

# An introduction to the textbook

This book is about the sources of English words, about their etymology and history, about their sound-structure, and about some formal properties of English word-formation rules. It is important to realize, however, that it is not about *all* possible origins, it is not about *all* the ways in which English has introduced new words into the language, but rather it is primarily about a particular subset, that portion of the vocabulary which is borrowed from the classical languages (Latin and Greek) either directly, or indirectly through French.

This (very large) portion of our vocabulary is a familiar subject. Greek and Latin roots in the English language have been studied and have been part of the core educational curriculum at least since the Renaissance. Departments of Classical Languages traditionally offer courses under titles like “Classical Roots in English,” and in the past a decent education necessarily included a full program in the classics. In the twenty-first century, however, it is extremely rare for students entering college to have a clear idea even of what Latin is – some ancient language, perhaps –, or whether English is derived from it or not, and even what it means for a language to be “derived,” in any sense, from another. The word *cognate* is not only generally unknown to undergraduate students, it often remains conceptually obscure, because it is simply not one of the topics we grow up with these days.

We take the view that people cannot call themselves “educated” who do not have a minimal acquaintance with the history and structure of the words in their own language. It doesn’t take much: if you are a word-lover and use a dictionary a lot, you will probably find much that is familiar in this book. But people don’t usually use a dictionary to do more than settle an argument about spelling, pronunciation, or origin. It should be used for *much* more. Learning to appreciate those additional uses is one of the benefits we hope to provide to our readers.

Another benefit is learning to appreciate relationships between words that even the best dictionaries don’t always make clear. These relationships are part of what linguists call *morphology*. Morphology, which addresses the patterns of word-formation and change, is not a very “regular” part of language. The forms that words take is largely the legacy of history, whereas both the sound structure of language and its syntactic organization are probably innate, for the most part.

A question which everyone wonders about, and often asks of instructors, is “How many words does English have?” And even more commonly, “How many words does the typical educated person know, approximately?” There is no

verifiable answer to this question. We can tell you how many headwords a given dictionary has (or claims to have), or how many words Shakespeare used in his plays (because it is a closed corpus of texts, and we can count the number of different words – about 21,000 if you count *play, plays, playing, played* as a single word, and all similar cases, almost 30,000 if you don't). A very generous estimate of the vocabulary of a really well-educated adult is that it may reach up to 100,000 words, but this is a wildly unverifiable estimate. We can quote the *Oxford English Dictionary's* statistics on the number of main entries: 231,100 (<http://dictionary.oed.com/about/facts.html>), but that figure is not particularly meaningful because it includes ancient as well as modern words, and most of the ancient words are unknown to us. They are obsolete and of antiquarian interest only.

One thing is certain: well over 80 percent of the total vocabulary of English is borrowed. The more we know about the sources and processes of linguistic borrowing, the better our chances of coping with technical vocabulary and educated usage in general.

**An introduction to dictionaries**

To use this book, one must have easy access to a good dictionary. Let us therefore start by asking, what makes a dictionary good? This prompts further questions: how did such books come into being? How do we get the most out of them?

All the major dictionaries of English are available in electronic form. The advantage of an electronic version is that it can be updated frequently, it allows easy cross-referencing, it allows complex searches on dates, etymology, author(s) of citations, it allows audio links to the pronunciation, etc. Beware: the electronic Thesaurus included in word-processing programs does not and cannot stand in for a dictionary which contains a full array of information relevant to the origins, the history, and the forms of words. Such electronic applications are primarily for spelling and for finding synonyms.

English dictionaries are a recent invention. Curiously, in Britain they started as an accidental by-product of ignorance. Anglo-Saxon monks often did not know Latin very well. Most of the texts they were copying were written in Latin, so they jogged their memories as any elementary language student might do today by writing translations (“glosses”) between the lines. By the beginning of the eighth century the first lists of Latin-to-English glosses appeared, organized either by topic or alphabetically. Such lists are known as *glossaries*. The earliest known glossary arranged in alphabetical sequence not by the Latin, but by the English word, was produced in the thirteenth century. These were, however, only bilingual aids. It took another three centuries before someone realized there might be money to be made by publishing lists of English hard words with explanations of their meanings also in English. The first such publication appeared at the beginning of

the seventeenth century. The first moderately complete English dictionary was another 150 years later: *A Dictionary of the English Language* in two volumes, by Samuel Johnson, published in 1755. Modern lexicography is therefore less than 300 years old.

The making of dictionaries has been a major scholarly occupation and a flourishing business enterprise for publishers in the last two centuries. In the twenty-first century electronic versions of the printed dictionaries have become an essential teaching and research tool. Since availability and access varies, here we will introduce three sources without which the exercises in the online Workbook accompanying this textbook cannot be completed.

