’, short for *procureur general*; beginning of the 15th cent. denoting a high-ranking public officer in charge of finances and tax collection), highest-ranking officer of an army (1549; short for *capitaine general*: see CAPTAIN-GENERAL *n.*), head of a religious order (1554, originally with reference to the Jesuits), and its etymon (ii) classical Latin *generālis* shared by or common to the whole of a class or kind, generic, forming a group or class, of universal application, concerned with the nature or character of a thing, in post-classical Latin also common, of or for everybody, equal for all, (of the church) universal (4th cent.), widespread, not localized (from 11th cent. in British sources), (of an official) having overall responsibility (11th cent.; frequently from 13th cent. in British sources), (of an assembly or court) having overall jurisdiction (from 12th cent. in British sources), not specific, imprecise (from 13th cent. in British sources) < *gener-*, *genus* class, kind, race (see GENUS *n.*) + *-ālis* -AL suffix¹. With use as noun compare post-classical Latin *generalis* head of a religious order (from late 13th cent. in British sources; also in continental sources), *generale* general principle (13th cent. in a British source). The word is somewhat rare in classical Latin; in the later language, when *genus* and *species* (after the use of ancient Greek γένος and εἶδος in Aristotle) had become familiar as the technical terms for classes respectively of greater and less extension, *generalis* came to be often used in contrast to *specialis*; the antithetic use of the two words remains in the modern Romance and Germanic languages.

The spelling forms, meanings, and collocations cited here for the French and Latin words (with supporting dates of first attestation) are not intended to be comprehensive, but rather are selected on the basis of their correspondence to elements of the data for the word’s use in English presented in the rest of the dictionary entry, and their plausibility (on the grounds of correspondence and availability within the contact situation) as partial or full explanations for the adoption of the relevant forms, meanings, or collocations in English. (In some instances, additional forms and meanings may be listed, very sparingly, for each language if these are needed in order to provide an overview of the main lines of development in that language, particularly when that is relevant as antecedent to developments that are reflected in English.) Openness to this type of French and Latin polygenesis depends on a number of factors.¹² Firstly, there is our knowledge of the trilingual contact situation, and the shifts in functions of each language over time – where a crucial factor is the quality of the lexicographical and other resources through which we can observe this. Secondly, there are formal factors, especially the way that the patterns for naturalising Latin words in Middle English are heavily influenced by the shape such words typically have in French; in the case of *general*, this is rather undramatic, Latin *-ālis* being represented as *-al* in both French and English; more striking are patterns such as those shown by Latin nouns in *-tāt-*, *-tās* typically ending in *-ty* when borrowed into English, as a result of the pattern established by loans via French (as e.g. *bounty*, *plenty*, *poverty*, *subtlety*), the same pattern being applied even in those relatively rare cases (as e.g. *actuality*, *ponderosity*, *mediety*) where transmission even partly via French seems unlikely (on the basis of comparison of the available lexical data for each language). Thirdly, there is semantic complexity: Middle English loanwords of this type very often show borrowing of multiple meanings, very possibly resulting in many cases from coalescence of distinct borrowings in different areas of technical discourse, where the balance of use of French and Latin was often rather different.

¹²For a more detailed consideration of these factors see Durkin (2014: 236–53).

Among other common words of this type we find for instance *use, culture, family, nature, person*, also *question* or *accept*. Or, to take an alphabetical sample from the *OED*, we can find for instance among words beginning with *ab-*: *ability, abject, abjection, abjuration, abjure, ablative, ablution, abnegation, abortive, absent, absolution, absorb, abstinence*, and others.

Those examples, like the data in all figures in this article, are taken from the *OED* as it stands today, at a point in a comprehensive revision programme where rather over half of the dictionary's entries have been fully revised during the course of the past two decades, but other entries remain in unrevised form, which means for entries for older words that the etymological discussions still in some instances date back to the late 1800s or early 1900s, although we do now have an accelerated programme in place for revising more of these etymologies sooner. As the *OED*'s revision work continues, the proportion represented by the third columns in each graph, input from French and/or Latin, will rise further. The *OED*'s original editors already sometimes entertained the possibility of borrowing from both French and Latin, but more often plumped for one or the other. The *Middle English Dictionary*, in its very concise but insightful etymologies, much more often opts for input from both languages; in the many decades of its compilation (with publication in fascicles extending from 1952 to 2001), it was surely greatly helped in this task by the steadily improving dictionary resources for French and (medieval British) Latin, and especially by the first edition of the *Anglo-Norman Dictionary* (first fascicle published in 1977, second edition currently nearing completion), and by the *Dictionary of Medieval Latin from British Sources* (published in fascicles between 1975 and 2013) and its precursor wordlists (the *Medieval Latin Word-List from British and Irish Sources* in 1934, and the *Revised Medieval Latin Word-List* in 1965). The new edition of the *OED* is in the privileged position of being able to devote a fair amount of space and resource to etymological analysis and discussion, and is the heir to the enormous amount of work that has been done on Middle English, Anglo-Norman, continental Old and Middle French, and medieval Latin in the past century, especially as reflected by a rich panoply of dictionaries. Among advances being made today I might highlight the constant stream of valuable data for the history of English words coming out of the second edition of Aberystwyth's *Anglo-Norman Dictionary* and the Nancy-based *Dictionnaire du Moyen Français (DMF)*, as well as the peerless scholarly work of the (very sadly recently discontinued) *Dictionnaire étymologique de l'ancien français*, complementing older resources, foremost among them the *Französisches etymologisches Wörterbuch*. With regard to medieval Latin, the documentation of the still relatively recently completed *Dictionary of Medieval Latin from British Sources* continues to transform the historical picture for countless English words.

