

# Introduction to Lexicography

Written by Assistant Professor Tetyana Skibitska

## An Introduction to Dictionaries

The theory and practice of compiling dictionaries is called *lexicography*.

There is some disagreement on the definition of lexicology, as distinct from lexicography. Some use "lexicology" as a synonym for theoretical lexicography; others use it to mean a branch of linguistics pertaining to the inventory of words in a particular language.

It is now widely accepted that lexicography is a scholarly discipline in its own right and not a sub-branch of linguistics, as the object of study in lexicography is the dictionary.

The pursuit of **lexicography** is divided into two related disciplines:

- Practical lexicography is the art or craft of compiling, writing and editing dictionaries.
- Theoretical lexicography is the scholarly discipline of analyzing and describing the semantic, syntagmatic and paradigmatic relationships within the lexicon (vocabulary) of a language and developing theories of dictionary components and structures linking the data in dictionaries. This is sometimes referred to as metalexicography.

Practical lexicographic work involves several activities, and it is important to note that the compilation of really crafted dictionaries require careful consideration of all or some of the following aspects:

- ✓ Profiling the intended users (i.e. linguistic and non-linguistic competences) and identifying their needs
- ✓ Defining the communicative and cognitive functions of the dictionary
- ✓ Selecting and organizing the components of the dictionary
- ✓ Choosing the appropriate structures for presenting the data in the dictionary (i.e. frame structure, distribution structure, macro-structure, micro-structure and cross-reference structure)
- ✓ Selecting words and affixes for lemmatization as entries
- ✓ Selecting collocations, phrases and examples
- ✓ Choosing lemma forms for each word or part of word to be lemmatized
- ✓ Defining words
- ✓ Organizing definitions
- ✓ Specifying pronunciations of words
- ✓ Labeling definitions and pronunciations for register and dialect, where appropriate
- ✓ Selecting equivalents in bi- and polylingual dictionaries
- ✓ Translating collocations, phrases and examples in bi- and polylingual dictionaries

Theoretical lexicography concerns the same aspects, but leads to the development of principles that can improve the quality of future dictionaries.

A **dictionary** is a book of alphabetically listed words in a specific language, with good definitions, etymologies, pronunciations, and other information; or a book of alphabetically listed words in one language with their equivalents in another, also known as a lexicon.

In many languages, words can appear in many different forms, but only the undeclined or unconjugated form appears as the headword in most dictionaries. Dictionaries are most commonly found in the *form of a book*, but some newer dictionaries, like StarDict and the *New Oxford American Dictionary* on Mac OS X (Macintosh Operating System) название операционной системы фирмы Apple Computer для ПК Power Macintosh. Первые версии появились в 1984 г. (тогда она называлась System). В 2001 г. появилась версия Mac OS X с анимированным GUI (Graphical User Interface) графический интерфейс пользователя, ГИП (произносится goo-ee) а) аппарат создания под ОС собственного графического интерфейса пользователя б) программа, позволяющая осуществлять визуализацию данных), are *dictionary software* running on PDAs (Personal Digital Assistant) персональный цифровой секретарь тип сверхлёгкого миниатюрного ПК с ЖК-экраном, клавиатурой и/или рукописным вводом) or computers. There are also many *online dictionaries* accessible via the Internet.

**Printed dictionaries** – Printed dictionaries range from small pocket-sized editions to large, comprehensive multi-volume works.

**Handheld Electronic dictionaries** – Electronic dictionaries are small devices that receive input via a miniature keyboard, voice recognition or a scanning device that reads printed text, and outputs the translation on a small LCD screen or speaks the translation audibly.

**Dictionary programs** – Computer software that allows words or phrases to be input and translated.

**Online dictionaries** – Online dictionaries similar to dictionary programs, these are often easy to search, but not always free to use, and in some cases lack the credibility of printed and electronic dictionaries.

**Glossaries.** A **glossary** is a list of terms in a particular domain of knowledge with the definitions for those terms. Traditionally, a glossary appears at the end of a book and includes terms within that book which are either newly introduced or at least uncommon.

A **bilingual glossary** is a list of terms in one language which are defined in a second language or glossed by synonyms (or at least near-synonyms) in another language.

In a more general sense, a glossary contains explanations of concepts relevant to a certain field of study or action. In this sense, the term is contemporaneously related to ontology (a subject of study in philosophy that is concerned with the nature of existence).

A **core glossary** is a simple glossary or defining dictionary which enables definition of other concepts, especially for newcomers to a language or field of study. It contains a small working vocabulary and definitions for important or

frequently encountered concepts, usually including idioms or metaphors useful in a culture. In computer science, a core glossary is a prerequisite to a core ontology.

**Searching glossaries on the web.** The search engine Google provides a service to only search web pages belonging to a glossary therefore providing access to a kind of compound glossary of glossary entries found on the web.

A *defining vocabulary* is a published, stable, and culturally accepted core glossary specifically used by dictionary publishers to standardize their use of simple words to explain complex words, and culture-specific idioms or metaphors. It can also be published as a **defining dictionary**, but the most common use of such dictionaries is to assist in creating new dictionaries. In English, the commercial defining dictionaries typically include only one or two meanings of under 2000 words. With these, the rest of English, and even the 4000 most common English idioms and metaphors, can be defined.

An example of a useful published vocabulary is Basic English (850 words). The defining vocabulary used by Longman's to define its 4000 most common English language idioms is about 2000 words long. The English variant E-Prime is designed to avoid any judgmental statements, and so also may be useful for a neutral defining vocabulary.

