

# The English Language in the Dictionary

---

IN THE OFFICES WHERE WEBSTER'S NINTH NEW COLLEGIATE Dictionary was edited, several thousand letters are received each year. The topics they articulate are enormously varied. Some letters merely ask for a particular bit of information about the English language that has been sought but not found in the dictionary. A few others are in hot pursuit of a special interest. Still others, their writers having come to think of the dictionary as an all-purpose reference book, ask questions about many other subjects besides words. A surprising number of correspondents, however, express considerable curiosity about how their dictionary – that formidably long and closely printed work with its many special abbreviations, symbols, and devices and its multitude of uses – came to be just the book it is.

They may ask quite directly such questions as how words make it into the dictionary or what it is that lexicographers do when they are editing or reediting a dictionary. But even questions of a very different sort on the surface may unwittingly reveal much the same interest. "Why did you fail to include word x in your book?" and "Why don't you still use the system of transcribing pronunciations with which I grew up instead of the present one with its 'up-side down e'?" may be in part expressions of annoyance at what is seen (not always accurately) as the dictionary's failure to do its job, but they are just as truly demands to know how dictionary editors make the decision to exclude some words from a given dictionary or to revise a long-standing feature of earlier editions. What follows is an effort to present a brief overview of the English language and its history and to provide brief and necessarily somewhat general answers to a few of the questions that users of this dictionary probably have about it, the processes that went into its making, and its relation to that fascinating and sometimes maddening marvel which we call the English language.

Language is the object of study of the academic discipline known as linguistics. Although the roots of linguistic science are found in earlier centuries, it is in most respects a modern creation, and the understanding of language that it offers us differs in a number of fundamental ways from the conceptions of language held by thinkers of the ancient and medieval

worlds, the Renaissance, or the Enlightenment. This understanding does not differ, however, in every way. The use of language is still seen by linguists as a peculiarly human activity. We often use the word *language* to refer to the limited stock of movements or utterances by which some animals communicate a limited number of messages, but in doing so we recognize that we are speaking of something different in kind from our own language. Moreover, modern definitions of language are just as likely as earlier definitions to emphasize its functional aspect: language enables human beings, at least those who share a particular language, to communicate with each other by stating ideas, expressing feelings, and exchanging information.

Modern definitions of language, though, are more likely than older ones to stress some other aspects. One is the arbitrary nature of the relationship between the conventional sounds or other signs which serve as the vehicle of language and the meaning being conveyed by them. A few naive souls may believe that domestic swine are called pigs because their habits are so dirty, but it is clear to most that no inherent or necessary connection exists between the sequence of sounds *\p\* plus *\i\* plus *\g\* and "any of various stout-bodied short-legged omnivorous mammals (family Suidae) with a thick bristly skin and a long mobile snout." A similarly naive Frenchman or German could insist with neither more nor less reason that *cochon* or *Schwein* is the word that naturally expresses the essential piggishness of the animal.

A linguistically oriented definition of language would also be likely to emphasize the systematic nature of language. Were it not for highly organized systems operating within any natural language – be it one with many millions of native speakers and enormous international importance like Spanish, Mandarin Chinese, or English or one spoken by a handful of people in a remote area – it could hardly be the subtle and effective tool of communication that it is. These systems are enormously complex both in themselves and in their mutual interaction, so complex indeed that no language has yet had its workings fully described; yet, paradoxically, and fortunately for the human race, any child with a normal ability to learn can, within a very few years, master at least the essentials of these systems for any one language with which it is in daily contact (or even two languages if the child's environment is bilingual). It is uncertain whether this is so because, as some linguists believe, at a profound level of structure the details that make English so different from other languages are unimportant

and the systems of all languages are largely the same, but the fact that we learn our native language almost effortlessly up to a certain basic level of control is hardly to be denied.

## **The Systems of Language**

The major systems that make up the broad comprehensive system of language itself are four in number: lexicon, grammar, semantics, and phonology. The one that dictionary editors and dictionary users are most directly concerned with is the vocabulary or *lexicon*, the collection of words and word elements which we put together in various ways to form larger units of discourse: phrases, clauses, sentences, paragraphs, and so forth. All languages have a lexicon, and all lexicons are governed by rules that permit some kinds of word formation, make others dubious, and render still others clearly impossible. In English we might say *versatileness* without hesitation if we needed such a word and could not for the moment think of *versatility*, even though the former is not normally part of our everyday working vocabulary; but *versatilize* might give us considerable pause, and *nessversatile* we would simply never utter. The size of the lexicon varies considerably from language to language. The language of an isolated people, for example, may be perfectly adequate with a relatively small and fixed vocabulary, since it has no need of the coinages attendant upon modern technology, while English and other major languages have enormous stocks of words, to which they add year by year at a great rate. Since the dictionary is concentrated upon the lexicon, our discussion of the other systems of language, as it proceeds, will be largely concerned with how they are related to the lexicon and thus are important within the dictionary.

The grammatical system of language governs the way in which words are put together to form the larger units of discourse mentioned earlier. Grammar, of course, varies a great deal from language to language just as the lexicon does: in English, word order is a dominant factor in determining meaning, while the use of inflectional endings to mark the grammatical function of individual words within a sentence plays a clearly subordinate role, though important in some ways (as in indicating the number of a noun, the case of a personal pronoun, and the tense of a verb). Other languages show markedly different patterns, such as Latin with its elaborate set of paradigms for nouns, verbs, adjectives, and pronouns and its highly flexible word order. The semantic system of a language has to do

