

English Words: History and Structure

ROBERT STOCKWELL AND
DONKA MINKOVA

 **CAMBRIDGE**
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS

The Edinburgh Building, Cambridge CB2 2RU, UK
40 West 20th Street, New York, NY 10011-4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

<http://www.cambridge.org>

© Robert Stockwell and Donka Minkova 2001

This book is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without
the written permission of Cambridge University Press.

First published 2001
Reprinted 2002

Printed in United Kingdom at the University Press, Cambridge

Typeface Monotype Times New Roman 10.5/13pt *System* QuarkXPress™ [SE]

A catalogue record for this book is available from the British Library

Library of Congress cataloguing in publication data

Stockwell, Robert P.

English words: history and structure / by Robert P. Stockwell and Donka Minkova.
p. cm.

Includes bibliographical references and index.

ISBN 0 521 79012 3 (hardback) ISBN 0 521 79362 9 (paperback)

1. English language – Etymology. 2. English language – Morphology. 3. English
language – Word formation. 4. Vocabulary – Problems, exercises, etc. I. Minkova,
Donka, 1944– II. Title.

PE1571.S76 2001

422 – dc21 00–052896

ISBN 0 521 79012 3 hardback
ISBN 0 521 79362 9 paperback

Contents

An introduction to the textbook	1
Chapter One: Word origins	3
1 Inheritance	4
2 Neologisms (Creation <i>de novo</i>)	5
3 Blending	6
4 Acronyms	7
4.1 Initialisms	8
4.2 Reverse acronyms	9
5 Creation by shortening	10
6 Derivation	11
6.1 Derivation by affixation	11
6.2 Derivation without affixation	12
7 Compounding	12
8 Eponyms	15
8.1 Based on personal names	15
8.2 Based on geographical names	16
8.3 Based on names from literature, folklore, and mythology	16
8.4 Based on commercial brand names	17
9 Other Sources	18
Chapter Two: The background of English	19
1 The family history of English	19
1.1 Indo-European	23
1.2 The Germanic branch	26
1.3 English	28
2 Historical influences on the early vocabulary of English	30
2.1 The indigenous vocabulary of Old English	30
2.1.1 Earliest loanwords	31
2.1.2 The Scandinavian element	33
2.2 English becomes a hybrid	34
2.2.1 French loanwords in Middle English	36
Chapter Three: Composition of the Early Modern and Modern English vocabulary	39
1 The Early Modern English cultural scene	39
2 Vocabulary enrichment during the Renaissance	41
3 Transmission, etymology, source identification	46
4 Summary of early British history and loanwords in English	54

Chapter Four: Smaller than words: morphemes and types of morphemes	56
1 The smallest meaningful units	56
1.1 Morphemes and syllables	57
1.2 The properties of morphemes	58
2 Types of morphemes	61
2.1 Roots	61
2.2 Affixes	63
2.3 Functions of affixes	64
3 Compounds	66
4 Hyphens	68
5 Cognates	68
5.1 Shared derivation	70
5.2 Shared form and meaning	71
6 Finding roots in a dictionary	71
Chapter Five: Allomorphy, phonetics, and affixation	73
1 Morphological rules	73
1.1 Types of allomorphy	73
1.1.1 Zero allomorphy	73
1.1.2 Irregular allomorphy	75
1.1.3 Regular allomorphy	75
1.1.4 Derivation	75
1.2 Origins of allomorphy	76
1.2.1 Phonetic change	76
1.2.1.1 Ease of pronunciation	77
1.2.1.2 Age or time of entry of the word into English	77
1.2.1.3 Frequency of use	78
1.2.1.4 Origin	78
1.2.1.5 Transparency	79
1.2.1.6 The fossilization of allomorphy	79
2 The sounds of English	80
2.1 Phonetic notation systems	80
2.2 Phonetic symbols and square brackets	80
2.3 Consonantal parameters	81
2.3.1 Place of articulation	82
2.3.2 Manner of articulation	82
2.3.3 Voicing	83
2.3.4 English consonants: summary	84
2.4 English vowels	84
2.4.1 Vowel variation	85
2.4.2 Vowel parameters	86
2.4.3 Reduction of vowels	88
3 The affixes of English	89
3.1 Prefixes	89
3.2 Suffixes	92
Chapter Six: Replacement rules	95
1 Assimilation and types of assimilation	95
2 Labial assimilation	96
2.1 Exceptions to labial assimilation	98
3 Voicing assimilation	99
3.1 Sound versus spelling	100
3.2 Left-to-right voicing assimilation	100