### Classification of dictionaries

**Classification of dictionaries.** All dictionaries are divided into **linguistic** and **encyclopedic** dictionaries. *Encyclopedic* dictionaries describe different objects, phenomena, people and give some data about them. *Linguistic* dictionaries describe vocabulary units, their semantic structure, their origin, their usage. Words are usually given in the alphabetical order.

Dictionaries can vary widely in coverage, size, and scope. One way to classify dictionaries is based on the number of entry words they contain and give information about, i.e. their coverage. A **maximizing dictionary** lists as many words as possible from a particular speech community. An example of a maximizing dictionary (also spelled maximising dictionary) is *the Oxford English Dictionary*, as it attempts to lemmatise (i.e. show as entry words) as many words as possible. A dictionary is **minimizing** if it attempts to include only a limited selection of words from a particular speech community. An example of a minimizing dictionary (also spelled minimising dictionary) is a dictionary containing the 2,000 most frequently used words in the English language, as it attempts to lemmatise (i.e. show as entry words) only a very limited number of English words using a specific principle for their selection. (e.g., a dictionary of Basic English words).

Linguistic dictionaries are divided into general and specialized. To **general dictionaries** two most widely used dictionaries belong: **explanatory** and **translation** dictionaries. There are a lot of *explanatory* dictionaries (*NED*, *SOD*, *COD*, *NID*, N.G. Wyld's *Universal Dictionary* and others). In explanatory dictionaries the entry consists of the spelling, transcription, grammatical forms, meanings, examples, phraseology. Pronunciation is given either by means of the

International Transcription System or in British Phonetic Notation which is different in each large dictionary, e.g. /o:/ can be indicated as /aw/, /or/, /oh/, /o/. etc.

*Translation dictionaries* give words and their equivalents in the other language. There are English-Russian dictionaries by I.R. Galperin, by Y. Apresyan and others. Among general dictionaries we can also mention *Learner's dictionaries*. They began to appear in the second half of the 20-th century. The most famous is *The Advanced Learner's Dictionary* by A.S. Hornby. It is a unilingual dictionary based on COD, for advanced foreign learners and language teachers. It gives data about grammatical and lexical valency of words.

A **specialized dictionary** is a dictionary that covers a relatively restricted set of phenomena. The typical type of specialized dictionary is that which in English is often referred to as a technical dictionary and in German as a Fachwörterbuch. Specialized dictionaries can have various functions, i.e. they can help users in different types of situation. Monolingual dictionaries can help users understand and produce texts, whereas bilingual dictionaries can help users understand texts, translate texts and produce texts. **Specialized dictionaries** include dictionaries of synonyms, antonyms, collocations, word-frequency, neologisms, slang, pronouncing, etymological, phraseological and others. The distinction between a minimizing dictionary and a maximizing dictionary is also important in connection with specialized dictionaries. A law dictionary that contains 2,000 words is minimizing in that it cannot reasonably be claimed to cover more than a limited number of legal terms. This should be contrasted with a law dictionary that contains more than 20,000 entry words, which is a maximizing dictionary, as it attempts to include nearly all legal terms.

According to the *Manual of Specialised Lexicography* a **specialized dictionary** (also referred to as a technical dictionary) is a lexicon that focuses upon a specific subject field. Specialized dictionaries can be classified in various ways. Following the description in *The Bilingual LSP Dictionary* Lexicographers categorize specialized dictionaries into three types. A *multi-field dictionary* broadly covers several semantic fields (e.g., a picture dictionary), a *single-field dictionary* narrowly covers one particular subject field (e.g., law), and a *sub-field dictionary* covers a singular field (e.g., constitutional law). For example, the 23-language *Inter-Active Terminology for Europe* is a multi-field dictionary, *the American National Biography* is a single-field, and *the African American National Biography Project* is a sub-field dictionary. In terms of the distinction between "minimizing dictionaries" and "maximizing dictionaries", multi-field dictionaries tend to minimize coverage across lexical fields (for instance, *Oxford Dictionary of World Religions*) whereas single-field and sub-field dictionaries tend to maximize coverage within a limited subject field (*The Oxford Dictionary of English Etymology*).

A **multi-field dictionary** is a specialized dictionary that has been designed and compiled to cover the terms within two or more subject fields. Multi-field dictionaries should be contrasted with single-field dictionaries and sub-field

dictionaries. The typology consisting of these three dictionaries is important for a number of reasons. First of all, a multi-field dictionary is an example of the ordinary technical dictionary, covering a large number of separate subject fields, e.g. banking, economics, finance, insurance and marketing. The main problem with multi-field dictionaries is that they tend to cover one or two subjects extensively, whereas the vast majority of subject are only represented by a very limited number of terms. Secondly, the typical multi-field dictionary tends to be a minimizing dictionary, i.e. it covers only a limited number of terms within the subjects covered. Thirdly, if the lexicographers intend to make a bilingual, maximizing multi-field dictionary they run into problems with the large amount of data that has to be included in the dictionary.

A **single-field dictionary** is a specialized dictionary that has been designed and compiled to cover the terms of one particular subject field. First of all a single-field dictionary is an example of a very specialized dictionary in that it covers only one single subject field. Examples of single-field dictionaries are a dictionary of law, a dictionary of economics and a dictionary of welding. The main advantage of single-field dictionaries is that they can easily be maximizing dictionaries, i.e. attempt to cover as many terms of the subject field as possible without being a dictionary in several volumes. Consequently, single-fields dictionaries are ideal for extensive coverage of the linguistic and extra-linguistic aspects within a particular subject field. Secondly, if the lexicographers intend to make a bilingual, maximizing single-field dictionary they will not run into the same problems with the space available for presenting the large amount of data that has to be included in the dictionary, cf. a multi-field dictionary.