with meanings and thus with the relation between the conventionalized symbols that constitute language and the external reality about which we need to communicate through language. The phonological system of a language is what allows a speaker of that language to transform a grammatical unit embodying a meaning into a flow of uttered sounds that can be heard and interpreted (accurately, if all goes well) by another speaker of the language. This system is always very tightly organized. The inventory of basic meaningful units of sound within a language (called *phonemes* by linguists) is never very large compared with the number of words and word elements in the lexicon; most speakers of English get by with about 40. Phonemes are identified by the fact that in some pairs of words they create a contrast that signals a difference in meaning: we consider the vowel sounds of *trip* and *trap* to be different phonemes because the difference in vowel sounds is the sole determinant of their being distinct words. Their consonant sounds are identical. Similarly, the initial consonant sounds of *pull* and *bull*, *tie* and *die*, and *come* and *gum* are contrasting phonemes. On the other hand, the sound at the beginning of *pit* and at the end of *tip* are phonetically quite different, but as they do not contrast meaningfully we do not perceive them as distinct phonemes. The combinations of these phonemes permitted in a given language are severely restricted, as are the ways in which speech sounds occur in conjunction with other significant elements of the phonological system such as stress (force or intensity) and intonation (the rise and fall in pitch of the voice as it moves through an utterance). In English, for example, it is possible for the consonants *bsstrbs* to occur in succession, but only at the beginning of a word (as in *strict*) or in the middle (as in *monstrous*), not at the end, and the sequence *bspgrbs* cannot occur at either the beginning or the end of a word but may occur in the middle (as in *upgrade*).

## **Variation and Change in Language**

All of the systems of language are constantly in operation, and in a given language at a given time they may seem almost to be monolithic or at least to have sufficient identity that it makes sense, for example, to talk of *the* grammar of English. And, indeed, how could it be otherwise? Language would be a far more imperfect tool of communication than it is if the speakers of a language were not functioning within a system sufficiently unified to permit almost constant mutual intelligibility. Yet the impression of unity which we receive when we take a broad descriptive look at a single language at a particular time (taking a *synchronic* point of view, as linguists

say) is more than a little misleading because it has failed to take account of the enormous variation that exists within the language.

Each of us speaks a distinctive form of English which is not identical in every particular with the form spoken by anyone else; linguists call this individual variety of a language an *idiolect*. Those who speak idiolects sharing certain features of vocabulary, grammar, and phonology that are distinctively different from corresponding features shared by others who live in a different geographical area or belong to a different social group or who differ in some other way that affects their language are said to speak a *dialect* of the language. Researchers have identified a number of different geographical dialect areas within the United States, rather clearly marked on the Eastern seaboard but progressively less well defined as one moves west; yet, the dialects of American English do not differ overall very greatly from one another; some dialects of Great Britain are more strikingly divergent in phonology, for example, than are any two dialects within this country, and some dialects of other languages approach the condition of mutual unintelligibility that is often taken to divide separate dialects from separate languages.

Nor is variation in language by any means confined to matters of idiolect or dialect. Variation may also be related to the several functional varieties of a language that people take up and discard as their roles and relationships change from moment to moment throughout the day. Such variation can involve vocabulary, pronunciation, and even grammar. A worker who queries one colleague concerning the whereabouts of another with "Seen John?" from which both the auxiliary verb *have* and the subject *you* have been deleted may not put the question in the same informal way to a superior.

If variation is one of the most prominent aspects of language as one considers it at the present time, the inescapable fact that emerges from considering language historically (taking a *diachronic* point of view, as linguists say) is change. No living language stands still, however much we might wish at times that it would. Change over the short run is most readily noticed in the lexicon, as a comparison of successive editions of any modern dictionary will show; in grammar and phonology the forces of change typically operate much more slowly. Still, the cumulative effect of changes that are imperceptible as they occur can be impressive when measured across the centuries. The English of one's great-great-

grandfather might not sound so very different from one's own. Perhaps it might seem a bit stiff and formal, a bit old-fashioned in its vocabulary, but the differences would not be dramatic. If we could somehow listen to an English-speaker of King Alfred's time, however, we would hear what all but a few scholars of historical English would take to be a foreign tongue.

## The History of English

The history of English is conventionally, if perhaps too neatly, divided into three periods usually called Old English, (or Anglo-Saxon), Middle English, and Modern English. The earliest period begins with the migration of certain Germanic tribes from the continent to Britain in the fifth century A.D., though no records of their language survive from before the seventh century, and it continues until the end of the eleventh century or a bit later. By that time Latin, Old Norse (the language of the Viking invaders), and especially the Anglo-Norman French of the dominant class after the Norman Conquest in 1066 had begun to have a substantial impact on the lexicon, and the well-developed inflectional system that typifies the grammar of Old English had begun to break down. The following brief sample of Old English prose illustrates several of the significant ways in which change has so transformed English that we must look carefully to find points of resemblance between the language of the tenth century and our own. It is taken from Aelfric's "Homily on St. Gregory the Great" and concerns the famous story of how that pope came to send missionaries to convert the heathen Anglo-Saxons to Christianity after seeing Anglo-Saxon boys for sale as slaves in Rome:

Eft he axode, hu ðaere ðeode nama waere bpe hi of comon. Him waes geandwyrd, bpaet hi Angle genemnode waeron. bpa cwaeð he, "Rihtlice hi sind Angle gehatene, for ðan ðe hi engla wlite habbað, and swilcum gedafenað bpaet hi on heofonum engla geferan beon."

A few of these words will be recognized as identical in spelling with their modern equivalents – *he, of, him, for, and, on* – and the resemblance of a few others to familiar words may be guessed – *nama* to *name*, *comon* to *come*, *waere* to *were*, *waes* to *was* – but only those who have made a special study of Old English will be able to read the passage with understanding. The sense of it is as follows: "Again he [St. Gregory] asked what might be the name of the people from which they came. It was answered to him that they were named Angles. Then he said, 'Rightly are

they called Angles because they have the beauty of angels, and it is fitting that such as they should be angels' companions in heaven.' " Some of the words in the original have survived in altered form, including *axode* (*asked*), *hu* (*how*), *rihtlice* (*rightly*), *engla* (*angels*), *habbað* (*have*), *swilcum* (*such*), *heofonum* (*heaven*), and *beon* (*be*). Others, however, have vanished from our lexicon, mostly without a trace, including several that were quite common words in Old English: *eft* ``again," *æode* ``people, nation," *cw'ð* ``said, spoke," *gehatene* ``called, named," *wlite* ``appearance, beauty," and *geferan* ``companions." Recognition of some words is naturally hindered by the presence of two special characters, *þ*, called ``thorn," and *ð*, called ``edh," which served in Old English to represent the sounds now spelled with *th*.