4	Total assimilation	101
4.1	Total assimilation of prefixes	101
4.2	Double consonant spellings	103
5	Other replacement rules	104
5.1	T-Lenition	104
5.1.1	T-Lenition and spelling	105
5.1.2	T-Lenition, palatalization, and affrication	105
5.1.3	Summary of palatalization and affrication after T-Lenition	106
5.2	D-Lenition	106
5.2.1	Summary: palatalization and affrication of dental stops	107
5.3	v-Vocalization	108
6	Vowel replacements	108
6.1	A-Lenition	109
6.2	E-Lenition	109
6.3	Multiple Lenition	110
7	Backness assimilation	110
Chapter Seven: Deletion rules and other kinds of allomorphy		112
1	Consonant deletion	112
1.1	S-Degemination	113
1.1.1	Spelling exceptions	114
1.1.2	Exceptions to S-Degemination	114
1.1.3	Other affixes in -s	115
1.2	X-drop	116
1.2.1	Exceptions to X-drop	118
1.3	N-drop	118
1.3.1	Pronunciation and boundaries	119
1.3.2	N-drop in other prefixes	120
2	Vowel deletions	121
2.1	V-drop in hiatus	121
2.1.1	Exceptions	122
2.2	Syllable syncope	123
2.1.1	Preservation of <-er> and <-or>	124
3	Expansion rules: vowel or consonant epenthesis	125
3.1	U-Epenthesis	125
3.2	P-Epenthesis	126
Chapter Eight: Fossilized allomorphy: false cognates and other etymological pitfalls		128
1	Fossilized allomorphy	128
2	Gradation	129
2.1	Gradation in Germanic	130
3	Rhotacism	130
3.1	Rhotacism in Latin	130
3.2	Rhotacism in Germanic	132
4	Metathesis (transposition)	132
5	Obscure cognates: completely unpredictable allomorphy	133
6	False cognates	135
6.1	Boundary misplacement	135
6.2	Homophony in roots and affixes	136
6.2.1	Root homophony	137
6.3	Affix homophony	139
6.3.1	Phonetic rules and homophony	140
6.3.2	Homophony of grammatical suffixes	141

6.3.3 Mixed homophony: affixes and roots	142
7 Pseudo-affixes	143
8 Semantic variation	144
9 Multiple derivatives – multiple meanings	145
10 Multiple affixes – same meaning	146
Chapter Nine: Semantic change and semantic guesswork	147
1 Terminology	147
1.1 Diversity of meanings	148
1.2 Other -onyms	148
2 How meanings change (“semantic change”)	149
2.1 External forces	149
2.1.1 Technology and current relevance	149
2.2 Accidental associations	151
2.3 Internal forces	152
2.3.1 Analogy	152
2.4 Loss of specificity	154
3 The results of semantic change	156
3.1 Scope	156
3.2 Status: amelioration and pejoration	156
3.3 Mixed examples	157
3.4 Narrowing/specialization	158
4 Types of status change	159
5 Changing cultural relevance	160
6 Semantic guesswork	161
Chapter 10: The pronunciation of classical words in English	163
1 Unassimilated classical words	163
1.1 The pronunciation of consonants in unassimilated classical words	164
1.2 The pronunciation of vowels in unassimilated classical words	164
2 The pronunciation of fully assimilated classical words	166
3 Finding the main stress	168
3.1 The role of final syllables	169
3.2 Steps in determination of main stress placement	171
4 Stress-changing affixes	175
4.1 Affixes which attract the stress to the syllable on their left	176
Appendix I: an introduction to dictionaries	177
1 The origins of dictionaries	177
2 Types of dictionaries	177
2.1 Unabridged	178
2.2 The <i>Oxford English Dictionary</i>	179
2.2.1 The editors	180
2.2.2 Reduced versions of the <i>OED</i>	181
2.3 Merriam-Webster	182
2.4 Webster’s competitors	183
2.5 Writing dictionaries	184
2.6 Desktop dictionaries	184
2.6.1 For British users	184
2.6.2 For American users	185
2.6.2.1 The <i>American Heritage Dictionary</i>	185
2.6.2.2 The <i>Merriam Webster’s Collegiate Dictionaries</i>	186
2.6.2.3 <i>Random House Webster’s Collegiate Dictionary</i>	186