8. Finding a principled way of investigating a huge amount of lexical data

I will return again to these multiple-input words shortly. But before I do so I would like to look in a little more detail at the problems inherent in trying to navigate an enormous amount of lexical data like that presented in Figures 6 and 7 in a principled way. Clearly, the numbers of loanwords are huge, and sheer numbers rule out some approaches that are possible for the Japanese loanwords (as outlined in section 5). In the past, people have often approached this problem with long lists of loans, grouped into semantic fields, reflecting areas of influence. Often there has been an – entirely understandable – tendency to cherry-pick those words likely to be most familiar to readers of

Middle English literary texts and to modern English speakers – but this can create a badly distorting picture.¹³ It is a worry I constantly wrestle with in my own work.

I am delighted at this point to be able to point to some recent systematic work looking at large areas of the Middle English lexicon. For instance, several research projects led by Louise Sylvester and Richard Ingham have looked at the range of terms available in Middle English and Anglo-Norman in a wide variety of domains of everyday life, interrogating dictionary-derived data in a thesaurus approach. This work has yielded fascinating insights into the spread of loans across different technical fields, highlighting factors such as how often there is lexical replacement, or how often the loan appears to be the first item to occupy a particular semantic slot.¹⁴ Scandinavian-derived vocabulary has not been a focus of this article, but mention must be made of the wonderful work of the Gersum project led by Richard Dance and Sara Pons-Sanz, looking in a systematic way at all vocabulary for which Scandinavian origins are possible in a specific body of texts, namely selected poems of the Middle English alliterative tradition, and in the process producing a database of enormous value (both in its content and from a methodological perspective) for anyone interested in the etymologies of English words.¹⁵ Also, it should not go unmentioned that, in addition to the ongoing work on the revision of the *OED*, there is important collaboration by teams at Glasgow and at the *OED*, revising and extending the content of the (unique and still underused) *Historical Thesaurus of English*. I am truly delighted that work of this type is currently taking place, and is being funded, and I sincerely hope that funding models enable such work to be sustained, continued, and enriched. This is a field where we all learn from one another.

My own work has been focused especially on trying to find approaches to examining effects of French and Latin borrowing in the Middle English period on the core of the lexicon. It is particularly challenging to find ways of doing this that are principled and do not smack of cherry-picking. What constitutes the ‘core’ of the lexicon is itself variously identified. I want to touch on two possible approaches: through ‘basic’ vocabulary, and through word frequency.

9. Exploring impact on the core lexicon: basic vocabulary

Basic vocabulary is notoriously difficult to identify. A recent very impressive approach by an international team of collaborators produced the Leipzig-Jakarta 100-Meaning List of Basic Vocabulary, consisting of those meanings which were found most resistant to borrowing in analysis of the data from an extensive cross-linguistic research project.¹⁶ This gives us an interesting set of meanings in which we can examine the extent of lexical borrowing in English, especially with the aid of historical dictionaries, and especially if we have a historical thesaurus to help us.¹⁷ If we do this, we quickly find that we need to clarify what we mean by borrowing of words with basic meanings. For instance, ‘sweet’ is one of the meanings in the Leipzig-Jakarta list. If we look at this meaning in the *Historical Thesaurus of English* (sitting in the thesaurus structure at ‘the

¹³ For some exploration of this see Durkin (2020: 349–51).

¹⁴ Among outputs of this work see especially *The Bilingual Thesaurus of Everyday Life in Medieval England* (available at <https://thesaurus.ac.uk/bth/>), and Sylvester (2018), Ingham *et al.* (2019), Sylvester (2020), Sylvester *et al.* (2021).

¹⁵ For access to the database see <https://www.gersum.org/>. On the etymological methodology employed see also detailed treatment in Dance (2019).

¹⁶ See Haspelmath and Tadmor (2009), Tadmor *et al.* (2010), and also data at <https://wold.cld.org/>

¹⁷ For a detailed exploration see Durkin (2014: 41–5, 400–23) and, from a methodological perspective, Durkin (2016b).

world > physical sensation > taste and flavour > sweetness > [adjective]’), we find *sweet* (first recorded in the Old English period, as is the variant *soot*), and twenty-three other words, some of them borrowed from French or Latin, three of these, *douce*, *dulcet*, and *mellite*, being first recorded in the Middle English period.¹⁸ None of these words for a moment ever remotely rivals the native, non-borrowed word *sweet* as the usual, default realisation of this meaning. Taking a slightly closer look at one of these, *douce*, I think sheds useful light on a typical pattern of lexical borrowing from these source languages in this period.

