

## CHAPTER I

*Introduction**Sarah Ogilvie*

Dictionaries and lexicography (the art and craft of dictionary-making) have existed as long as humans have been writing. When one considers that the first dictionaries were carved into clay tablets by Sumerians over 4,000 years ago, then the first monolingual English dictionary, which appeared in 1604, could be considered positively ‘recent’. However, the four centuries since then present a fascinating story of evolution, innovation, devotion, plagiarism, and controversy.

Dictionaries are the kinds of books that are always ‘just there’. Alongside religious texts they have acquired, throughout history, a sense of sacredness and authority. There are reasons for this, and this volume traces how this became so. How did a single genre of text have the power to standardise the English language across time and region, to rival the Bible in notions of authority, and to challenge our understanding of objectivity, prescription, and description?

The story of English dictionaries is neither neat nor tidy, and certainly cannot be told as a straight linear progression from wordlist to spelling book to multi-volume dictionary. Rather, it is better understood as the story of the development of a whole ecosystem in which dictionaries of varied sizes and types co-existed for decades, or in some cases over a century, in multiple print runs and editions. Hence, the linear chronology which appears at the front of this volume necessarily presents a more cohesive picture than the reality. This is not to say that there were not observable trends in the content and coverage of dictionaries throughout the centuries. Many dictionaries of the sixteenth century advertised the coverage of ‘hard words’; those of the eighteenth century were prescriptive in their approach and prided themselves on ‘completeness’; whereas dictionaries in the nineteenth century pioneered the descriptivist approaches which are now the norm amongst major English dictionaries.

Economics has played an important part in the history of English dictionaries: the large, conspicuous, and expensive texts were never as

popular to readers as the smaller, cheaper texts. The story of the English dictionary since the early seventeenth century, as described by John Considine in this volume, has always primarily been a story of cheap, unpretentious, and fairly portable texts. This is perhaps not surprising to us in today's world in which millions of people act similarly by choosing to access a free online or mobile dictionary of mediocre quality, rather than pay a subscription to a comprehensive, scholarly one.

This volume attempts to tell the story of this thriving ecosystem. An international team of twenty-seven leading scholars and lexicographers presents chapters that are divided into three sections: first, an overview of essential issues pertaining to dictionary style and content; secondly, a fresh narrative of the development of English dictionaries throughout the centuries right up to current-day applications of technology, corpus linguistics, natural language processing, machine learning, and artificial intelligence; and thirdly, essays on the regional and global nature of English lexicography and its power to help standardise varieties of English and to define nations seeking independence from the British Empire.

These essays engage critically with the dictionaries they document, contextualising them historically and asking theoretical and methodological questions relating to the role that dictionaries have played, or do play, as tools of standardisation, prestige, power, education, literacy, and national identity. Readers who may want to delve more deeply into specific topics covered in the volume are encouraged to take advantage of the comprehensive list of dictionaries and the guide to further reading at the back of the volume.

It is important to note that the title 'English Dictionaries' refers to monolingual English dictionaries as distinct from bilingual dictionaries. Hence, although the first three essays of the second section address influences on monolingual dictionaries from the (earlier) bilingual tradition, it is the former that is the focus of this volume. This reflects a general distinction within the field in which lexicographic policies and practices of both traditions are generally kept separate. In essence, it would not be possible to do justice to either tradition in a single volume on both.

It was not long after the printing press was invented that the first dictionaries included English, but they were not monolingual. The *Promptorium parvulorum*, the first English-to-Latin dictionary, was published in 1499. It was another century before the first book generally regarded as the first stand-alone, monolingual English dictionary was published, *A Table Alphabeticall* (1604) by Robert Cawdrey. Cawdrey made use of wordlists that had appeared earlier in educational texts such as Richard

*Introduction*

3

Mulcaster's *The First Part of the Elementarie* (1582) and Edmund Coote's *The English Schoole-Maister* (1596).

It is intended that this *Cambridge Companion* will serve as a guide both to those who are studying this subject for the first time and to those who are already engaged in the study of dictionaries, especially those who might use them as barometers of culture and ways of gauging the social and cultural practices and biases of a particular period or region.