2.6.2.4	<i>Webster's New World Dictionary of the American Language</i>	187
2.6.3	Important differences between dictionaries	187
2.6.3.1	Historical order vs. logical order	187
2.6.3.2	The position of etymologies in dictionary entries	189
2.6.3.3	Dating of earliest examples	189
2.6.3.4	Specialized dictionaries	189
2.6.3.5	Thesaurus	191
2.6.3.6	Dictionaries of synonyms	191
2.7	Dictionary shelf-life	191
	Appendix II: morpheme list	193
	Index	205

1 Word origins

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

The vocabulary of English is not an unchanging list of words. New words enter the language every day, and words cease to be used. The two sources of new words are borrowing and word-creation. In fields of higher learning, like the life sciences, physical sciences, medicine, law, and the social sciences, English has usually borrowed words from other languages to get new words to cover new concepts or new material or abstract phenomena. Words referring to notions and objects specific to other cultures are often borrowed wholesale. We may borrow a word as a whole, or just its central parts (the roots). We have borrowed mainly from Latin, Greek, and French. We will leave the discussion of borrowing for later chapters. In this chapter, we will focus on the patterns of vocabulary innovation – the creation of new words – that occur within English.

We now address this topic: where do our new words originate – how do they get created – when we don't borrow them? Other than borrowing, we can count ten main sources of words in English. All but the first involve the creation of new words. These are by inheritance, by creative imagination, by blending, by joining initial letters of a phrase, by shortening, derivation, conversion, compounding, by using names as ordinary words, and by some rare echoic processes.

1 Inheritance

For the most part, the core vocabulary has been part of English for many centuries, passed down with minor changes. Much of it is shared with closely related languages like Dutch, the Scandinavian languages, and the classical languages Latin and Greek. The notion of what it means to be a closely related language is the topic of Chapter 2. For the moment the notion of relationship can be understood in a pre-scientific sense, as in “family relationship.”

The core vocabulary includes all of the common prepositions (*by, for, to, on, in, of, with, among, etc.*). They are learned well before the age of five. Similarly conjunctions like *and, but, or*. They are an essential part of the glue that holds sentences together. Other core words are the auxiliary and linking verbs (*be, is, was, were, are, am, have, can, could, may, might, will, would, shall, should, must, ought to*), and many common verbs having to do with perception and the senses (*feel, think, touch, hear, see*), and common names of body parts and kinship (*face, mouth, eyes, hand, foot, leg, mother, father, brother, sister*). If we look just at the 1,000 most common words of English, over 800 of them are of this type. Many of them can be traced back as far as language history allows us to go – about 8,000 years before the present time. Some of the others have popped up in the language during more recent times – the last two or three millennia – and in many instances their origins remain mysterious. For instance, *brunt* as in “to take the brunt of the attack,” has been in the language since 1325, but it remains of unknown origin; *blear(y)*, from the fourteenth century, origin also unknown; *duds*, as in “to wear fancy duds,” from the middle of the fifteenth century, also unknown. Closer to our times, *copacetic, posh* are from the beginning of the twentieth century; their etymology is unknown.¹ *Snazzy* is from the first Roosevelt administration starting in 1933, but no one knows its ancestry.

In addition to its core vocabulary, English has a rich supply of learned words (*learned*, in this meaning, is pronounced as two syllables). The learned vocabulary is different from the core vocabulary in that most of it is acquired through literacy and education. It tends to be associated with technical knowledge and professional skills, though there is also a large part of it which is associated with humanistic education and the literary tradition. Vocabulary enrichment in all of those areas has drawn heavily on borrowed words and roots. Most of this

¹ *Posh*, it should be pointed out, has been claimed to be a blend of “Port Out, Starboard Home,” the wealthy way to travel on a Mediterranean cruise, to avoid having the sun in your porthole. The *Oxford English Dictionary* considers this etymology to be without foundation.

book is devoted to finding out when and how such vocabulary came into English. But first we need to examine the sources of other words, words that are not part of the inherited core vocabulary and that are not borrowed from the classical languages. These are words which are created by inventive minds, and they follow a small number of patterns.