A **sub-field dictionary** is a specialized dictionary that has been designed and compiled to cover the terms of one (or possibly more) sub-fields of a particular subject field. It is therefore a sub-division of single-field dictionaries. First of all a sub-field dictionary is an example of a very specialized dictionary in that it covers only a limited part of one single subject field. Examples of sub-field dictionaries are a dictionary of contract law and a dictionary of fusion welding. The main advantage of sub-field dictionaries is that they can easily be maximizing dictionaries, i.e. attempt to cover as many terms of the sub-field as possible without being a dictionary in several volumes. Consequently, sub-fields dictionaries are ideal for extensive coverage of the linguistic and extra-linguistic aspects within a particular subject field. Secondly, if the lexicographers intend to make a bilingual, maximizing sub-field dictionary they will not run into the same problems with the space available for presenting the large amount of data that has to be included in the dictionary, cf. a multi-field dictionary. Consequently, the best coverage of linguistic and extra-linguistic aspects within the subject field covered by a dictionary will be found in a sub-field dictionary. The best coverage of a subject field will then be to compile a number of sub-field dictionaries that together cover the entire subject.

All types of dictionaries can be **monolingual** (excepting translation ones) if the explanation is given in the same language, **bilingual** if the explanation is given

in another language and also they can be **polylingual**. *Monolingual dictionaries* contain entries in one language and the data related to those entries are in the same language. These dictionaries can have a number of different, though interrelated functions. Monolingual dictionaries can assist users who produce texts, help users read and understand texts, and assist users who write texts.

**Monolingual learner's dictionaries** are written for learners of a foreign language. Most such dictionaries are aimed at advanced learners, but in English there are ones for elementary and intermediate users too. These tools of language education are based on the supposition that learners must move from a bilingual dictionary to a monolingual one as they advance in their study of the target language, but that general purpose dictionaries compiled for native speakers are too complex and indeed confusing for their needs. Learners' (or learner's) dictionaries include a lot of information on grammar, usage, common errors, false friends, collocations, and so on, which a native speaker knows intuitively. Conversely, these dictionaries leave out etymology and quotations, although they do include example sentences.

The first English monolingual learner's dictionary was *The Idiomatic and Syntactic Dictionary of English* by A. S. Hornby published in 1942. This was republished as *A Learner's Dictionary of Current English* by Oxford University Press in 1948. The second edition came in 1963, the third in 1974, both in several impressions. The dictionary was a huge financial success. This unparalleled success was, of course, the result of the boom in the English language teaching industry worldwide. It is now in its seventh edition as the *Oxford Advanced Learner's Dictionary*.

The *Longman Dictionary of Contemporary English* was published in 1978. The editors, led by Paul Proctor, introduced several innovations. The most striking was the use of a restricted defining vocabulary, which has now become a standard feature of learners' dictionaries. Almost a decade later another new player, the *Collins Cobuild English Dictionary*, came out, a significant milestone in corpus-based lexicography.

1995 was the 'year of the dictionaries': Oxford published its fifth edition, Longman its third, Cobuild its second, and yet another player appeared, the *Cambridge International Dictionary of English*. 2002 saw the entrance of yet another competitor: the *Macmillan English Dictionary for Advanced Learners*.

The current editions in 2008 are the seventh for OUP (2005), the fourth for Longman (2003, reprinted in 2005 with Writing Assistant), the fifth for Collins Cobuild (2006), and the third for CUP (2008). In May 2007, Macmillan released its new (second) edition of advanced learner's dictionary.

In *bilingual dictionaries*, each entry has translations of words in another language. In dictionaries between English and a language using a non-Roman script, entry words in the non-English language may be either printed and sorted in the native order, or romanized and sorted in Roman alphabetical order. Bilingual dictionaries may can have several functions: communicative functions, e.g. they can help users read and understand foreign-language texts, help users to translate texts and help users to produce texts in a foreign language. They can also have

cognitive functions, e.g. they can help users who want to know something about a foreign language in general or about a specific issue such as the inflectional paradigm of a foreign-language word.

A **bilingual dictionary** is a dictionary that is usually used to translate words or phrases from one language to another. Bilingual dictionaries are sometimes used to understand texts read, often, in a foreign language. Bilingual dictionaries can be *unidirectional*, meaning that they list the meanings of words of one language in another, or can be *bidirectional*, allowing translation to and from both languages. Bidirectional bilingual dictionaries usually consist of two sections, each listing words and phrases of one language alphabetically with their translation. Other features sometimes present in bilingual dictionaries are definitions, lists of phrases, usage and style guides, verb tables and maps.

**The most important challenge** for practical and theoretical lexicographers is to define the function(s) of a bilingual dictionary. A bilingual dictionary may have as its function to help users translate texts from one language into another, or its function may be to help users understand foreign-language texts. In such situations users will require the dictionary to contain different types of data that have been specifically selected for the function in question. If the function is understanding foreign-language texts the dictionary will contain foreign-language entry words and native-language definitions, which have been written so that they can be understood by the intended user groups. If the dictionary is intended to help translate texts, it will need to include not only equivalents but also collocations and phrases translated into the relevant target language.