Other points worth noting include the fact that the pronoun system did not yet, in the late tenth century, include the third person plural forms beginning with *th-*: *hi* appears where we would use *they*. Several aspects of word order will also strike the reader as oddly unlike ours. Subject and verb are inverted after an adverb – *bpa cw'ð he* ``Then said he" – a phenomenon not unknown in modern English but now restricted to a few adverbs such as *never* and requiring the presence of an auxiliary verb like *do* or *have*. In subordinate clauses the main verb must be last, and so an object or a preposition may precede it in a way no longer natural: *bpe hi of comon* ``which they from came," *for ðan æ he engla wlite habbað* ``because they angels' beauty have."

Perhaps the most distinctive difference between Old and Modern English reflected in Aelfric's sentences is the elaborate system of inflections, of which we now have only remnants. Nouns, adjectives, and even the definite article are inflected for gender, case, and number: *æære æode* ``(of) the people" is feminine, genitive, and singular, *Angle* ``Angles" is masculine, accusative, and plural, and *swilcum* ``such" is masculine, dative, and plural. The system of inflections for verbs was also more elaborate than ours: for example, *habbað* ``have" ends with the *-að* suffix characteristic of plural present indicative verbs. In addition, there were two imperative forms, four subjunctive forms (two for the present tense and two for the preterit, or past, tense), and several others which we no longer have. Even where Modern English retains a particular category of inflection, the form has often changed. Old English present participles ended in *-ende* not *-ing*, and past participles bore a prefix *ge-* (as *geandwyrd* ``answered" in the passage above).

The period of Middle English extends roughly from the twelfth century through the fifteenth. The influence of French (and Latin, often by way of French) upon the lexicon continued throughout this period, the loss of some inflections and the reduction of others (often to a final unstressed vowel spelled -e) accelerated, and many changes took place within the phonological and grammatical systems of the language. A typical prose passage, especially one from the later part of the period, will not have such a foreign look to us as did Aelfric's prose; but it will not be mistaken for contemporary writing either. The following brief passage is drawn from a work of the late fourteenth century called *Mandeville's Travels*. It is fiction in the guise of travel literature, and, though it purports to be from the pen of an English knight, it was originally written in French and later translated into Latin and English. In this extract Mandeville describes the land of Bactria, apparently not an altogether inviting place, as it is inhabited by ``full yuele [evil] folk and full cruell."

In bpat lond ben trees bpat beren wolfe, as bpogh it were of scheep; whereof men maken clothes, and all bping bpat may ben made of wolfe. In bpar contree ben many ipotaynes, bpat dwellen som tyme in the water, and somtyme on the lond: and bpei ben half man and half hors, as I haue seyde before; and bpei eten men, whan bpei may take hem. And bpere ben ryures and watres bpat ben fulle byttere, bpree sithes more bpan is the water of the see. In bpat contree ben many griffounes, more plentee bpan in any other contree. Sum men seyn bpat bpei han the body vward as an egle, and benethe as a lyoun: and treuly bpei seyn soth bpat bpei ben of bpat schapp. But o griffoun hath the body more gret, and is more strong, bpanne eight lyouns, of suche lyouns as ben o this half; and more gret and strongere bpan an hundred egles, suche as we han amonges vs. For o griffoun bpere wil bere fleyng to his nest a gret hors, 3if he may fynde him at the poynt, or two oxen 3oked togidere, as bpei gon at the plowgh.

The spelling is often peculiar by modern standards and even inconsistent within these few sentences (*contree* and *contree*, o [griffoun] and a [gret hors], *bpanne* and *bpan*, for example). Moreover, there is in addition to thorn another old character ȝ, yogh, to make difficulty. It can represent several sounds but here may be thought of as equivalent to y. Even the older spellings (including those where *u* stands for *v*) are recognizable, however, and there are only a few words like *ipotaynes* ``hippopotamuses" and *sithes* ``times" that have dropped out of the language altogether. We may notice a few words and phrases that have meanings no longer common such as *byttere* ``salt," *o this half* ``on this side of the world," and *at the poynt* ``to hand," and the effect of the centuries-long dominance of



French on the vocabulary is evident in many familiar words which could not have occurred in Aelfric's writing even if his subject had allowed them, words like *contree*, *ryueres*, *plentee*, *egle*, and *lyoun*.

In general word order is now very close to that of our time, though we notice constructions like *hath the body more gret* and *three sithes more bpan is the water of the see*. We also notice that present tense verbs still receive a plural inflection as in *beren*, *dwellen*, *han*, and *ben* and that while nominative *bpei* has replaced Aelfric's *hi* in the third person plural, the form for objects is still *hem*. All the same, the number of inflections for nouns, adjectives, and verbs has been greatly reduced, and in most respects Mandeville is closer to Modern than to Old English.

The period of Modern English extends from the sixteenth century to our own day. The early part of this period saw the completion of a revolution in the phonology of English that had begun in late Middle English and that effectively redistributed the occurrence of the vowel phonemes to something approximating their present pattern. (Mandeville's English would have sounded even less familiar to us than it looks.) Other important early developments include the stabilizing effect on spelling of the printing press and the beginning of the direct influence of Latin and, to a lesser extent, Greek on the lexicon. Later, as English came into contact with other cultures around the world and distinctive dialects of English developed in the many areas which Britain had colonized, numerous other languages made small but interesting contributions to our word-stock.