2 Neologisms (Creation *de novo*)

Though one might think it an easy matter to create a new word (without basing it on some pre-existing word or part of a word) for some new idea or new artifact, such creations are extremely rare. *Blurb* is such a word, created in 1907 to refer to the embellished descriptions on the jackets of books. *Kodak* was created by George Eastman, founder of the camera company that bears his name. Of the word itself, Eastman is reported to have said that it was “a purely arbitrary combination of letters, not derived in whole or in part from any existing word.”² *Nylon*, *Orlon*, *Dacron*, *Kevlar*, and *Teflon* are others, invented by wordsmiths within the companies that manufacture these products. Probably except for *nylon* these are not part of the core vocabulary. Even the *-on* ending of these words is obviously by analogy with words like *electron* and therefore, unlike *Kodak*, these words are not completely made up from scratch. Another word like *Kodak* is *quark*, which first appears in Joyce’s *Finnegans Wake* in the phrase “Three Quarks for Muster Mark,” taken over by physicists to mean “Any of a group of sub-atomic particles (originally three in number) conceived of as having a fractional electric charge and making up in different combinations the hadrons, but not detected in the free state” (*OED*). In the world of marketing, such creations generally are the result of massive commercial research efforts to find a combination of sounds that does not suggest something they do not want to suggest, words that have a pleasant ring to them and that are easy to pronounce. But most of the new words that even advertising experts come up with are derived from old words. For instance, the headache remedy named *Aleve* clearly is intended to suggest *alleviate*. The skin cream called *Lubriderm* is intended to suggest lubricating the *derm*, which suggests skin because of its occurrence in familiar forms like *dermatology*, *epidermis*, *dermatitis*. On the other hand another famous headache remedy, *Tylenol*, is

² The original source of this quotation appears to have been in a letter quoted in a biography of Eastman by Carl W. Ackerman (*George Eastman*, New York 1930). It was picked up by H. L. Mencken, the great Baltimore journalist whose major contributions to scholarship were his monumental studies of the distinctive words and phrases of American English (*The American Language* and two *Supplements*, New York, 1936 [4th edn., the supplements in 1945 and 1948]).

like *Kodak*, created *de novo*. *Frigidaire* is a clever coinage for a particular brand of refrigerating device. *Kleenex* is a similarly catching proprietary commercial name based on clean and the pseudo-scientific suffix *-ex*

3 Blending

Creations by blending are also called *portmanteau words*, following Lewis Carroll (Charles L. Dodgson), the author of *Through the Looking Glass*. He wrote:

Well, “slithy” means “lithe and slimy” . . . You see it’s like a portmanteau – there are two meanings packed up into one word. . . . “Mimsy” is “flimsy and miserable” (there’s another portmanteau).

Of course, to appreciate what Carroll was saying, you have to realize that portmanteau itself is a rather old-fashioned word for “suitcase,” originally designed for carrying on horseback. Other examples of blends created by him are *chortle*, from *chuckle* and *snort*; and *galumph*, from *gallop* and *triumph*. In blending, parts of two familiar words are yoked together (usually the first part of one word and the second part of the other) to produce a word which combines the meanings and sound of the old ones. Successful examples, in addition to Lewis Carroll’s whimsical literary examples above, are *smog*, a blend of *smoke* and *fog*, *motel* from *motor* and *hotel*, *heliport* from *helicopter* and *airport*, *brunch* from *breakfast* and *lunch*, *flurry* from *flutter* and *hurry*, *flush* from *flash* and *gush*. Sometimes we lose track of the components of the new blend. The origin of the word is then no longer transparent. *Vaseline* is such a word. It was based on German *wasser* “water” and Greek *elaion* “oil.” It was made up in 1872 by the man who owned the company that produced it. It is still a “proprietary term” (as *Kodak* and *Tylenol* and the other commercial terms above are), that is, it is trademarked and owned by the company that manufactures it. It is not uncommon for new technical terms to be created by blending. *Medicare*, the Social Security term covering medical care for the elderly in the United States, is now totally established, though it dates from as recently as 1965. *Medicaid* is the same sort of blend. In medical practice, a term like *urinalysis*, obviously from *urine* plus *analysis*, is so transparent in its derivation that one hardly notices that it is a separate blended word. In the field of chemistry, developing rapidly in the nineteenth century, new compounds and chemical substances required new names, which were chiefly blends: *acetal* (*acetic* and *alcohol*), *alkargen* (*alkarsin* and *oxygen*), *carborundum* (*carbon* and *corundum*), *chloral* (*chlorine* and *alcohol*), *phospham* (*phosphorus* and *ammonia*), and many