In French and in Latin the usual words in the core basic vocabulary meaning ‘sweet in taste’ are French *doux* and its etymon Latin *dulcis*. The *OED* records a loanword *douce* from the mid 14th century, and another, *dulce*, from the beginning of the 16th century. There are also intermediate semi-classicised forms such as *doulce*, so we could imagine splitting the material into a different number of dictionary entries, and precisely how many distinct ‘words’ we are looking at here is not an entirely categorical matter. The *OED* entries for each word record a number of typical uses, divided into three separate senses for *douce* and two for *dulce* (as adjective). Both words are above the (very low) threshold of evidence to be labelled ‘rare’ in the *OED*, but neither is common at any point in their history. For *douce* there is a bare handful of uses in Middle English, and for both words combined in the full-text searchable portions of *Early English Books Online*, where we can do a bit more systematic number crunching, we find only around 80 uses in total, compared with over 200,000 examples of *sweet*. If we look further at the actual uses of *douce* and *dulce* in English, it is hard to find a single example that has reference straightforwardly to sweetness in taste: instead, what uses there are parallel technical uses in French or Latin (designating varieties of fruit, or non-saline water, or showing special uses in chemistry), or they are metaphorical; this is true even of a very small seam of use we find in Older Scots poetry.¹⁹

This is a very typical pattern, for which we could – with sufficient labour – find thousands of more or less close parallels: some technical uses, closely tied to the expansion of technical registers in English, and some (mostly metaphorical) uses as a stylistically marked term, closely tied to stylistic and rhetorical norms of late Middle English and Early Modern English literary composition. As such, it is useful in exploring some of what makes up a lot of the word count in my graphs in Figures 6 and 7.

Clearly though, this is not what we mean by impact on the ‘basic’ vocabulary of English. We are looking really for words that are the usual realisation of a particular basic meaning. The 100-meaning Leipzig-Jakarta list yields some firm Scandinavian-derived items: *root*, *wing*, *hit*, *leg*, *egg*, *give*, *skin*, *take* (although there are qualifying comments that should be made about several of these, not least the unusual and still somewhat disputed process by which in cases like *egg* or *give* a Scandinavian-derived form replaces a native Old English cognate). A few other items from the Leipzig-Jakarta list lead us to French borrowings close to the lexical core of modern English. One of the meanings in the list is ‘soil’: *soil* is a French loanword, but is it the usual way of expressing this meaning for all speakers even in standard English? If we are British perhaps, but if we are in the USA we may wonder whether we should be looking instead at the Scandinavian borrowing *dirt*, and there is further variation in other varieties of English, national, regional, or local. A different problem is presented by ‘to cry/weep’: we may be content that the French loan *to cry* is the usual word in this meaning except in marked literary or archaic use: but here we encounter

¹⁸Data cited here from the version linked to the *OED* on oed.com

¹⁹For more detailed examination of this example and its implications see Durkin & Allan (2016).

a different issue, that the meaning ‘to weep, shed tears’ of the verb *to cry* is almost certainly an innovation within English, rather than a meaning of the Anglo-Norman donor. Maybe at this point ‘basic’ vocabulary is starting to look a bit of a false start. But there is a more promising, and interesting, instance in the form of the verb *carry*.²⁰

After its first attestations in the 14th century, *carry* ‘competes’ only with native (i.e. non-borrowed) *bear* in this meaning, and, at least in some varieties, it gains ground very quickly: old-fashioned concordances show that already in the late 14th century in Chaucer’s writings it is outnumbered by *bear* only by a ratio of about two to one. (Chaucer in his role as Controller of the Customs of Hides, Skins, and Wools in the port of London was of course not unacquainted with the keeping of business records and its associated linguistic practice, as well as being immersed in the literary culture of several languages.) The *Historical Thesaurus of English* confirms that this is essentially a two-horse race, and we may I think legitimately speculate that there could be a functionalist motivation in the steady adoption of *carry*, in establishing an isomorphic one word–one core meaning relationship, in preference to the extensive polysemy of *bear*.²¹

Closer examination of the early history of *carry* raises some further interesting and relevant issues. Formal considerations show that it is a borrowing from Anglo-Norman *carrier*, not from continental French *charrier* – although even a brief look at the *Dictionary of Medieval Latin from British Sources* shows us that there is a fair quantity of evidence also for British medieval Latin *carriare*, an Anglo-Norman influenced form alongside the etymologically expected *carricare*. Etymologically these words are all derived from a word for a cart, and indeed this gives the semantic core of the word in Anglo-Norman: to transport things, especially goods in bulk, especially by cart. This is also the earliest recorded sense in English, and it is very frequent in late Middle English, in business records and related texts but also more widely; within a few decades we find also the broadened use which comes to take up a basic meaning semantic slot, which the *OED* defines as ‘to take or transport (a person or thing) from one place to another by means of physical effort, esp[ecially] using the hands or arms’; as already noted, there is evidence that this broadened sense rapidly became rather frequent, at least in some varieties. The use to do with carting gradually fades into the background; quite when it came generally to be understood as a specialised use of the broader sense, rather than vice versa, I doubt we can really recover. In my presentation of this story I have so far avoided one key wrinkle: there is no firm evidence that the broadened sense existed in Anglo-Norman, either in evidence in the *Anglo-Norman Dictionary* or in other source texts consulted by me and by others better qualified than me to do so. Is the broadening in meaning, from ‘to transport by cart’ to the broader meaning familiar today, therefore entirely an English phenomenon, and should the word therefore be placed in a similar category to *cry*? It is entirely possible that this is the case, but I would be very hesitant to ascribe too much significance to the lack of Anglo-Norman evidence: and in either case, intrusion of a loanword into the basic vocabulary almost simultaneously with its borrowing is a phenomenon worthy of our notice, whether the crucial semantic shift occurred in English or in Anglo-Norman.