The historical aspect of English really encompasses more than the three stages of development just under consideration. English has what might be called a prehistory as well. As we have seen, our language did not simply spring into existence; it was brought from the Continent by Germanic tribes who had no form of writing and hence left no records. Philologists know that they must have spoken a dialect of a language that can be called West Germanic and that other dialects of this unknown language must have included the ancestors of such languages as German, Dutch, Low German, and Frisian. They know this because of certain systematic similarities which these languages share with each other but do not share with, say, Danish. However, they have had somehow to reconstruct what that language was like in its lexicon, phonology, grammar, and semantics as best they can through sophisticated techniques of comparison developed chiefly during the last century. Similarly, because ancient and modern languages like Old

Norse and Gothic or Icelandic and Norwegian have points in common with Old English and Old High German or Dutch and English that they do not share with French or Russian, it is clear that there was an earlier unrecorded language that can be called simply Germanic and that must be reconstructed in the same way. Still earlier, Germanic was just a dialect (the ancestors of Greek, Latin, and Sanskrit were three other such dialects) of a language conventionally designated Indo-European, and thus English is just one relatively young member of an ancient family of languages whose descendants cover a fair portion of the globe. (For more detail on the Indo-European languages and their relationships, see the table having that title in the dictionary.)

### **The Dictionary and the Systems of English**

By far the largest part of this volume is called "A Dictionary of the English Language" and so is naturally concerned with the systems of English that we have cursorily surveyed in their synchronic and diachronic aspects. In fact, information related to all four systems is given at most entries in the dictionary, as well as information related to what could reasonably be considered a fifth system of English and many other (though not all) languages – writing. The writing system provides an alternative to speech that permits long-distance transmission and visual reception of a communication and also enables a record to be kept for much longer than human memory can keep it. The writing system of Modern English allows for considerable variation, as is shown by the persistence of variant spellings like *veranda* and *verandah* or *judgment* and *judgement* and by the fact that many compound words have solid, hyphenated, and open stylings all in common use (as *decision maker*, *decision-maker*, and *decisionmaker*). At the same time, however, it tends to be a force for standardization and unification because recorded language creates a precedent for future language use and provides a basis on which language use can be taught to the younger members of a community. This conservative effect is one reason why spelling reformers have so far met with but modest success in their efforts.

We may now begin to look at the ways in which the specific systems of our language are treated in the dictionary and at the processes of lexicography which produce the information about these systems that the dictionary user encounters. A dictionary is necessarily and obviously concerned with the lexicon above all, and the information it can convey about the language

systems is confined to the level of the word or short phrase. The result is that no dictionary of English, however good it may be, can provide all of the information about the English language that one might wish to have at one time or another. Thus, for example, details about such important aspects of phonology as the patterns of sentence stress and sentence intonation cannot readily be accommodated in a work of reference organized in terms of discrete words, nor can grammatical topics such as word order in subordinate clauses or the structural relation of interrogative to declarative sentences.

## The History of English in the Dictionary

A similar limitation applies to the treatment of the historical aspect of English; yet, Webster's Ninth New Collegiate Dictionary is able to offer a good deal of historical information about words. What we earlier called the prehistory of English is encountered in the etymologies that appear in square brackets ahead of the definitional material at many entries. An etymology tells us what is known of an English word before it became the word we enter in the dictionary; that is, if the word was created in English the etymology shows, to whatever extent is not already obvious from the shape of the word, what materials were used to form it, and if the word was borrowed into English the etymology traces the steps of the borrowing process backward from the point at which the word entered English to the earliest recorded ancestral language. Where it is most relevant, note is made of one or several words from other languages that are related ("akin") to the entry word but are not in the direct line of borrowing. Thus, a word like Aelfric's *heofon* (ignoring for the moment the dative plural inflection *-um* that it bears in the passage we looked at earlier) appears as part of this dictionary's etymology for the modern word *heaven*: [ME *heven*, fr. OE *heofon*; akin to OHG *himil*]. Since *heaven* is a native English word, it has only two recorded ancestors, Middle English *heven* and Old English *heofon*. Beyond those forms lie only the hypothetical, reconstructed forms of West Germanic, Germanic, and Indo-European. In this case one West Germanic cognate is shown, Old High German *himil*, which is the parent of Modern German *Himmel* but only a second cousin of our English word. Similarly Mandeville's *contree* appears as the first element in the etymology of its modern descendant *country*: [ME *contree*, fr. OF *contree*, fr. ML *contrata*, fr. L *contra* against, on the opposite side]. Here we see that our word can be traced back through three nouns of Middle English, Old French, and Medieval Latin (all of which had the same basic meaning as

the Modern English noun and so are not glossed) to a Latin preposition (which has a different meaning and so is glossed). The two etymological patterns are, as we would expect from what we know of the history of the English vocabulary, among the most common and are repeated with differing details at entry after entry throughout the book. Of course, borrowings that have occurred within the Modern English period are more various, and we find such exotic language names as Nahuatl (at *chocolate*), Taino (at <sup>2</sup>*barbecue*), Tagalog (at *boondocks*), Malay (at <sup>3</sup>*amok*), and Kimbundu (at *banjo*) as well as the more familiar Russian (at *troika*), Italian (at <sup>1</sup>*ballot*), Arabic (at *mullah*), Spanish (at <sup>1</sup>*macho*) and Japanese (at *tycoon*).