Perhaps the most difficult aspect of creating a bilingual dictionary is the fact that lexemes or words cover more than one area of meaning, but these multiple meanings don't correspond to a single word in the target language. For example, in English, a **ticket** can get you into the movie theater, or can be given to you by a police officer for exceeding the posted speed limit. In Ukrainian these two meanings are not covered by one word as in English, but rather there are several options: **квиток** and **квитанція** or **талон**.

To combat the above problem, the user can perform *a reverse lookup*. In the above-mentioned example in English and Ukrainian of the word **ticket**, after finding that ticket is translated into *квиток* and *квитанція* in the English-Ukrainian dictionary, both of those Ukrainian words can be looked up in the Ukrainian-English section to help to identify which one has the meaning being sought. Reverse lookups can usually be performed faster with Dictionary programs and online dictionaries.

Bilingual dictionaries are available in a number of formats, and often include a grammar reference and usage examples.

**Visual dictionaries** – A visual dictionary is a printed dictionary that relies primarily on illustrations to provide the user with a reliable way of identifying the correct translation. Visual dictionaries are often multi-lingual rather than bilingual – instead of containing translations between two languages they often cover four or more languages.

A **picture dictionary** is a dictionary containing word entries that, for all or most such entries, are accompanied by photos or drawings illustrating what the words mean. Picture dictionaries are usually used with young children. Related to this, many picture dictionaries exist to help children learn foreign languages. There do exist, however, several foreign-language-teaching picture dictionaries that are geared towards, or also suitable for, older audiences.

Picture dictionaries explain concepts from soup-tureen in the 1904 *Engelska bild-glosor med textöfningar* ... to supersonic in the 1998 *Visual Encyclopedia*.

Another beneficial use of picture dictionaries, aside from the aforementioned, is for when one knows (or has an idea of) what something looks like, but lacks the correct term for it. For example, an adult or teenager may not be familiar with the term "platen," but wants to know what a particular part of an old typewriter is called, which happens to be (it will be learned) the platen. To find out the term, one consults a "comprehensive, image-sleuthing" picture dictionary (usually using a table of contents or index) and finds an image of a typewriter – then locates the part of the dictionary-image typewriter that one wants a word/descriptor for. If the pre-existing image *in the term-hunter's mind* reasonably verifiably matches the typewriter-image part *in the dictionary* that's labeled "platen," he or she now has a word learned (and available) for the researched item/concept – as well as having found its correct spelling.

### **Unabridged dictionaries**

General-purpose dictionaries are of two types also: (1) so-called unabridged dictionaries, and (2) desk dictionaries, which are shortened forms of the full dictionaries, either for college use or for use at lower educational levels. Desk dictionaries are the ones that we consult most of the time, in part because the unabridged dictionaries are ungainly and over-sized, in part because most of us don't have access to an unabridged dictionary at home or in our offices.

**Unabridged.** What does "unabridged" mean? First, it does not mean, as one might think, that an "unabridged dictionary" contains every English word. Nobody knows how many words English has. The blurbs on the jackets of various dictionaries may state that the dictionary contains "more than" 200,000 words, but that is difficult to determine. All one can count is "entries" or "headwords," and even that turns out to be a slippery notion because what is a headword in one dictionary may be subordinated – listed below the main entry – in another. Landau (*Dictionaries: The Art and Craft of Lexicography*, p. 84) characterizes the American system of entry counting thus:

- 1) Every word or phrase that is explicitly or implicitly defined, so long as it is clearly identifiable, usually by appearing in bold-face type, is an entry.
- 2) The more entries one has or can claim, the better.

He goes on to point out that in a particular dictionary the entry for *parachute* (*n.*) counts as five entries because the forms *parachuted*, *parachuting*, *parachute* (*v.*) and *parachutist* all appear down inside the entry. But there is surely a large difference in the "counting value" of some of these "countable" entries. Size alone,



measured by number of entries, does not make a dictionary better. In fact entry-counts are good mostly for publicity purposes. "Unabridged" means only this: the dictionary is not a shortened version of some other dictionary. It was compiled from scratch, which is to say, largely from its own files of citations, with all definitions and arrangements of meanings and examples determined by its own editors. However, dictionary producers are notorious plagiarists, and in fact have to be: every dictionary of the last 250 years has depended heavily on its predecessors, simply because the job is too big to be done really from scratch. The extremely high degree of originality of the *Oxford English Dictionary* (discussed below), the only one certainly compiled from its own files of citations, is in part due to necessity: it was the first (and still the only) dictionary ever to try to include every word that had appeared in English since the Norman Conquest, barring only technical terms that had not become common parlance. Probably the best understanding of "unabridged" is therefore something like "too big to serve easily as a desk dictionary, and having considerably more entries than desk dictionaries typically do, normally at least twice as many."

***The Oxford English Dictionary.*** The *OED*, as it is generally called (or simply *The Oxford*), is the only English dictionary compiled totally from its own citation files. Its editors, wisely, also consulted the work of their predecessors, especially Samuel Johnson. Though it excludes most technical words, it nevertheless has to be viewed as the greatest of all unabridged dictionaries – not just in English but in any language. Nothing exactly comparable to it exists for Russian, German, Spanish, French, or Italian. Its size cannot be compared with other modern dictionaries of English because it includes, in principle, all the words that have ever appeared in the English language subsequent to 1150, a date which corresponds roughly to the beginning of the Middle-English period (the period of Geoffrey Chaucer, who died in 1400). The other great modern unabridged dictionaries like the *Merriam-Webster's* have excluded older obsolete and obsolescent words, but they considerably exceed the *OED's* coverage of technical words from all the major fields of knowledge. Of the 291,627 entries in the *OED*, half or more than half are older words that no longer occur in modern usage. To say that more than half the words are no longer in contemporary use is not a criticism: the *OED* set out to create a record of the history of the English vocabulary and the historical development of the meanings of English words. It is a historical work par excellence.