²⁰ For a detailed account of *carry* see Durkin (2018). See also Durkin (2019) on some interesting issues raised by Middle English borrowing from Anglo-Norman in another area of basic vocabulary, kinship terms.

²¹ Explanations of lexical change, as of many other areas of linguistic change, are frequently controversial, a key problem being identifying why a particular change has affected a certain word and not others, or has occurred at a particular time rather than earlier or later, and a key difficulty being in establishing what is and is not strictly comparable, and which factors should be taken into account. Some important accounts (some sceptical) are found in Ullmann (1959), Samuels (1972), Lass (1980), Fortson (2003), Traugott & Dasher (2005), Geeraerts (2009).

The main thing I think that the example of *carry* illustrates is that basic vocabulary, in narrower definitions of that term, often behaves in somewhat surprising ways: this is an area where, above all others, every word has its history.

10. Exploring impact on the core lexicon: high-frequency words

As well as occurring in basic vocabulary lists, the word *carry* is also a high-frequency word in contemporary English, but in this company it is something of an atypical outlier, both in its tightness of focus around a core meaning, and in its spread very specifically from the language of trade. Much more typical of the high-frequency borrowings is the type I have already touched on illustrated by *general*, *person*, *culture*, *family*, or *nature*, and it is with a very brief consideration of *general* and its ilk that I will end this article.

Good quality frequency information is still something of a Holy Grail for the study of earlier periods in the history of the English lexicon; there are enormous challenges from issues such as the skewed nature of our surviving texts, the range of spelling variation shown by individual lexemes, and the messy full and partial overlap with the range of spelling variation shown by other lexemes. Even for contemporary English there are difficulties enough. However, if we take a published, carefully vetted list of the 1,000 most frequent words in a corpus of contemporary published written British English, we find that a large proportion of these words are loanwords from French or Latin, as shown in the bar chart in Figure 8.²² If we then look at when these words are first

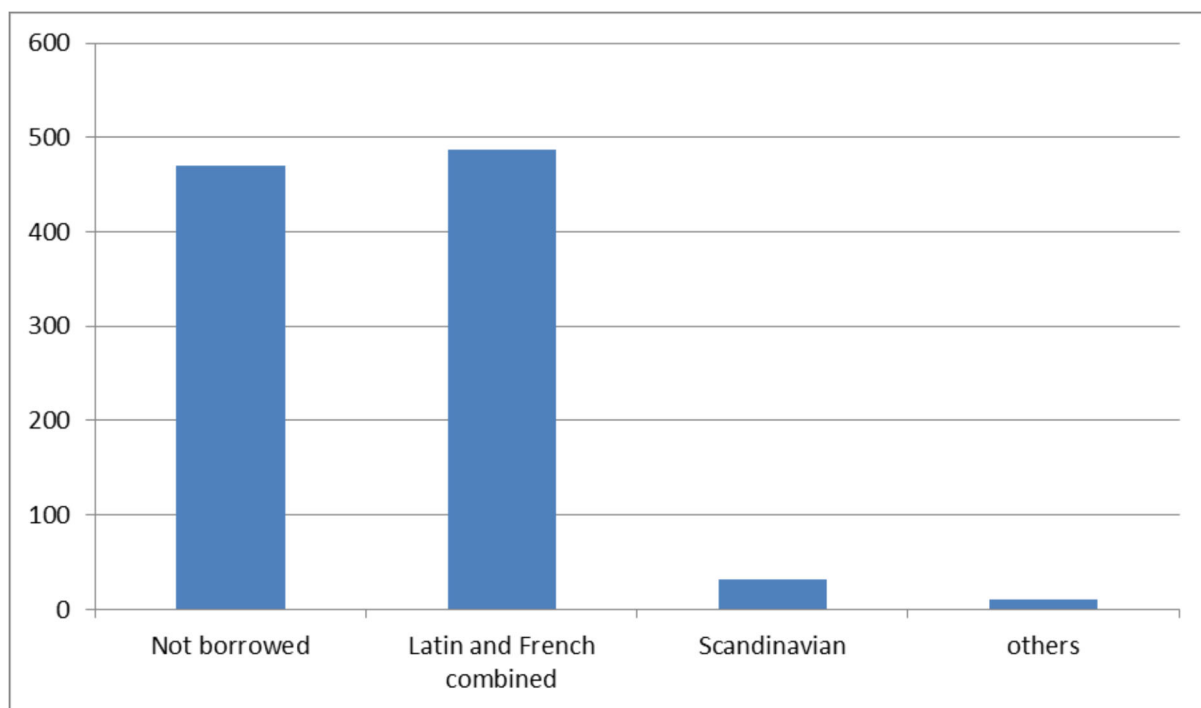


Figure 8. Etymological origins of the 1000 most frequent words in the British National Corpus (after Durkin 2014: 38).