The ***Oxford English Dictionary (OED)*** (<http://dictionary.oed.com/>) will be your primary reference aid. The *OED* online is unsurpassed in the richness and flexibility of use of its database. The wealth of information and search options available on the website can, however, be intimidating. For the college or university student unfamiliar with the organization of the entries, the website provides a helpful guide; we recommend also the very useful article entitled “Working with the online version of the Oxford English Dictionary” by T. T. L. Davidson (<http://dictionary.oed.com/learning/university/worksheet.html>). The article offers easy-to-follow step-by-step advice on searches with the following sample aims (we cite):

- Look up the meanings of a word and how they have developed.
- Look up when words and meanings were first used.
- Find out the etymological source of a word.
- Find parts of words (e.g. the uses of prefixes such as *pre-*, *arch-*, or *peri-*, or suffixes such as *-ology*, *-nik*, or *-ate*) and generally to investigate word-formation in English.
- Find out how far the *OED* systematically records relationships between words such as synonyms (e.g. the relevant senses of *reel* and *spool*).
- Secure some support for the ideas that speakers have about likely collocations of words (e.g. that we can say *notable collector* and *distinguished collector*, but only *notable frequency* and not *\*distinguished frequency*).
- Examine the details of processes in English such as the emergence of “zero-derivation” forms such as when *paper* started being used as a verb, presumably having been around for some time as a noun.
- Dig out meanings of words which poets might have been using when the poem was written but which have since disappeared.
- Try to recreate the vocabulary “fields” of political and social discourses of the past.
- Find out what contribution particular writers have made to the development of new words and meanings.

*The American Heritage Dictionary of the English Language (AHD)* is another important source of information related to the contents of this book.

Its fourth edition (Boston: Houghton Mifflin, 2000) is available at [www.bartleby.com/61/](http://www.bartleby.com/61/). Although it does not have dates or citations, the *AHD* has many additional valuable features. The main entries are accompanied by sound-files, so the student can hear the word pronounced. Many items have links to color illustrations of the word, so one can actually *see* what a *pricket*, a *quoin*, or a *rabbet* looks like. Entries can be accompanied by notes on regional usage, current recommended usage, synonyms, and word-histories. Most importantly for the student interested in the earliest word-connections, the *AHD* provides links to Indo-European roots for words whose etymologies can be traced all the way back to reconstructed proto-forms. There is also an online article on Indo-European and the Indo-Europeans by Calvert Watkins, Indo-European sound correspondences, and a searchable appendix of Indo-European roots. Since the *OED* policy is not to provide reconstructed Indo-European forms, the *AHD* will be indispensable to the user of this book for mastering the material in Chapter 8.

*Webster's Third New International Dictionary, Unabridged*, is found at <http://unabridged.merriam-webster.com/>. The (current) electronic version carries a 2002 copyright. It is a fully searchable version which contains over 450,000 vocabulary definitions, each with etymological and phonological information, and very extensive usage examples. The *Webster's Third* provides some dates of the appearance of new senses; it allows searches of authors cited, and in addition to the usual types of searches it has the potential to search for rhymes, homophones, cryptograms.

We have only mentioned three online resources here – we believe that between them anyone using the book will be sufficiently well served. They have “aged” well, if a decade is enough for that qualification. This is not to say that there are no other good resources on the web – the electronic-based records of language are constantly growing, but we have no doubt that the *OED*, the *AHD*, and *Webster's Third* will continue to provide most reliable scholarly lexicographical information.

# 1 Word-origins

The two general themes of this book are the origins and the structure of English words. Our word-stock is huge. It is useful to divide it up between words that belong to the common language that everybody knows from an early age and words that are learned in the course of our education. The former, the core vocabulary, is nearly the same for everyone. The latter, the learned vocabulary, is peripheral and certainly not shared by everyone. The core vocabulary is not an area where we need special instruction – the core vocabulary is acquired at a pre-educational stage. Our learned vocabulary is a different matter. It varies greatly in size and composition from one individual to another, depending on education and fields of specialization. No single individual ever controls more than a fraction of the learned vocabulary. Often the extent of one’s vocabulary becomes a measure of intellect. Knowledge about the history and structure of our words – both the core and the learned vocabulary – is a valuable asset.

The vocabulary of English is not an unchanging list of words. New words enter the language every day, words acquire or lose meanings, and words cease to be used. The online *Oxford English Dictionary (OED)* is updated quarterly with at least 1,000 new and revised entries; this is a fair measure of how dynamic our vocabulary is. The two sources of new words are borrowing and word-creation. In fields of higher learning, like the life sciences, physical sciences, medicine, law, the fine arts, and the social sciences, English has usually borrowed words from other languages to get new words to cover new concepts or new material or abstract phenomena. Words referring to notions and objects specific to other cultures are often borrowed wholesale. We may borrow a word as a whole, or just its central parts (the roots). We have borrowed mainly from Latin, Greek, and French. The discussion of borrowing will be a central theme in later chapters; in this chapter, we focus on the patterns of vocabulary innovation – the creation of new words – that occur within English. Before we identify the many ways of vocabulary enrichment, however, we want to address briefly the whole notion of lexical heritage.