An etymologist must know a good deal about the history of English and also about the relationships of sound and meaning and their changes over time that underlie the reconstruction of the Indo-European family, but even that considerable learning is not enough to do all that must be done to provide etymologies of English words in a dictionary such as this. A knowledge is also needed of the various processes by which words are created within Modern English: among the most important processes are shortening, or clipping (see <sup>1</sup>*stereo*), functional shift (as the noun *commute* from the verb *commute*), back-formation (see *grid*), combination of initial letters (see *radar*), transfer of personal or place names (see *silhouette* and *denim*), imitation of sounds (see <sup>1</sup>*whiz*), folk etymology (see *Jerusalem artichoke*), and blending of two words (see *motel*). Also available to one who feels the need for a new word to name a new thing or express a new idea is the very considerable store of prefixes, suffixes, and combining forms that already exist in English. Some of these are native and others are borrowed from French, but the largest number have been taken directly from Latin or Greek, and they have been combined in many different ways often without any special regard for matching two elements from the same original language. The combination of these word elements has produced many scientific and technical terms of Modern English. Once in a while a word is created spontaneously out of the creative play of sheer imagination. (For examples of the latter sort of creation see the etymologies of *boondoggle* and *googol* in the dictionary. Such invention is common, as Merriam-Webster editors know from their mail, which frequently includes requests from coiners that their brand-new words be entered in the dictionary. Very few coinages of this kind ever come into common enough use to justify dictionary entry, however.)

An etymologist working on a new edition of the Collegiate Dictionary must review the etymologies at existing main entries and prepare such etymologies as are required for the main entries being added to the new edition. In the course of the former activity adjustments must sometimes be made either to incorporate a useful piece of information that has previously been overlooked or to revise the account of the word's origin in the light of new evidence. Such evidence may be unearthed by the etymologist or may be the product of published research by scholars of historical linguistics and others. In writing new etymologies this editor must, of course, be alive to the possible languages from which a new term may have been borrowed and to the possible ways in which one may have been created. New scientific and technical terms sometimes pose special difficulties. While they are most often formed from familiar word elements, occasionally a case like *methotrexate* presents itself in which one element (here *-trexate*) resists identification.

When all attempts to provide a satisfactory etymology have failed, the editor has recourse to the formula "origin unknown." This formula seldom means that the editor is unaware of various speculations about the origin of the term but instead usually means that no single theory conceived by the etymologist or proposed by others is well enough backed by evidence to include in a serious work of reference, even when qualified by "probably" or "perhaps." Thus, our editors frequently have to explain to correspondents that the dictionary fails to state that the origin of *posh* is in the initial letters of the phrase "port out, starboard home" – supposedly a shipping term for the cooler accommodations on steamships plying between Britain and India from the mid-nineteenth century on – not because the story is unknown to us but because no evidence to support it has yet been produced. Some evidence exists that casts strong doubt on it; the word is not known earlier than 1918 (in a source unrelated to shipping), and the acronymic explanation does not appear until 1935. It therefore seems reasonable to consider the acronymic explanation a modern invention and assign *posh* the etymology [origin unknown]. The etymologist must sift such theories, often several conflicting theories of greater or lesser likelihood, and try to evaluate the evidence conservatively but fairly in arriving at the soundest possible etymology that the available information permits. Occasionally time will prove the result to be somewhat (or even quite) mistaken, and the etymology will need to be replaced by something better. This can happen even when the etymologist felt quite certain of the

soundness of the original etymology, and it is just one among many reasons why dictionaries must be reedited from time to time if they are to remain reliable.

Historical information about words is also provided by the date appearing in parentheses just before the first or only definition at most main entries. The date given is for the earliest recorded use known to our editors of the first entered sense of that entry. In most cases the date is also, in effect, for the earliest use of the word itself that we know of. Some words, however, had early senses that later passed from common use without gaining special literary importance, and these senses are omitted from this dictionary. Because it would be misleading to give a date for a sense that the dictionary does not show, the date is always for the first sense actually defined at the entry. Because the senses of any word having more than one are always presented in historical order, with the one known to have been used first given first, the date serves as a link between the prehistory of the word shown in the etymology and its later recorded history of semantic development within the language as reflected in the order of definitions.

Evidence for the dates has come from a number of sources. Especially for words that have been a part of the language since before this century, the most important sources have been the major historical dictionaries of English. These works include for each sense dated examples of use from one or several authors including the earliest one available to the editors. Chief among these dictionaries is the majestic thirteen-volume Oxford English Dictionary and its supplements. Also of great importance have been the Middle English Dictionary, A Dictionary of American English, A Dictionary of Americanisms, The Scottish National Dictionary, and A Dictionary of the Older Scottish Tongue. Other dictionaries that include a greater or lesser number of dated quotations and that have proved helpful in particular cases include Hobson-Jobson (a glossary of Anglo-Indian terms), The Stanford Dictionary of Anglicised Words and Phrases, Wright's English Dialect Dictionary, Cassidy and Le Page's Dictionary of Jamaican English, Branford's Dictionary of South African English, Avis's Dictionary of Canadianisms on Historical Principles, Wentworth's American Dialect Dictionary, Wentworth and Flexner's Dictionary of American Slang, The Barnhart Dictionary of New English Since 1963, and The Second Barnhart Dictionary of New English. The Century Dictionary and Cyclopedia and the successive editions of Merriam-Webster's unabridged dictionaries and their

supplements of new words have also provided much assistance, for while these dictionaries do not incorporate dated quotations, an entry in one or another of them is sometimes earlier than any example of the word from running text that we have been able to find.

The other major source of dates, especially for the period from 1890 to the present, is the Merriam-Webster file of examples of words used in context, which are called citations. More will be said of this collection later. Here it need only be noted that among the 13,000,000 slips which the file contains frequently appear one or more examples of a given word that are earlier than any quoted in our reference sources. And, of course, our citations have been essential to the dating of a considerable number of entries not included in any of the dictionaries mentioned above. The date of 1949 at *classical conditioning* is a case of the first sort, the earliest example in a reference source being from 1964, while the date of 1974 at *earth tone* is of the second sort. Some of the older books in our editorial library and in other libraries to which our editors have access have occasionally supplemented the resources of the citation file in supplying dates.

Almost from the appearance of the first volume of The Oxford English Dictionary, scholars have been discovering earlier dates for particular words and senses by examining works not searched for examples by the dictionary's readers or by reading some works a second time and publishing the results of their findings in various journals. Many hundreds of entries in this dictionary include a date derived from one of these articles, and while far too many scholars and other interested students have participated in this work for a listing here to be practical, some collective acknowledgment of our debt to them is necessary. The date of 1676 at *menagerie* may be cited as an example of one derived from a source of this kind; the newly discovered quotation is 36 years older than the earliest example that had previously been found.