²²For more details of the analysis employed here see Durkin (2014: 34–40), as also for brief comparison with the results found when a similar analysis is made of other contemporary corpora of written English. From a methodological perspective see also Durkin (2016b). For listing of all the words concerned, with etymological analysis, see appendix to Durkin (2020).

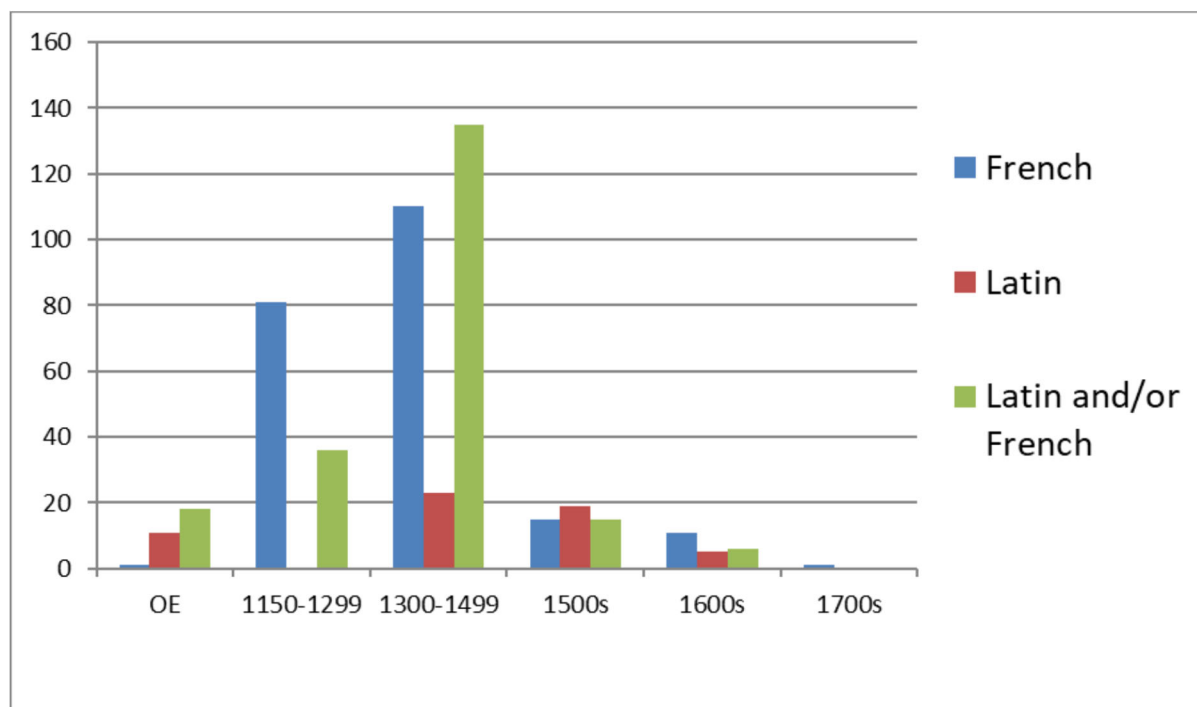


Figure 9. Loanwords from French, Latin, and French and/or Latin among the 1000 most frequent items in the *British National Corpus*, arranged chronologically by date of first attestation (after Durkin 2014: 39).

attested, the concentration in the Middle English period is overwhelming, and the proportion that can be assigned etymologies from both French and Latin is also very striking, as shown in Figure 9. Of course, this approach is teleological, looking at the origins of those words in most frequent use today, and most of these words appear to have shown a fairly steady increase in frequency of use, and of spread to more general contexts of use, over their first several centuries of use in English. However, few of them appear to have been exactly slow starters; one way we can see this, apart from simply number of attestations, is the considerable degree of polysemy shown from an early date. If we turn back to Figure 3, showing the main sense divisions of the adjective *general* as presented in the *OED*, it will be seen that five of the eight main senses are first attested before the end of the 14th century, and the remaining three before the end of the 16th century. What particularly interests me about this range of meanings is quite how many of them are paralleled earlier in French or Latin. In fact, for *general* all but one of the main senses are paralleled earlier in both French and Latin. (The exception is use of the type shown by *general practitioner* designating general, non-specialist competence, which is first attested in English from 1551, and which appears to have no clear antecedents in French or Latin.) We can see similar phenomena with *person* (where all of the major strands of meaning that occur before the late 19th century have clear antecedents in French or Latin), and with many other of the high-frequency French and Latin dual origin borrowings cited earlier.

In the *Historical Thesaurus of English* data, *general* sits alone, without near synonyms, in many of its major senses, and its main Middle English near or full synonyms in others are other Middle English borrowings, especially *common*.²³ The last of these is particularly interesting, since

²³ See for example (in the version linked to the *OED* in oed.com) ‘the world > relative properties > kind or sort > generality > [adjective]’, ‘the world > relative properties > kind or sort > generality > [adjective] > generally applicable’.

in its wider range of meanings, *common* frequently replaces the native term *i-mene*, cognate with German *gemein*, which perhaps hints to us how English may well have realised much of the conceptual area of both *general* and *common* without these medieval borrowings. But treating the near and full synonyms of *general* in the detail they deserve would take at least the whole of an article of this length, and this is an area where assessing the correspondence of two senses of different words is much less categorical than it is with for instance the use of *carry* or *bear* to mean ‘transport bodily from place to place especially with the hands’.