## 1 Lexical heritage

Our lexical heritage consists of all those words which we as speakers receive from our predecessors when we acquire our native language. These

inherited words in turn have originated, one generation, or many, back from the present, through borrowing or any of the types of word creativity listed below. The inherited words divide into **core** vs. **periphery**. For the most part, the core vocabulary has been part of English for many centuries, passed down with minor changes. Much of it is shared with closely related languages like Dutch, the Scandinavian languages, and the classical languages Latin and Greek. The notion of what it means to be a closely related language is the topic of Chapter 2. For the moment the notion of relationship can be understood in a pre-scientific sense, as in “family relationship.”

The core vocabulary includes all of the common prepositions (*by, for, to, on, in, of, with, over, among*, etc.). They are learned well before the age of five, and so are conjunctions like *and, but, or*. They are an essential part of the glue that holds sentences together. Other core words are the auxiliary and linking verbs (*be, is, was, were, are, am, have, can, could, may, might, will, would, shall, should, must, ought to*) and many common verbs having to do with perception and the senses (*feel, think, dream, touch, hear, see*) and common names of everyday essentials and properties, body parts, kinship, colors (*food, drink, water, bread, mouth, eyes, hand, foot, leg, mother, father, brother, sister, black, white, green*). If we look just at the 1,000 most common words of English, over 800 of them are of this type. Many of them can be traced back as far as language history allows us to go – at most about 6,000 years before the present time. Among the top ten most frequent nouns identified by researchers at Oxford University Press: *time, person, year, way, day, thing, man, world, life, hand*,<sup>1</sup> only *person* is a historically borrowed word, and it has been in the language for over seven centuries. All five most frequent verbs in the language: *be, know, say, make, get* have existed in English as long as the language itself, and are shared with genetically related languages. Some of the other words have popped up in the language during more recent times – the last two or three millennia – and though more recent, in many instances their origins remain mysterious. For instance, the base of the word *penny*, which has been around for as long as English has existed, since early Old English, is completely unknown, *brunt* as in ‘to take the brunt of the attack,’ has been in the language since 1325, but it remains of unknown origin; *blear(y)*, from the fourteenth century, origin also unknown; *duds*, as in ‘to wear fancy duds,’ from the middle of the fifteenth century, also unknown. Closer to our times, *copacetic, posh* are from the beginning of the twentieth century; their etymology is unknown.<sup>2</sup> *Snazzy* is from the first Roosevelt administration starting in 1933, but no one knows its ancestry. Even words first recorded within our own lifetimes: *wazoo* (1961), *glitch* (1962), *ditsy* (1978), *full monty* (1985), *wazzock* (1984) have unknown or highly speculative etymologies.

<sup>1</sup> The information is found on the BBC News website: [http://news.bbc.co.uk/2/hi/uk\\_news/5104778.stm](http://news.bbc.co.uk/2/hi/uk_news/5104778.stm).  
<sup>2</sup> *Posh*, it should be pointed out, has been mistakenly claimed to be a blend of ‘Port Out, Starboard Home,’ the wealthy way to travel on a Mediterranean cruise, to avoid having the sun in your porthole. The *Oxford English Dictionary* considers this etymology to be without foundation.



In addition to its core vocabulary, English has a rich supply of learned words (*learned*, in this meaning, is pronounced as two syllables). The learned vocabulary is different from the core vocabulary in that most of it is acquired through literacy and education. It tends to be associated with technical knowledge and professional skills, though there is also a large part of it which is associated with humanistic education, with literature and the arts. Vocabulary enrichment in all of those areas has drawn heavily on borrowed words and roots. We return to the notions of core and periphery in Chapter 3, Section 3, when we will know more about the entry of borrowed words into English. Indeed, most of the rest of this book is devoted to finding out when and how the learned vocabulary came into English. But first we need to examine the sources of other words, words that are not part of the inherited core vocabulary and that are not directly drawn from the classical languages. These are words which are created by inventive minds, and they follow a small number of patterns. The next sections in this chapter address this topic: where do our new words originate? How do they get created and integrated into the language? The two general headings under which the specific types of new words can be grouped are “Regular word-formation” and “New word creation.”