The style of the date is determined by the period of English to which the sense being dated belongs: for entries from Old English we indicate simply that the example is from the period before the twelfth century (bef. 12c); for those from Middle English we indicate their century, as (14c); for those from Modern English we give a single year, as (1742).

Some caution needs to be exercised in interpreting the significance of a date. It is never meant to indicate the exact point at which a word entered

the language. For one thing, words have often been in spoken use for many years before they come to be written down. Then, too, not all surviving texts, even for the earlier periods, have been read to collect examples for any historical dictionary, and obviously for the modern era only a very small sample of all published material has been examined in that way. One can perhaps with some justification think of the date as indicating a time by which one can be sure that the word was in use, but it will be safest simply to remember that the date actually belongs to the earliest occurrence known to the editors of this dictionary of the first entered sense of the word.

Leaving the historical aspect of English aside now, we may consider how information about the systems of English as they presently exist is recorded in this dictionary. The phonological system needs little more than a mention here. Its role in this dictionary is discussed in some detail in the "Guide to Pronunciation," which immediately follows this section, as is the way in which the pronunciations shown in the dictionary have been determined.

## **Semantics in the Dictionary**

In turning to consider the coverage of the semantic system in the dictionary, we face several difficult problems. If one function of a dictionary is more important than its many others, surely that function is to define the meaning of words. But while definition is central to the dictionary and quite obviously is involved with semantics, for the most part it deals with individual words in isolation from other words and thus ignores, to a considerable extent, the systematic, relational side of English semantics. Another problem is that although we know quite a lot about the system of English phonology and a good deal (though less) about the grammatical system, our understanding of the semantic system is very imperfect, and much of what we do know about it does not come very obviously into play in a dictionary. Still, we will have a glimpse of this system when we consider the dictionary treatment of synonyms, and in the meantime there is much to be said about the defining of words. Perhaps the first thing that we need to remind ourselves of is that when we speak of the meaning of a word we are employing an artificial, if highly useful, convention. Meaning does not truly reside within the word but in the minds of those who hear or read it. This fact alone guarantees that meaning will be to a great degree amorphous; no two people have had exactly the same experience with what a word refers to and so the meaning of the word will be slightly or greatly different for each of us. It is obvious,



then, that a dictionary which set itself the task of defining the meanings of words in their entirety would be a foolhardy enterprise. So dictionary editors invoke the traditional distinction between *denotation* – the direct and specific part of meaning which is sometimes indicated as the total of all the referents of a word and is shared by all or most people who use the word – and *connotation* – the more personal associations and shades of meaning that gather about a word as a result of individual experience and which may not be widely shared. The dictionary concerns itself essentially with the denotations of words.

For the editors of this dictionary the defining process began long before they actually sat down to examine critically the definitions of the last edition and to formulate trial definitions. It began with an activity that is called in our offices “reading and marking.” Ordinarily each editor spends a portion of the working day reading a variety of newspapers, magazines, and books, looking for anything that might be useful to a definer of English words. Because both time and staff are limited and the scope of English seems nearly unlimited, changes in subject matter, geographical area covered, and individual publications must be made from time to time in a way carefully calculated to ensure the breadth and depth as well as the continuity of our coverage of the vocabulary of English. An editor who is reading and marking will, of course, be looking for examples of new words and for unusual applications of familiar words that suggest the possible emergence of a new meaning but will also be concerned to provide evidence of the current status of variant spellings, inflected forms, and the stylings of compound words, to collect examples that may be quotable as illustrations of typical use in the dictionary, and to record many other useful kinds of information. In each instance the reader will underline the word or phrase that is of interest and mark off as much context as is considered helpful in clarifying the meaning. This example of a word used in context is called a *citation* of the word. Ideally the editor would like all citations to illuminate the meaning of the word, but some passages will remain obscure no matter how far they extend, and sometimes one must mark a citation simply for the occurrence of the word or meaning (especially when it is new), trusting that the reading-and-marking process will yield more helpful examples in the long run. In the case of ephemeral words, of course, this may never happen, but truly ephemeral words will not need to be defined for a dictionary. At this early stage of the dictionary-making process editors do not make judgments about the likelihood of a word's establishing itself in the language. If a possible citation has even the barest potential to be

useful at a later time, it is marked.

These samples of words in bracketed context are put onto 3x5 slips of paper, and the citation slips are placed in alphabetical order in rows of filing cabinets. They will be used, as needed, by the editors in their roles as writers of definitions and certain other parts of dictionary entries. The editors engaged in this ninth edition of the Collegiate reviewed every one of the million and a half citations that had been gathered since the eighth edition was prepared in the early 1970s. When necessary, they also drew upon the additional resources of what are called the "consolidated" files, those that contain all the citations (over eleven and a half million) that had been accumulated in our offices since the late nineteenth century and had been used in the editing of the many dictionaries this company published before the present one.

The actual defining process begins with a number of special assignments called "group defining projects," which may range from a small set of words like those for the days of the week or the letters of the English alphabet (for which parallel, formulaic definitions are required) to the vocabulary of a large subject area such as music or anthropology. When these assignments have been completed, defining proceeds alphabetically, with the editors responsible for the terminology of the life sciences or the physical sciences and related technologies working independently of the editors responsible for defining the general vocabulary.