The fully-up-dated second edition of 1989 is available in three formats: (1) twenty very large heavy printed volumes, which one is likely to find only in libraries; (2) a two-volume "compact edition" in which four regular printed pages of the full-sized version are reduced to one-quarter size and printed together on a single page – and a magnifying glass is provided; and (3) a compact disk, containing the whole dictionary as well as search programs which enable you to bring up onto your computer screen information which would take days to assemble from the printed versions. Unfortunately, the only one of these three versions which might be called "inexpensive" or even "moderately priced" is

the compact edition, which has on several occasions been made available at a very reasonable price as a bonus for joining one book club or another. The CD-ROM version is between \$200 and \$400, depending on which version you choose; the hard-copy version is about three times that much. A third edition in electronic form was brought out in 2005.

This great dictionary is so important to all work on the history of the English language that one should know how it came in existence. The first edition of the *OED* was compiled between 1884 and 1928; it contained about 240,000 entries. Recall, however, that this number included all the earlier as well as current words of English, so probably half the headword entries were obsolete. Furthermore, the *OED* explicitly chose not to include technical terminology from the sciences and medicine unless these terms had become common parlance outside the jargon of specialists. The policies of later dictionaries like *Merriam-Webster's* have been somewhat inconsistent on this issue, but they have generally included much more such terminology than the *OED*.

**Reduced versions of the *OED*.** The *OED* has twice been the source of highly selective reduced-size versions. The first of these is *The Shorter Oxford English Dictionary* published in 1933. It has been revised twice, once in 1944 and most recently in 1993 under the title *The New Shorter Oxford English Dictionary*. This version was released on CD-ROM in 1997.

The etymological portion of the *OED* – just the etymological portion – was the basis for the second selective version, *The Oxford Dictionary of English Etymology* (1966). This version is wonderful for etymology, and it is the right size for a desk dictionary, but in fact since it has neither extended definitions nor illustrative quotations, it is not useful as a desk dictionary and is useful even for etymological purposes only if you can't get your hands on the *OED2* CD-ROM. In 1986 Oxford published *The Concise Oxford Dictionary of English Etymology*, with a paperback reprint in 1993.

***Merriam-Webster.*** *Webster's Third New International Dictionary of the English Language*, published by the Merriam-Webster Company in 1961, is the only other relatively complete unabridged English dictionary of recent times. It differs from the *OED* in that it does contain very large numbers of technical words. It has some 450,000 entries. The fact that it is more than forty years old says something about the incredible expense and time required to update or replace a great unabridged dictionary. It replaced *Webster's New International Dictionary of 1934*, which remains the largest of all English dictionaries, having over 600,000 entries.

As of 2003, the company's two best known dictionaries are:

- *Webster's Third New International Dictionary, Unabridged*, the most complete current non-specialist American dictionary of English.
- *Merriam-Webster's Collegiate Dictionary*, Eleventh Edition, the largest and most popular college dictionary, which is available in CD-ROM format for use on personal computers.

Merriam-Webster has also published dictionaries of synonyms, English

usage, geography (*Merriam-Webster's Geographical Dictionary*), biography, proper names, medical terms, sports terms, slang, Spanish-English, and numerous others. Non-dictionary publications include, *Collegiate Thesaurus*, *Secretarial Handbook*, *Manual for Writers and Editors*, *Collegiate Encyclopedia*, *Encyclopedia of Literature*, *Encyclopedia of World Religions*.

On February 14th, 2007 Merriam-Webster announced it was working with mobile search and information provider AskMeNow to launch a mobile dictionary and thesaurus service enabling consumers to access definitions, spelling and synonyms via text message. Services also include Merriam-Webster's Word of the Day and Open Dictionary, a wiki service promising subscribers the opportunity to create and submit their own new words and definitions.

The name "Webster's," at least in America, is almost synonymous with "dictionary." One should know, however, that the name "Webster's" is in the public domain. The only publishing company whose work is directly descended from that of the nineteenth-century American lexicographical giant, Noah Webster, is the G. and C. Merriam Company of Springfield, Massachusetts. Its founders, after Webster's death in 1843, bought out the rights to the 1841 edition of Webster's *American Dictionary* (first edition 1828). But the Merriam-Webster dictionaries are not the only ones that use the Webster name to add prestige to their product. One of the best desk dictionaries with the Webster name, *Webster's New World Dictionary of the English Language* (first edition 1953) is totally unrelated to the Merriam-Webster company or to the Webster family. Another great desk dictionary (also unrelated to the earlier Webster's), the *Random House Webster's College Dictionary* (1992), simply has the name "Webster's" inserted into its earlier title, which was *The Random House College Dictionary* (1968, 1975).