2 Regular word-formation

2.1 Derivation by affixation

Unlike the coining of new words which are not immediately transparent to the native speaker, discussed below in Section 3, deriving new words by affixation is usually completely transparent. This way of generating new lexical items is “regular” in the sense that it relies on pre-existing and recyclable language units that are familiar to any native or fairly proficient speaker of the language. The processes under the umbrella of regular word-formation, jointly known also as **derivational morphology** are, in many instances, so obvious that significant numbers of derivations are not even treated by dictionaries as separate entries. Since most of this book is about the complexities of derivational morphology, we do not want to anticipate details here. Roughly, by way of introduction, derivation by affixation consists in making up new words by adding affixes, or endings, to more basic forms of the word. Mostly these derivations require no special definition or explanation because they follow regular rules. For example, from the *Chambers Dictionary*, under the headword *active*, we find these derived words: *activate*, *activation*, *actively*, *activeness*, *activity*, *activism*, *activist*. Four of them are given no further explanation at all, two of them are given only the very briefest explanation because the meaning has become slightly specialized, and one – *activate* – is treated at more length because it has a technical sense that requires explanation. The question is, when is a derived form merely that, predictable and comprehensible by general rules of the language, and when does the derived form





are nouns, in these cases; but there are also adjective/adverb–verb pairs like *near*, *idle*, *clear*, *smooth*, *obscure*, and many more). Since this process allows one word to acquire a new function, we can also think of it as *functional extension*, or *functional shift*.

The process of conversion is extremely productive today: we can *chair a meeting*, *mask our intentions*, *air our opinions*, *panel the walls*, *stage a protest*, *weather the storm*, *storm the gates*, *e-mail the students*, *floor our enemies*, *polish the car*, *fish in troubled waters*, and so on. Conversions that have been around long enough are normally shown with a single entry in many dictionaries, with the identification n., a., v., meaning that the form occurs as noun, adjective, and verb all three. Recent, or surprising, conversions often get separate entries in the dictionaries. Like other word-formation patterns, conversion may produce short-lived or nonce words: *to history*, *to beetle* (both found in Shakespeare), *to conversation*, *to dead*, *to ditty*, *to maid*, *to nighthawk*, *to perhaps*, *to proverb*, *to wool*, *to word*, and even *to Devonshire* (1607) ‘to clear or improve land by paring off turf, stubble, weeds, etc., burning them, and spreading the ashes on the land’ (*OED*).

A relatively recent pattern of derivation without suffixation, which preserves the sounds, but changes the stress, of verbs vs. nouns and adjectives, is typified by pairs such as *convict*, *present*, *refuse*, *torment*, in which the last syllable is stressed in the verbs, but the stress shifts to the initial syllable in the other forms. Such words are known as *diatonic words*. Only two words showed such stress shifting before 1570, *rebel* and *record*, but the number of words, mostly borrowed, that have undergone such shifting in the last three centuries has been growing steadily.

2.3      **Compounding**

This is a very large, and therefore very important, source of new words. To produce new words by compounding, what we do is put together two words in a perfectly transparent way, and then various changes take place which may cause the compound to lose its transparency. A clear example from very early English is the word *Lord*, which is an opaque form of *loaf* ‘bread’ (you can see the *l* and the *o* still), and *ward* ‘guardian’ (you can see the *rd* still). A less extreme example, without the phonetic complication, is a word like *hoe-down* ‘noisy dance associated with harvests and weddings in the old South and West.’ The *OED* gives it as the equivalent of an earlier sense of *breakdown*, now obsolete in the relevant meaning. In neither case can one infer the meaning from knowing the meanings of the constituent parts. It is therefore an opaque compound. Other examples of the “Lord” type which were once compounds and are now recognizable only as fully assimilated single words include *woman* from *wife* + *mon* ‘female’ + ‘person,’ *good-bye* from *God be with you*, *holiday* from *holy day*, *bonfire* from *bone fire*, *hussy* from *house wife*, *nothing* from *no thing*.

A full description of compounds is far beyond our scope, but because it is the largest and most important source of new words in the English vocabulary, outside

We begin by distinguishing between *syntactic compounds* and *lexical compounds*.<sup>3</sup> One can always figure out what a syntactic compound means. Such compounds are formed by regular rules of grammar, like sentences, and they are not, therefore, listed in a dictionary. So if someone were to say,

We know, just from the rules of grammar, that they could also say,

*Quartet playing* is therefore a syntactic compound. Other transparent syntactic compounds are *birthplace* (a place of birth), *bookkeeper* (someone keeps the books in order), *washing machine* (we wash things with the machine), *moonlight* (light provided by the moon), *sunrise* (the rising of the sun), *policymaker* (someone who devises policies). In fact the majority of compounds we use on a daily basis are the transparent syntactic ones.

*They iced the cream.*

<sup>3</sup> Students interested in finding out more about this topic should consult a truly great piece of scholarship, Hans Marchand, *The Categories and Types of Present-Day English Word-Formation*, 2nd edn (Munich: Beck, 1969).