If you were a definer, you would typically be working at a given moment with a group of citations covering a relatively short segment of the alphabet, *grio-* to *gror-*, for example, and with the entries of the dictionary being reedited that fall within the same segment. Your job would be to determine, under the guidance of the citations, which existing entries could remain in the new edition essentially unchanged because their usage showed no significant alteration, which entries needed to be revised either by modification of existing definitions or by the addition of new ones, which old entries were expendable for the new edition, and what new entries should be added to keep coverage of the lexicon up-to-date. You would begin by reading and sorting out the citation slips, first by grammatical function, in the case of a word like *groom* that is both noun and verb, and then by meaning within each part of speech. For each group of citations that was covered by an adequate existing definition, you would need only to indicate that you had examined them and would do nothing to the definition. For

definitions needing adjustment, you would indicate the change to be made. In many cases, you would have some citations left over that were not covered by an existing definition, and it would then be your job to determine whether that segment of meaning was perhaps relatively uncommon and not backed by a sufficient range and number of citations and so not needed for the dictionary or whether in fact it was a sense that dictionary users are entitled to find suitably defined when they come looking for it. In the former case you would reject the citations, and eventually they would find their way back to the files to await review for another dictionary (by which time perhaps the citational backing would be stronger and a definition needed). In the latter case yours would be the responsibility to frame the kind of definition that will adequately convey that particular segment of meaning to the dictionary user.

In writing that definition, you could follow any of a number of paths marked out by the instructions given to each definer. These include both the general policies and practices that govern all Merriam-Webster dictionaries and the more specific directions and prohibitions contained in the ``style file," as it is called, for this particular dictionary.

The kind of definition that you would write in most cases is called an analytical definition. It consists in its purest form of the statement of a class to which the term being defined is assigned and a number of characteristics which differentiate the individual from other members of the class. For example, the first sense of *grove* is defined in this Collegiate as ``a small wood without underbrush," assigning a grove to the more general class of woods and using ``small" and ``without underbrush" to indicate in what ways a grove is unlike other kinds of woods. Another possibility would be for you to define a synonym, as is done at the sixth sense of the noun *grip*, where the definition is ``STAGEHAND." Defining by synonym tends to be inexact because even true synonyms do not have just the same meaning and is perhaps most useful in cases like the one just mentioned where one kind of referent has two or more names, a situation that occurs frequently with the common names of plants and animals. For this reason we link any synonym definition to an analytical definition by making the synonym a cross-reference (in small capital letters) to another entry where an analytical definition suitable for both words is given: at *stagehand* is the definition ``a stage worker who handles scenery, properties, or lights," which is also a good definition of the sixth sense of *grip*.

Within these basic defining patterns many variations are permitted. Some analytical definitions may justifiably be truncated by the use of a related word within the definition in order to save precious space for more entries. For example, *grievance committee* can be defined as ``a committee formed by a labor union or by employer and employees jointly to discuss and where possible to eliminate grievances" because the second sense of *grievance* is ``a cause of distress (as an unsatisfactory working condition) felt to afford reason for complaint or resistance," and so the definition of *grievance committee* need not give that information a second time.

It is also possible to add a synonomous cross-reference to an analytical definition and thereby incorporate at little cost of space a second version of the meaning that looks at it from a slightly different aspect. It is possible, as in the definition of *grievance* just cited, to add a parenthetical element that specifies one or several of the typical referents of the word or that indicates the sole or a typical object of a transitive verb. One may begin an adjective definition with one of a wide variety of formulas but others are forbidden. It is clear already that definers' instructions are elaborately detailed, and it would be tedious to rehearse them here. Their purpose is to assist in developing the definer's native talent so that the definitions that he or she writes are consistently good ones. What is a good definition? Many qualities could be mentioned, and probably different definers would rank the relative importance of those on any list differently; but all definers want a sufficient range and number of their definitions to be objective in reflecting what the word means as it is actually used rather than what the definer or someone else thinks it ought to mean, and they want their definitions to be accurate, clear, informative, and concise. In short, they want their definitions to have the qualities that users have in mind when they call a dictionary they admire ``authoritative."

In the course of your defining, you would have an opportunity fairly often to make another kind of decision: whether to include or omit a new candidate for main entry. Let us take as an example the word *gentrification*, which is one of many entries new to this edition of the Collegiate Dictionary. If you had been the definer who handled that word, you would have been faced with a group of 26 citations to read covering a span of five years and including extracts from such publications as *The New York Times*, *Scientific American*, *Playboy*, *American Demographics*, *The Boston Globe*, *Smithsonian*, *Money*, *The Christian Science Monitor*, *Saturday Review*, *The Wall Street Journal*, and *Harper's*. In reading the citations you would notice

that while they varied in many details of context, they seemed (with only a single exception) to embody one meaning. There was a significant variable, however; in some citations the immigration was taking place in run-down areas and in others the process of restoring the neighborhood was already well advanced. You might then have produced the following definition incorporating that variation: ``the immigration of middle-class people into a deteriorating or recently renewed city area."

The number and time span of the citations and the variety of the sources would already have told you that this was a very strong, and perhaps even an essential, candidate for entry in the new edition. There is no magic number of citations that guarantees entry and no particular span of years that must be reached. To a great extent the judgment made here must rest on your insight and experience as a definer who has seen the citational backing for many words, who has most likely defined words for other Merriam-Webster dictionaries in the past, and who thus has some sense of the relative importance and degree of establishment of new entries within the lexicon and of their likely staying power.

You would have noticed that in addition to the evidence for *gentrification* there were also eight citations for a verb *gentrify*, and, seeing both that it was less important than the noun (though also well backed by citations) and that its meaning was easy to infer from the meaning of the noun, you would have added it to your new main entry as an undefined run-on.

To take one further example of a somewhat different kind, if you had been the life-sciences definer responsible for handling the term *Reye's syndrome*, you would have read 13 citations. Many of these would have been from sources such as *The Journal of the American Medical Association*, *Biological Abstracts*, *Science*, and *Emergency Medicine* likely to be seen chiefly by people with specialized interests; but you would also have seen examples from *Newsweek*, *Parade*, *The New York Times*, and an encyclopedia yearbook. In other words, the term is likely to be encountered by people with general interests and, given its nature, will probably be looked up in a dictionary fairly often. Such considerations would have led you to propose entry for the term and with a much higher priority than if the citational backing had been nearly all technical. You would also have noted several citations for the spelling *Reye syndrome* and would have appropriately added that as a secondary variant.