PART I

*Issues in English Lexicography*

## CHAPTER 2

*How a Word Gets into an English Dictionary**Kory Stamper***What is a Dictionary?**

Before we can discuss how a word gets entered into an English dictionary, we first need to examine what purpose a dictionary serves.

Many people believe that a dictionary acts as a sort of gatekeeper for proper English, listing only the words that are considered ‘correct’ or ‘worthwhile’. The corollary to that belief is the assumption that a word is not an ‘official’ word or a ‘real’ word until it has been entered into a professionally edited dictionary. Nothing could be further from the truth. The job of a dictionary is to record, as much as possible, the language as it is *actually used*, and not as people *think it should be used*. This may seem like a minor distinction, but as we will see, it’s an important one and forms the basis for how modern dictionaries chronicle the language.

**A Short History of the Early English Dictionaries**

People can be forgiven for assuming that dictionaries only record ‘proper English’ or ‘elegant English’, because for much of the history of English dictionaries, that is exactly what they did. The book considered to be the first English monolingual dictionary was written in 1604 by a schoolmaster named Robert Cawdrey and is called *A Table Alphabeticall*. (The full title of the work is much longer, in keeping with book-naming conventions of the 1600s, but modern scholars use the shortened form.) Cawdrey’s focus was on listing and defining what he called ‘hard usual’ words that he felt would improve plainspoken communication, though it’s clear in reading through his wordlist that his ideas of what constituted ‘plainspoken communication’ were a bit loftier than the modern person’s idea of simple speech. Cawdrey’s dictionary is rudimentary and contains none of the features that we associate with modern dictionaries: no comprehensive pronunciations or etymologies, no example sentences showing the word in

use, and no extensive or multi-sense definitions. He does mark some of his entries as coming from French or Greek where appropriate, and in the introduction, he explains how to look up a word without knowing how to spell it – proof that dictionary use was not universal even among the educated gentry. There are 2,543 defined words in his book, and they cover everything from general vocabulary like *abhorrent* and *confidence*, to legal terms (*misprision* and *rejoinder*), scientific terms (*meteor* and *comet*), and theological terms (*tabernacle* and *sanctification*). In writing *A Table Alphabeticall*, Cawdrey consulted previously published wordlists, bilingual dictionaries of English, and specialised glossaries – sometimes borrowing entries wholesale from those other sources. But his dictionary is not just a compilation of those works: it is written in its own style and is intended to be as systematic and helpful as possible. It is a remarkable lexicographical work, and soon other dictionaries like it followed (most notably, John Bullokar's 1616 *English Expositor* and Henry Cockeram's 1623 *English Dictionary*).

*A Table Alphabeticall* did, however, have one significant shortcoming that was shared by these other early seventeenth-century English dictionaries. It was written by one person, and though Cawdrey did his research, everything presented in *A Table Alphabeticall* is based ultimately on Cawdrey's thoughts and opinions about English. There are words that he chose not to enter because they were not 'the plainest & best kind of speech', including many foreign terms he deemed to be nothing but hot air used only to impress the hearer with the speaker's intelligence, as well as any terms that he considered to be 'low' (that is, common among the speech of the lower classes). Simple terms were also omitted. The target audience for Cawdrey's book and these other early English dictionaries was educated gentlepersons. Literacy was not as widespread in the early seventeenth century as it is now, and a comprehensive education was generally reserved for the wealthy or the well-connected. Early dictionary writers (or lexicographers, as they are properly called) had no reason to include simple words in their dictionaries, since their intended audience already knew these words. These early English dictionaries were not general surveys of English, nor were they intended to be. They were instead meant to polish the already-decent English of the educated.