**Webster's competitors.** Although the name "Webster's" has great visibility in the modern marketplace, and though the cachet of the name certainly helps to sell dictionaries in modern America, it is worth pointing out that this is due to a considerable extent to hype and mythology. Noah Webster was not the best lexicographer even of his own time, though he was the most influential one because of his *Speller* – which was the text-book of choice throughout most of the century. In his own time the best American lexicographer was probably Joseph Worcester, whose *Universal and Critical Dictionary of the English Language* appeared as the only American competitor for Webster in 1846, the final revised version in 1860. At both dates it was superior to Webster's in almost every way, but in 1864 a vastly improved version of the Webster's appeared (reworked by two scholars hired by Webster's son-in-law, and consequently known as the *Webster-Mahn* in deference to the German scholar who totally replaced the Webster etymologies). This was really the first "unabridged" Webster's dictionary, and it won the competition against Worcester in the marketplace. Near the end of the century William Dwight Whitney, a Sanskrit scholar at Yale University, produced the great *Century Dictionary*, which, in the words of Sidney Landau "is surely one of the handsomest dictionaries ever made." It was never revised, however, and is now of historical interest only. But Whitney was not the only end-of-century

competition for Webster's place in lexicography: There was also the 1893 Funk and Wagnalls unabridged *Standard Dictionary of the English Language*, revised and enlarged in 1913 as the *New Standard Dictionary*, with 450,000 entries, making it a true competitor for the unabridged Webster's. Though it was never later fully revised, and it therefore dropped out of competition, this dictionary made many important changes in dictionary practice which are continued in the various dictionaries connected with the name of Clarence Barnhart and with the dictionaries published by Random House.

**Writing dictionaries.** All modern dictionaries draw much of their historical and etymological information from the *OED*. Etymologies and definitions are based on citations. What is a **citation**? It is an index card (or, these days, a computer file) which lists a word and a quotation containing that word – if possible in a context that clearly implies a specific meaning – and gives the source, author, and date of the citation. As Landau says, "In spite of other sources [such as earlier dictionaries, either your own or your competitors]", a large ongoing citation file is essential for the preparation of any new general dictionary or for the revision of an existing dictionary." We have already mentioned the citation file of the *OED*, and a bit about how it came into existence. In America, the G. and C. Merriam Company is reputed to have the largest continuously updated and current file of citations of the words they enter into their dictionaries. Both Random House and Barnhart have independent citation files. The quality of a dictionary ultimately depends on the quality of the writing and editing.

### **Desktop dictionaries**

**For British users.** There is really only one desktop dictionary likely to be satisfactory in Britain - *The Chambers Dictionary*. This great dictionary is available in many editions, with small variations in the title. An edition called *The Chambers 21<sup>st</sup> Century Dictionary* was ambitiously published in 1997, three years in advance of the millennium bug. Its ultimate ancestor, *The Chambers 20<sup>th</sup> Century Dictionary*, first edition, came out in 1901. The 1998 edition does away with the centennial puffery and goes simply under the name *The Chambers Dictionary*. The one-page discussion (p. xx) of what American English is like (i.e., how it differs from British English) is about as useful as a comparable American one-page explanation of British English would be that was supposed to include the southern counties of Britain, the north country, Scotland, and Ireland. However, *Chambers* often records American usage in pronunciation, a favor which is not reciprocated by some American dictionaries. For instance, *schedule* is recorded by *Chambers* with the [sk-] pronunciation marked as "esp. US," but *The American Heritage Dictionary* (see below) does not record the British sh- pronunciation at all, even though it is widely favored in Canada. *Merriam-Webster's* (every modern edition), however, does record the difference.

The most conspicuous feature of *Chambers* is that all derived forms are listed within the entry under a single headword. Thus if you want to find the

computer term *descriptor*, you have to look under *describe*. If you want to find *repentance* you look under *repent*. Thus there are many fewer headwords in *Chambers* than in typical American dictionaries, though the total number of words defined in *Chambers* is actually somewhat larger than we find in any American desk dictionary. *Chambers* also has an appendix that lists common phrases and even quotations from the classical languages and modern foreign languages, and another appendix which gives the origins of many first names. *Chambers* does not give the dates when a word entered English, which is a useful feature of several American dictionaries and of the *OED*. In general, etymology is treated with minimal detail in *Chambers*.

**For American users.** At least four possible choices have to be considered.

- 1) *The American Heritage Dictionary*
- 2) *Merriam-Webster's Collegiate Dictionary*
- 3) *Random House Webster's College Dictionary*
- 4) *Webster's New World Dictionary of the American Language*

***The American Heritage Dictionary*** is an American dictionary of the English language published by Boston publisher Houghton Mifflin, the first edition of which appeared in 1969. This dictionary was innovative in two important ways:

(1) Rather than placing all the etymological information in the entry, in case the word contained a root derived from Proto-Indo-European (the parent language of most European languages) the entry provided a reference to an appendix called *Indo-European Roots*, where one can find, for every root, not only the word in question but often dozens of other words which are related by virtue of being derived from the same point of origin. Most readers found the appendix of little value because they did not know how to use it. It is unlikely ever to be valued highly by the general public.

(2) Since there had been much negative publicity about the usage labels in *Merriam-Webster's Third New International Dictionary*, the *American Heritage Dictionary* took advantage of the bad publicity to step into the breach and created a "Usage Panel" who made judgments, reported in the dictionary, about their preferences in several hundred instances of disputed usage (e.g., as between "He laid down on the bed" and "He lay down on the bed"). The panel's recommendations were sometimes too sensitive to "establishment" usage; they were often keen to protect the language from decay and corruption, metaphorically speaking. But the *Heritage* received lots of good publicity from this ploy: as a merchandising technique it was successful. As a record of actual usage, which is what dictionaries are obligated to report, it is dubious, at best, and cannot be viewed as especially authoritative.

The AHD broke ground among dictionaries by using corpus linguistics in compiling word-frequency and other information. The AHD made the innovative step of combining *prescriptive* elements (how language *should* be used) and *descriptive* information (how it actually *is* used); the latter was derived from text corpora.