It is worth noting briefly that in the course of your work as a definer you would have been concerned with what the citations reveal about a word in addition to its meaning. The definer is initially responsible for most of the framework of the entry including not only spelling variants and run-ons but also inflected forms, usage notes, verbal illustrations and illustrative quotations, and temporal, regional, stylistic, and subject labels.

The other important part of the entries in this dictionary that is concerned with English semantics is the synonym paragraph. These paragraphs are not written by each individual definer as particular entries are encountered but are rather the special assignment of one, or sometimes several, editors who decide which words will be included in a single paragraph and at which entry the paragraph will be placed. The synonym editor has a number of responsibilities in addition to the actual writing or revising of the synonym paragraph. Each entry for a term discussed in a paragraph must be checked to ensure that the definition of a given sense is fully consonant with its treatment in the paragraph, and the editor has the authority to make small adjustments of definitions so that no discrepancies which might puzzle a user remain.

Like the definer, this editor must read citations very carefully to see that the opening statement of the core meaning shared by the synonyms includes neither too much nor too little, that each discrimination of one word from the others is accurate, and that where space permits typical examples may be selected to include as illustrative quotations or as verbal illustrations. It is particularly in these paragraphs that the dictionary user comes into contact with the systematic side of English semantics because here the concern is with the relationship of meanings instead of the meanings themselves as discrete entities. For example, the synonym paragraph at *splendid* in this dictionary states that *splendid, resplendent, gorgeous, glorious, sublime, and superb* mean "extraordinarily or transcendently impressive." This statement of meaning is at once too broad and too narrow to be a good definition for any of the words; it trims away the particular elements of meaning that make each word distinctive (the most important of which are stated in the following discussion). It does, however, give us an accurate notion of the point at which these words come into a precise semantic relationship with each other.

## **Grammar and Usage in the Dictionary**

The last of the four systems of English whose reflection we may see, at least briefly and occasionally, in the dictionary is the grammatical system. As we saw earlier, this system involves chiefly the relationship between words as they form more complex units rather than individual words themselves. A descriptive grammar of English is a very different kind of book from a dictionary. Nevertheless, virtually every entry in this dictionary contains at least one piece of information about its grammatical nature and the kinds of relationships it can enter into, namely, the functional label which typically indicates the part of speech of the entry or, in the case of terminal word elements, the part of speech of the words that they form. If an entry is labeled *adv*, we know that it can describe the action of a verb but cannot itself be the main verb of a sentence, while an entry labeled *n* cannot link the subject of a sentence with a predicate adjective but can be the subject. Other parts of the entry also give us information that is grammatical in nature. One sort of information is offered by the boldface inflected forms that are shown at every entry for which they are irregular exceptions to the ordinary patterns of English inflection or may present some other sort of problem to the dictionary user. Another is offered by the undefined run-on entries. They illustrate the complex patterns by which one word or a number of words can be derived from a single base by means of affixation or functional shift. Certain kinds of usage notes following or standing in place of definitions also present grammatical information. Typical of the former kind of usage note is the one given at sense 2b of *boy*, "often used interjectionally," and the one given at sense 2b of the verb *conk*, "usu. used with *off* or *out*." Typical of the latter are the several usage notes at the entry for the preposition *for*, "used as a function word to indicate duration of time or extent of space" at sense 9, for example.

Usage is a concept that embraces many aspects of and attitudes toward language. Grammar is certainly only a small part of what goes to make up usage, though some people use one term for the other, as when they label what is really a controversial point of usage a grammatical error. Usage guidance is offered in this dictionary in many ways; it would be little exaggeration to say that any information a user seeks and finds in this book can offer some guidance as to usage. But usage information is chiefly conveyed through three devices: usage notes; temporal, regional, and stylistic labels; and usage paragraphs. The first two are developed by definers from their examination of citations, including sometimes (and particularly in the case of the labels) citations found in historical, dialect, and slang dictionaries as well as those in Merriam-Webster's citation file.

The usage paragraphs like the synonym paragraphs are the result of a special project chiefly in the hands of two editors with assistance from several others. The editors attempted to select particular problems of confused or disputed usage that would be of broad general interest and could be treated at individual entries in the dictionary. The great majority of them involve words that have traditionally been points of dispute (a few of these are now probably more traditional than truly the subjects of heated dispute), but some are relatively new items for this kind of consideration. Several paragraphs deal with pronunciation, a subject rarely treated in books about usage.

The editors who wrote the paragraphs used several kinds of material: books describing one or another aspect of the history of usage as a problem in English; books and articles ruling on particular points of usage, whether the product of one person or a group; historical and other dictionaries, and above all citations of usage itself from our file. In digesting this mass of information and presenting it in a very brief compass, the editors have typically combined information on the history of the controversy, the current state of expressed opinion, illustrations of both old and modern use (often quoted), and practical advice.

It has been close to 250 years since Dr. Johnson published his great dictionary and over 150 years since Noah Webster's American Dictionary of the English Language appeared. Even the more modest Collegiate series will be approaching its hundredth birthday before many years have passed. It seems clear that the long tradition of English dictionaries is not likely to wither and die. Indeed, dictionaries are likely to become, if anything, even more important to the general public in the future, at least as long as the vocabulary of English continues the rapid growth which began earlier in this century and which seems now to intensify year by year. As long as they are edited with a proper regard for the right of the dictionary user to have accurate information about what English words actually mean and how they are actually used, those dictionaries will continue to serve a useful purpose and to be needed. Though they are incomplete as descriptions of the systems of English and are edited by fallible humans whose best intentions sometimes fall short of the mark, such dictionaries will continue to form, as the best dictionaries have always done, a helpful bridge between what we know about language and how we use it. Movement across such a bridge is, of course, in both directions: our use of language furnishes the basis for our knowledge of it, but our knowledge of it also helps to use it more



effectively.