The concepts denoted by *general*, or *person*, *family*, *nature*, or dozens of other words of this type, are to a large extent European internationalisms, with West European cross-currents of influence affecting both the concepts and the words denoting them right up to the present day.²⁴ But what I think is most remarkable in terms of the history of the English lexicon is the appearance of words of this type in a rich range of meanings more or less simultaneously. What are the implications of the borrowing of a range of interrelated senses, by people who were multilingual, accustomed to making daily use of each of these languages? Do we have convergence of multiple separate borrowings, or does semantic borrowing, subsequent to borrowing of the word form, have a more significant role here? I think the answer must be that there is some of each, but I am not at all sure how we ascertain the proportions of each. Where multiple meanings are closely linked to one another conceptually, and where an individual is accustomed to employing the equivalent Anglo-Norman or Latin word in precisely the same contexts, it may well be that replication of the range of senses of the donor is to be expected as the normal result, except where there is a blocker, such as an established native term in frequent use in precisely the same meaning.

I am touching here on questions that bother me as I am ploughing my lexicographical furrow, and to which I suspect there must be insights to be found from work on a variety of different areas, such as: the history of concepts and ideas, European intellectual history, work on multilingualism, or on the mental lexicon, or indeed on the multilingual mental lexicon. But perhaps this (incomplete) list also indicates a methodological problem. Work on lexis often forces us to look at data that are clearly categorical and quantifiable alongside other factors that push into many other fields of expertise, far from the quantifiable, far from the familiar. In particular, interrogating lexical data, and attempting explanations of lexical change, can push into areas of linguistics that may delight some but be anathema to others, and it can also draw in perspectives from far beyond the world of linguistics. I suspect many of us are more comfortable if this involves words for parts of a plough or for the bulk importation of goods than if it touches on the history of complex concepts cutting across language boundaries, but I think there is interesting territory here for the intrepid lexicologist, and if the dictionary and its etymologies can act as a bridge, I will be happy to have done my bit.

²⁴ For very different, but perhaps ultimately complementary, approaches to investigating the relationships between words and concepts in their historical development, compare, on the one hand, Williams (1983) and the recent revisiting of the same approach in MacCabe *et al.* (2018), and, on the other, new methodological approaches developed by the research project *Linguistic DNA: Modelling concepts and semantic change* (<https://www.linguisticdna.org/>) and associated methodologies explored in Fitzmaurice *et al.* (2017).

11. Conclusions

I hope to have highlighted some of the sorts of issues that cannot help but matter to historical lexicographers today: the different sorts of multiple starts shown by many words, and questions of whether a particular word spreads rapidly into more general use; issues of who words are used by and how, especially in comparison with other words of the same or similar meaning; issues of shifts in these patterns over time, and how these can be identified and documented. These issues matter for lexicographers, and they also matter for linguists and historians of lexis using dictionaries: even if our interest is, for instance, in how widespread a particular phoneme or a particular word-forming process was in a language at a particular date, we ideally want a nuanced approach that tells us something about the prominence and frequency of the lexemes containing that element, rather than a bare inventory approach. The more deeply interested we are in the history of words and the history of the lexicon, the more detail we want available to us, even when that pushes us into areas that are necessarily interdisciplinary.

References

- The Anglo-Norman Dictionary* (1977-), 1st edn (1977–1992), ed. Stone, L.W., Reid, T.B.W. & Rothwell, W. (London, The Modern Humanities Research Association); *Anglo-Norman Dictionary: revised edition, AC; D-E* (2005), ed. Rothwell, W., Gregory, S. & Trotter, D., 2 vols (London, MHRA); (online publication) *Anglo-Norman Dictionary: revised edition* (2003-), at time of access published A-S, plus U and Z. www.anglo-norman.net
- Cannon, G. (1996), *The Japanese Contributions to the English Language: An Historical Dictionary*, associate ed. Warren, N. (Wiesbaden, Harrassowitz).
- Dance, R. (2019), *Words Derived from Old Norse in Sir Gawain and the Green Knight: An Etymological Survey*, 2 vols (Publications of the Philological Society, 50; Chichester, Wiley-Blackwell).
https://doi.org/10.1111/1467-968X.12148_02
- Dictionary of Medieval Latin from British Sources* (1975–2013), ed. Latham, R.E., Howlett, D. & Ashdowne, R. (Oxford, Oxford University Press, for the British Academy). For online publication see
<http://www.dmlbs.ox.ac.uk/web/online.html>
- Dictionary of Old English: A to I online* (2018), ed. Cameron, A., Amos, A.C. & Healey, A.D. (Toronto, Dictionary of Old English). <https://www.doe.utoronto.ca/pages/index.html>
- Dictionnaire étymologique de l'ancien français* (1971–2021), ed. Baldinger, K., Möhren, F. & Städtler, T., published in full form for letters D (in part) and E-K (Tübingen, Max Niemeyer). Dictionary and extensive bibliography available at www.deaf-page.de
- Dictionnaire de moyen français* (version 2020), ATILF – CNRS & Université de Lorraine. <http://www.atilf.fr/dmf>
- Durkin, P. (2006), 'Lexical splits and mergers: some difficult cases for the *OED*', in Caie, G.D., Hough, C. & Wotherspoon, I. (eds), *The Power of Words: Essays in Lexicography, Lexicology and Semantics in Honour of Christian J. Kay* (Amsterdam, Rodopi), 57–66. https://doi.org/10.1163/9789401203722_006
- Durkin, P. (2009), *The Oxford Guide to Etymology* (Oxford, Oxford University Press).
- Durkin, P. (2014), *Borrowed Words: A History of Loanwords in English* (Oxford, Oxford University Press).
<https://doi.org/10.1093/acprof:oso/9780199574995.001.0001>
- Durkin, P. (2016a), 'Etymology, Word History, and the Grouping and Division of Material in Historical Dictionaries', in Durkin, P. (ed.), *The Oxford Handbook of Lexicography* (Oxford, Oxford University Press), 236–52.
<https://doi.org/10.1093/oxfordhb/9780199691630.013.15>
- Durkin, P. (2016b), 'Using the OED and HTOED as tools in practical research: a test case examining the impact of loanwords on areas of the core lexicon', in Kytö, M. & Pahta, P. (eds), *The Cambridge Handbook of English Historical Linguistics* (Cambridge, Cambridge University Press), 390–406. <https://doi.org/10.1017/CBO9781139600231.024>
- Durkin, Philip (2018), 'Exploring the penetration of loanwords in the core vocabulary of Middle English: *carry* as a test case', *English Language and Linguistics*, 22(2): 265–82. <https://doi.org/10.1017/S1360674318000138>
- Durkin, P. (2019), 'New light on early Middle English borrowing from Anglo-Norman: Investigating kinship terms in *grand-*', *Anglia*, 137(2): 1–23. <https://doi.org/10.1515/ang-2019-0024>