The second edition, published in 1980, omitted the Indo-European

etymologies, but they were reintroduced in the third edition, published in 1992. The third edition was also a departure for the publisher because it was developed in a database, which facilitated the use of the linguistic data for other applications, such as electronic dictionaries. The fourth edition (2000) added Semitic language materials, including an analogous appendix of roots.

***The Merriam-Webster's Collegiate Dictionaries.*** Produced by The G. and C. Merriam Co. of Springfield, Mass. The latest edition is the 10th (1993). The 9th (1983) and the 8th (1973) are also excellent dictionaries, but the 7th (1963) is too old to use today. These dictionaries, depending on when they were printed, go by slightly different names, such as *Webster's New Collegiate Dictionary*, *Webster's Ninth New Collegiate Dictionary*, *Webster's Seventh New Collegiate Dictionary*. New printings with minor revisions come out almost every year, but as the dates above indicate (1963, 1973, 1983, 1993), major re-editing to produce a really new *Collegiate* takes about ten years. Several editors have been responsible for these superb dictionaries over the years, beginning with Philip Babcock Gove. The important thing to realize about all the *Collegiate* dictionaries that the G. and C. Merriam Company has produced is that they are based squarely on the citation files of the two greatest unabridged American dictionaries of this century, namely Second (1934) and Third (1961) *Webster's New International Dictionaries*, and of course all of them draw on the *OED* for etymological information and much else.

***Random House Webster's Collegiate Dictionary.*** Based on *The Random House Dictionary of the English Language*, 1966 and 1973. The latter is claimed to be an unabridged dictionary, and is the basis of the 1993 *Random House Unabridged Dictionary*. But this excellent dictionary is just too large to serve as a desk dictionary, and one is probably better served by the 1991 College version. Both for etymology and for general use, the College version is hard to improve upon.

Versions of the dictionary have been published under other names, including *Webster's New Universal Dictionary* (which was previously the name of an entirely different dictionary) and *Webster's Encyclopedic Unabridged Dictionary*.

***Webster's New World Dictionary of the American Language.*** The third edition is available in both full and college versions, like the *Heritage*. In spite of the gimmicky title (it has no special connection with Webster, and there is nothing specific to the New World or to American English about it except for the fact that it gives etymologies for American place names, a feature which is not found in other general-purpose dictionaries). The first edition was published by the World Publishing Company of Cleveland, Ohio in two volumes or one large volume, including a large encyclopedic section. In 1953, World published a one-volume college edition, without the encyclopedic material. It was edited by Joseph H. Friend and David B. Guralnik and contained 142,000 entries, said to be the largest American desk dictionary available at the time.

The second college edition, edited by Guralnik, was published in 1970. World Publishing was acquired by Simon and Schuster in 1980 and they continued the work with a third edition in 1989 edited by Victoria Neufeldt. A fourth edition

was published in 1998 and contains 160,000 entries.

One of the salient features of *Webster's New World* dictionaries has been an unusually full etymology, that is, the origin and development of words and the relationship of words to other Indo-European languages. The work also labels words which have a distinctly American origin.

## Specialized Dictionaries

**Specialized dictionaries.** The number of specialized dictionaries is vast and one cannot even judge whether a specialized dictionary is good or not unless one is a specialist in the field. There is virtually no end to specialized dictionaries – dictionaries of Old English, of Middle English, pronouncing dictionaries, reverse dictionaries, chronological dictionaries, frequency dictionaries, rhyming dictionaries, dictionaries of proverbs, dictionaries of loanwords, bibliographical dictionaries, legal terms, medical terms, music, astronomy, geography, computer terms.

**Thesaurus.** There is one type of dictionary which categorizes words only according to their semantic similarities, without regard for shared form or ancestry: this is called a thesaurus. The most famous such listing is *Roget's Thesaurus of English Words and Phrases*, first published in 1852 and in many editions subsequently. For expanding one's vocabulary, a thesaurus is likely to be even more useful than a standard dictionary, because it is arranged according to a universal set of concepts (e.g. *space, matter, intellect, abstract relations*) and then each of these is divided further and further until finally all the words can be grouped together which refer to closely similar meanings. Definitions are not given, or at least not normally very detailed definitions, just synonyms; and much of the book is an elaborate index to help you find the head entry under which all the semantically similar words of a particular category are listed.

**Dictionaries of synonyms.** Besides *Roget*, there are dictionaries of synonyms in which the headword is more or less arbitrarily chosen, and of course alphabetically listed: i.e., the editor's choice of headwords is not part of an elaborate universal classificatory system, and in the entry all the semantically similar words are listed with explanations of the distinctions among them. *Webster's New Dictionary of Synonyms* is an excellent such dictionary, as also is the Funk and Wagnall's *Modern Guide to Synonyms, A Dictionary of English Synonyms and Synonymous Expressions* by R. Soule. In 2008 *Oxford Learner's THESAURUS A dictionary of synonyms* was published. These are basically tools for writers, to help them avoid repeating the same word in different contexts (since English style has always placed a premium on variation and non-repetition). These are monolingual dictionaries. The best known bilingual dictionary of synonyms is *English Synonyms* compiled by Y. Apresyan.

**A production dictionary (activator).** A production dictionary guides you to exactly the right word you need for the context. Let's say you're writing an essay about how you spent your weekend. You had a "good" time, you went to a "good" party, the film you saw was "good" and the meal in the restaurant was "good".

You've used the word "good" four times! Wouldn't it have been better if you could have used some different words with similar and more appropriate meanings? If only you'd known how to find them.