This began to change in the late seventeenth and early eighteenth centuries, as literacy rates in England began to increase sharply. Scholars give several reasons for this increase: urbanisation, as many rural and less-educated people moved to London to seek work; a newfound class mobility, as the educated gentry lost influence and power to a rising – and comparatively less

### *How a Word Gets into an English Dictionary*

9

well-educated – merchant class; and the establishment of ‘dissenting academies’, which, unlike the traditional universities at Oxford and Cambridge, taught English as a subject and also taught other subjects in English (as opposed to the scholarly language of Latin). As literacy increased, so did the interest in (and market for) didactic books like grammars and dictionaries. Lexicographers, though, tended to continue to write their dictionaries primarily for the well-educated, though some of them grew savvy to the advantages of marketing a book well: Edward Phillips’s *A New World of English Words* (1658) primarily focused on hard words, as was the custom, but he called his book ‘a general dictionary’ in deference to the broader interest in dictionaries. (He also came under fire from another lexicographer, Thomas Blount, who accused him of heavily plagiarising Blount’s 1656 *Glossographica*.)

Elisha Coles, a schoolmaster, was the first lexicographer to make substantial moves towards creating a general dictionary. His 1676 *An English Dictionary* contains the usual lexical suspects – hard words and specialised vocabulary from mathematics, law, science, and theology – but also includes regional terms from all over England, as well as some ‘canting terms’, or criminal jargon of the time. It still did not include many simple words, nor did it include etymologies or quotations, but its extended scope was a hint of things to come.

The general dictionary came into its own in the early eighteenth century. John Kersey, a philologist and trained lexicographer, wrote one of the first truly general dictionaries of English, the 1702 *A New English Dictionary*. Kersey’s dictionary included hard words but focused primarily on common words that were in use, as his intended audience was ‘Young Persons, Tradesmen, Artificers, and the Female Sex’ – about as broad an audience as any writer could hope to get in the early eighteenth century. Nineteen years later, another philologist and lexicographer, Nathaniel Bailey, released his own general dictionary called *An Universal Etymological English Dictionary*, also written for a broad audience. (Despite its title, Bailey’s dictionary did not focus on etymology.) They were truly general dictionaries: Kersey was the first to include the words *cat* and *dog* in later editions of his *New English Dictionary*, and Bailey was the first to enter the definite article *the*. Bailey’s dictionary, especially, was very popular: its last edition, the thirtieth, was printed in 1802.

### **Samuel Johnson and the Modern English Dictionary**

Despite the proliferation of dictionaries, there was nonetheless an expressed desire among writers and grammarians for *The Dictionary*, not just

a *dictionary* – an authoritative reference work. Many of these same grammarians and writers were in a state of panic about the English language: it was too profligate, borrowed from other languages too easily, and allowed for the creation of words without regard to elegance of style. *The Dictionary*, then, would not just chronicle the language, but help set borders around it and point its users towards a more elegant and lasting manner of expression. In the 1740s, a group of London booksellers banded together to commission the creation of this work and tapped Samuel Johnson to create it.

Samuel Johnson was not the obvious choice for this project. He did not have a university degree; he was not well connected among the aristocracy; he was not a teacher, schoolmaster, or well-regarded scholar. What he was, however, was available and interested. In 1746, he agreed to write what was hoped to be the first authoritative dictionary of English.

Johnson did not, like lexicographers before him, rely solely on his own sense of the language to come up with his list of headwords (that is, the list of words that would appear as main entries in his dictionaries) or his definitions. Instead, he systematically read hundreds of sources as preparatory work for his dictionary – everything from Shakespeare and Milton to legal texts, educational treatises, geology texts, and poems – and as he did, he watched for interesting words or passages that he wanted to use in his dictionary. He underlined the word to be quoted and defined, then marked the surrounding context of that word, and finally put the initial letters of that word in the margin. When Johnson was finished reading a volume, his assistants would go through it and copy all the marked passages onto individual slips of paper, which were then organised alphabetically according to the underlined word. In the end, Johnson had hundreds of thousands of these slips, which modern lexicographers call ‘citations’, and he used them as the source of the headwords, definitions, and illustrative quotations in his dictionary.

Johnson’s *Dictionary of the English Language*, published in 1755, was remarkable in its scope and is still considered to be a masterwork of lexicography. The definitions were extensive; unlike in earlier dictionaries, Johnson did not simply give each entry one or two definitions that he felt were most important, least understood by the educated public, or most common in use. Johnson based every definition in his dictionary on the contextual meaning that a word was given in each of the citations he had collected, which meant his dictionary often gave numerous meanings for each headword in an attempt to catalogue the totality of the word’s use. For instance, the noun *light* has two separate meanings in Kersey’s dictionary and just one in Bailey’s, but Johnson gives fourteen; the adjective *general* has two discrete meanings in Bailey and one in Kersey, but ten in Johnson.