- Durkin, P. (2020), 'The relationship of borrowing from French and Latin in the Middle English period with the development of the lexicon of Standard English: Some observations and a lot of questions', in Wright (2020) 343–64. <https://doi.org/10.1515/9783110687545-012>
- Durkin, P. & Allan, K. (2016), 'Borrowing and copy: A philological approach to Early Modern English lexicology', in Auer, A. González-Díaz, V., Hodson, J. & Sotirova, V. (eds), *Linguistics and Literary History: In honour of Sylvia Adamson* (Amsterdam, Benjamins), 71–86. <https://doi.org/10.1075/lal.25.05dur>
- Evans, T.M. (1997), *A Dictionary of Japanese Loanwords* (Westport CT, Greenwood Press).
- Farrington, A. (1991), *The English factory in Japan, 1613–1623*, 2 vols (London, The British Library).
- Farrington, A. (2008), 'Cocks, Richard (bap. 1565, d. 1624), merchant and East India Company servant', in *Oxford Dictionary of National Biography*, online version published 3 January 2008. Retrieved 28 June 2022, <https://doi.org/10.1093/ref:odnb/>
- Fitzmaurice, S., Robinson, J.A., Alexander, M., Hine, I.C., Mehl, S. & Dallachy, F. (2017), 'Linguistic DNA: Investigating Conceptual Change in Early Modern English Discourse', in Kytö, M., Smith, J.J. & Taavitsainen, I. (eds), *Interfacing Individuality and Collaboration in English Language Research World (Studia Neophilologica, 89)*, 21–38. <https://doi.org/10.1080/00393274.2017.1333891>
- Fortson, B.W. (2003), 'An approach to semantic change', in Joseph, B.D. & Janda, R.D. (eds), *The Handbook of Historical Linguistics* (Oxford, Blackwell), 648–66. <https://doi.org/10.1002/9780470756393.ch21>
- Französisches etymologisches Wörterbuch: Eine Darstellung des galloromanischen Sprachschatzes*, (1922–1978), 2nd edn in course of publication, founding ed. von Wartburg, W. (Basel, Zbinden).
- Geeraerts, D. (1997), *Diachronic Prototype Semantics: A Contribution to Historical Lexicology* (Oxford, Oxford University Press).
- Geeraerts, D. (2009), *Theories of Lexical Semantics* (Oxford, Oxford University Press). <https://doi.org/10.1093/acprof:oso/9780198700302.001.0001>
- Hane, M. (2000), *Japan: A Short History* (London, Oneworld).
- Haspelmath, M. & Tadmor, U. (eds) (2009), *Loanwords in the world's languages: a comparative handbook* (Berlin, De Gruyter Mouton). <https://doi.org/10.1515/9783110218442>
- Hayakawa, I. (2014), *A Historical Dictionary of Japanese Words Used in English*, rev. edn (Tokyo, Texnai).
- Historical Thesaurus of English*, 2nd edn (2020-), ed. Alexander, M. <https://ht.ac.uk/>. 1st edn published as *Historical Thesaurus of the Oxford English Dictionary* (2009), ed. Kay, C., Roberts, J., Samuels, M. & Wotherspoon, I. (Oxford, Oxford University Press).
- Ingham, R. (ed.) (2010), *The Anglo-Norman Language and its Contexts* (York, York Medieval Press).
- Ingham, R. (2012), *The Transmission of Anglo-Norman: Language History and Language Acquisition* (Amsterdam, Benjamins). <https://doi.org/10.1075/lfab.9>
- Ingham, R., Sylvester, L. & Marcus, I. (2019), 'Penetration of French-origin lexis in Middle English occupational domains', in Cennamo, M. & Fabrizio, C. (eds) *Historical Linguistics 2015: Selected papers from the 22nd International Conference on Historical Linguistics, Naples, 27–31 July 2015* (Amsterdam, John Benjamins), 460–78. <https://doi.org/10.1075/cilt.348.22ing>
- Ingham, R., Sylvester, L. & Marcus, I. (2020), 'Lone other-language items in later medieval texts', *Journal of Historical Sociolinguistics*, 7(2): 179–205. <https://doi.org/10.1515/jhsl-2019-0030>
- Institute of Historical Research (Shiryō Hensan-jō), University of Tokyo (eds) (1978–81), *Diary of Richard Cocks, 1615–1622*.
- Lass, R. (1980), *On Explaining Language Change* (Cambridge, Cambridge University Press).
- MacCabe, C., Yanacek, H. & The Keywords Project (2018), *Keywords for Today: A 21st Century Vocabulary* (Oxford, Oxford University Press).
- Medieval Latin Word-List from British and Irish Sources* (1934), ed. Baxter, J.H., Johnson, C. & Abrahams, P. (London, Oxford University Press.)
- Middle English Dictionary* (1952–2001), ed. Kurath, H., Kuhn, S. & Lewis, R.E. (Ann Arbor, University of Michigan Press). <http://quod.lib.umich.edu/m/med/>
- Morpurgo Davies, A. (1963), *Mycenaeae Graecitatis Lexicon* (Roma, Ateneo).
- The Oxford English Dictionary* (1884-), 1st edn (1884–1928), first issued as *A New English Dictionary on Historical Principles*, ed. Murray, J.A.H., Bradley, H., Craigie, W.A. & Onions, C.T.; *Supplement and Bibliography* (1933); *Supplement* (1972–1986), ed. Burchfield, R.W.; 2nd edn (1989), ed. Simpson, J.A. & Weiner, E.S.C.; 3rd edn (in progress), as part of *OED Online* (March 2000-), ed. Simpson, J.A. (2000–2013) and Proffitt, M. (2013-). www.oed.com
- Revised Medieval Latin Word-List from British and Irish Sources* (1965), ed. Latham, R.E. (London, Oxford University Press for the British Academy).