This is when a production dictionary is needed. You could have looked up the word "good", and realised that actually you went to a "brilliant" party, the film you saw was "excellent" and the meal in the restaurant was "fantastic"! It's that easy, because when you look up a word in a production dictionary, you'll find lots of related words in the same section, so you can find exactly what you're looking for straightaway.

*The Longman Language Activator* takes you from a key word or basic idea, like 'good', and shows you more precise words or phrases with information on register, context and grammar structures. Detailed definitions help students choose the correct word and natural, corpus-based examples show words in typical usage, all the collocations and phrases you need to write correctly, index at the back of the book for easy cross-referencing. *The Longman Essential Activator* is a dictionary with a difference. It's a production dictionary.

*The Longman Essential Activator* works in the same way. The main difference is that the Longman Essential Activator is perfect for intermediate level students, while the Longman Language Activator is designed for students at a higher level. Besides, it includes **Essential word banks** (covering 25 topic areas based on a variety of exam task types, such as describing people, films/movies, education, environment, technology and computers. Within each section, you'll find all the vocabulary and phrases you'll need to be able to write correct English on that topic), **Essential communication section** (using clear flowcharts, it details all the essential phrases you'll need to communicate in various situations, such as saying goodbye, making hotel reservations and sending invitations), **Essential grammar help** (covering all the major areas that cause the biggest problems for intermediate students).

**Dictionaries of collocations.** Judging from the hugely increased use of the term in ELT-related publications, teachers today should be very much more aware of the prevalence of collocation in language use than they were twenty years ago. We are far better equipped to recognize phraseological expressions in all varieties of language, from the most spontaneous everyday conversation to the most carefully crafted literature, and easily able to find example material in journalism, letters to the press, crossword clues, broadcast interviews etc. Trained teachers also have access to a range of descriptive frameworks for analysing such phenomena, even if alternative theoretical viewpoints can be confusing and the terminology inconsistent. As a result of this greater consciousness, they should be better able to recognize the collocational problems of their learners and to answer the question posed by Allerton in 1984: 'So often the patient language-learner is told by the native speaker that a particular sentence is perfectly good English .... but that native speakers would never use it. How are we to explain such a state of affairs?'. Not only is there a substantial and growing literature on phraseological theory, there are also very many ELT textbooks that introduce and practise a wide range of



collocational patterns, and a general improvement in understanding their significance can be seen in the rubric of two vocabulary books published in 1989 and 1997: 'Some pairs of words often occur together. If you see one of them, you can expect to see the other. This makes listening and reading easier! Here are some partnerships.' (Flower and Berman 1989: 36); 'Collocation is the placing together ... of words which are often associated with each other, so that they form common patterns or combinations' (Watson 1997: 7). To find the way a word collocates, teachers can look in dictionaries of collocations such as the *LTP Dictionary of Selected Collocations* (Hill and Lewis), or the *Oxford Collocations Dictionary* (2002).

**Phraseological dictionaries** describe idioms and colloquial phrases, proverbs. Some of them have examples from literature. Some lexicographers include not only word-groups but also anomalies among words. In «*The Oxford Dictionary of English Proverbs*» each proverb is illustrated by a lot of examples, there are stylistic references as well. The dictionary by Vizetelli gives definitions and illustrations, but different meanings of polisemantic units are not given. The most famous bilingual dictionary of phraseology was compiled by A.V. Koonin. It is one of the best phraseological dictionaries.

An **idiom dictionary** explains idiosyncratic stock phrases and metaphors in language. Typical English idiom dictionaries, e.g. that published by Longmans, define about 4000 phrases, e.g. "buy the farm", "hit the road", "canary in a coal mine". Of these, a tiny subset, generally involving prepositions or action verbs, are very basic to the language, and are closely related to fundamental conceptual metaphors. These include forms like *out of* or *turn into*.

Idiom dictionaries, as well as dictionaries in general, may or may not rely on a defining vocabulary of terms (Longman's uses 2000) which are used only in their simplest senses, to minimize the number of such basic conceptual metaphors and polymorphic word uses, and make definitions easier for someone unfamiliar with the language to comprehend, such as children or students of English as an additional language. Example: "bite the bullet" "let the cat out of the bag"

**Pronouncing dictionaries** record only pronunciation. The most famous are D. Jones' *Pronouncing Dictionary*. (CUP, 2006) and J. C. Wells' *Pronunciation Dictionary* (Longman, 2008), both available in printed and CD-ROM forms.

**Dictionaries of neologisms** are: a four-volume *Supplement to NED* by Burchfield, *The Longman Register of New Words/1990/*, *Bloomsbury Dictionary of New Words /1996/*.

**Etymological dictionaries** trace present-day words to the oldest forms of these words and forms of these words in other languages. One of the best etymological dictionaries was compiled by W. Skeat. Why should one study etymology? In view of the fact that etymology often concerns itself with aspects of language that are sometimes fossilized and no longer relevant to our

ordinary synchronic understanding of what words mean or how they are used, one may legitimately ask why one should bother. The study of the etymology of words:

- ✓ enlightens us as to interesting accidents in their history;
- ✓ from a practical point of view, it gives us insights into their present meanings and into the meanings of other words which are related to the same sources, thereby expanding our vocabularies substantially and sharpening our awareness of the meanings of complex words;
- ✓ enables us to guess correctly at the meaning of a new word we have never encountered before, which happens to contain some of the parts of words we have learned.
- ✓ the most important reason is to know our language history, just as we want to know the history of our social institutions, our technology, our ancestry, our government, and so on.