### *How a Word Gets into an English Dictionary*

11

Johnson's dictionary was also the first to make extensive use of illustrative quotations. These quotations help orient a word's definition, which is an abstracted statement of meaning without context, within the word's actual use in print, which is its native habitat. These quotations also inadvertently defended Johnson against any detractors who complained that his dictionary included terms considered inelegant or low. In the preface to his dictionary, Johnson writes, 'Some of the examples have been taken from writers who were never mentioned as masters of elegance or models of stile; but words must be sought where they are used'. *The Dictionary of the English Language* immediately established the basic defining method and set the scholarly standard for all dictionaries that followed.

By the nineteenth century, dictionaries were a growth industry, not just in Britain, but in America as well. Widespread literacy and general education created a booming market for dictionaries of all kinds, and by the mid-1800s, dictionary publishers moved away from a single-author model and began to recruit and retain editorial staffs to keep up with the popular demand for more extensive and comprehensive dictionaries. They also established reading programmes to make sure that the raw material they were collecting for their dictionaries covered as much thematic, geographical, and sometimes chronological territory as possible. It was during this time that the modern template for how a word gets into a dictionary was solidified.

### **The Nuts and Bolts of the Defining Process**

Entering a word into an English dictionary generally consists of two processes: the collection of written evidence, and the analysis of that evidence. We'll look at each process separately.

#### *Gathering Evidence of Use*

There are several ways to gather the written evidence of a word's use. Until the late-twentieth century, the primary way this was done was through a process that was based on the preparatory work Samuel Johnson did for his dictionary. This process is often called 'reading and marking'.

Each dictionary company compiles a list of written, edited prose sources that they have their editors or trained readers go through in order to find new words or new uses of existing words. These lists can be as targeted or as comprehensive as necessary, according to the type of dictionary being written. For a general English dictionary, lexicographers try to formally read as much edited, published work as possible: books from all genres,

magazines, newspapers, speciality trade journals, monographs, pamphlets, and so on. For a more specialised dictionary, the reading list may focus on one particular type of source (such as medical texts) or one particular era (like eighteenth-century science works). The list is then assigned to readers or editors, and they will read through the sources, looking for new words, new uses of old words, or sometimes just words that catch their eye. The word is then underlined and the context around that word bracketed, and then the page is somehow marked for the assistants who will go through each source and copy each bracketed citation into a database (or, formerly, on 3x5 inch index cards, which were then filed alphabetically). Just as it was with Johnson's dictionary, these citations comprise the raw materials used to create a dictionary.

In addition to formal reading and marking, lexicographers often read and mark additional sources that they find on their own: the marketing copy on TV dinner boxes, menus, phone books, playbills, catalogues. If it has print on it and is widely distributed, there's a very good chance that a lexicographer has read it for citations. Some dictionary publishers also solicit or accept citations from the general public. The goal is to collect the biggest possible cross-section of the language. Lexicographers cannot hope to record everything in print – they are lucky if they can record even a measurable fraction of what makes it into print. But the goal is to at least have a good representative sampling of the language to draw upon when writing dictionaries.

This axiom applies to the geographic reach and types of books, magazines, trade journals, and informal materials read as well. Most major regional newspapers from around the English-speaking world are often included in a comprehensive reading programme and marked for regional variations. A balanced reading programme is not snobbishly academic: while technical fields like medicine and computer science do add to the language, lexicographers recognise that non-academic fields such as cooking and pop culture give us just as much – if not more – new language as academia. A representative cross-section of English includes everything from legal texts to romance novels, from *Today's Chemist at Work* to *Thrasher Magazine*, from California to Australia and back again.

### *The Internet and the Lexicographer*

This process has changed slightly with the rise of the Internet, of course. Just as most people now get their news online, so, too, does the dictionary: many lexicographers now comb through news sites, popular blogs, and well-known public social media feeds as they look for new words.