- Samuels, M.L. (1972), *Linguistic Evolution: With Special Reference to English* (Cambridge, Cambridge University Press).
<https://doi.org/10.1017/CBO9781139086707>
- Sylvester, L. (2018), 'Contact effects on the technical lexis of Middle English: A semantic hierarchic approach', *English Language and Linguistics*, 22(2): 249–64. <https://doi.org/10.1017/S1360674318000126>
- Sylvester, L. (2020), 'The role of multilingualism in the emergence of a technical register in the Middle English period', in Wright (2020) 365–79. <https://doi.org/10.1515/9783110687545-013>
- Sylvester, L., Tiddeman, M. & Ingham, R. (2021), 'Lexical borrowing in the Middle English period: a multi-domain analysis of semantic outcomes', *English Language and Linguistics*, 26(2): 237–61. <https://doi.org/10.1017/S1360674321000113>
- Tadmor, U., Haspelmath, M. & Taylor, B. (2010), 'Borrowability and the notion of basic vocabulary', *Diachronica*, 27: 226–46. <https://doi.org/10.1075/dia.27.2.04tad>
- Thompson, E.M. (1883), *Diary of Richard Cocks Cape-Merchant in the English Factory in Japan 1615–1622 with Correspondence*, 2 vols (London, Hakluyt Society).
- Traugott, E.C. & Dasher, R.B. (2005), *Regularity in Semantic Change* (Cambridge, Cambridge University Press).
- Trotter, D. (ed.) (2000), *Multilingualism in Later Medieval Britain* (Cambridge, D.S. Brewer).
- Ullmann, S. (1959), *Principles of Semantics*, 2nd edn (Oxford, Blackwell).
- Williams, R. (1983 [1st edn 1976]), *Keywords: A Vocabulary of Culture and Society* (London, Harper Collins).
- Wogan-Browne, J., with Collette, C., Kowaleski, M., Mooney, L., Putter, A. & Trotter, D. (eds) (2009), *Language and Culture in Medieval Britain: The French of England, c.1100-c.1500* (York, York Medieval Press).
- Wright, L. (ed.) (2020), *The Multilingual Origins of Standard English* (Berlin, Mouton de Gruyter).
<https://doi.org/10.1515/9783110687545>

To cite the article: Durkin, P. (2022), 'Tracking the history of words: changing perspectives, changing research', *Journal of the British Academy*, 10: 67–91.
<https://doi.org/10.5871/jba/010.067>

This article is licensed under a
 Creative Commons Attribution-NonCommercial-NoDerivs 4.0 Unported License.

Journal of the British Academy (ISSN 2052–7217) is published by
 The British Academy, 10–11 Carlton House Terrace, London, SW1Y 5AH
www.thebritishacademy.ac.uk

The Editors welcome 'Responses' to articles published in the Journal of the British Academy. Offers of 'Responses' should be sent to journal@thebritishacademy.ac.uk