more. Blending is an area of word formation where cleverness can be rewarded by instant popularity: *sexploitation* from the seventies, the *Chunnel* from the eighties are common words now. On a lighter note, the reward can even be amusement: unpleasant as the phenomena they describe are, the words *guesstimate*, *testilying*, *pagejacking*, *spamouflage*, *compfusion*, and *explornography* will probably elicit a smile.³

4 Acronyms

Acronyms (*acr-o* “tip, point” + *onym* “name”) are a special type of blend. A typical acronym takes the first sound from each of several words and makes a new word from those initial sounds. If the resulting word is pronounced like any other word it is a true acronym. True acronyms are, for example, ASCII (pronounced [ass-key]) (American Standard Code for Information Interchange), NASA (National Aeronautics and Space Administration), WAC (Women’s Army Corps pronounced to rhyme with lack, sack, Mac), SHAPE (Supreme Headquarters Allied Powers Europe), and NATO (pronounced to rhyme with Cato) is for North Atlantic Treaty Organization. *Laser* stands for Light Amplification by Stimulated Emission of Radiation. Some of the most famous acronyms of World War II included FU-, as in FUBAR (F***ed Up Beyond All Recognition) and the GI favorite SNAFU (Situation Normal All F***ed Up). Often, however, to make an acronym pronounceable, we take not just the initial sounds but, for example, the first consonant and the first vowel together. Thus *radar* comes from *radio detecting and ranging*. *Sonar* is from *sound navigation and ranging*, where the first two letters of each of the first two words form the basis of the acronym. Few of us realize that the now very common noun *modem* was similarly formed from *modulator–demodulator*. Sometimes acronyms are based on even larger chunks of the words they abbreviate: COMECON stands for the Council for Mutual Economic Assistance – the organization of the pre-1990 East European counterpart to the Common Market. A similar formation is the name of the computer language FORTRAN (Formula Translation). These are half-way between blends and acronyms. When an acronym becomes fully accepted as a word, it often comes to be spelled with lower-case letters, like other words:

³ *Pagejacking* is an Internet scam by which web porn operators clone legitimate web pages. The last three words are from the list of new words available on the web page of the *American Dialect Society*: <http://www.americandialect.org/adsl.shtml>. They are defined as follows: *spamouflage* as “the non-spam-like header on a spam email message,” *compfusion* as “confusion over computers,” and *explornography* as “tourism in exotic and dangerous places.”

modem, *radar* came to be treated that way, as well as *okay*; and indeed in the case of *snafu* some young people may not even realize that it disguises an obscene word.

4.1 Initialisms

If the letters which make up the acronym are individually pronounced, like COD, such acronyms are called *initialisms*. America seems to have been the great breeding ground of initialisms. They are rare in English before the twentieth century (GOP and OK are early examples, both dating from the middle of the nineteenth century). TNT (trinitrotoluene) dates from just before World War I. That war produced only a smallish number of acronyms – for example WAAC (Women’s Army Auxiliary Corps) and WREN (Women’s Royal Naval [Service]). It was during the first administration of Franklin Delano Roosevelt, starting in 1933, and then during World War II, that the fashion for acronyms and initialisms really got moving. The name for American soldiers was GI’s (for General Issue), and the vehicle they drove, the Jeep, was a pronunciation of GP – General Purpose (vehicle). UFO (Unidentified Flying Object) is from the early 1950s. Roosevelt created many new government agencies, nearly all of which were referred to by initialisms (WPA Works Progress Administration, NRA National Recovery Administration, CCC Civilian Conservation Corps, FCC Federal Communications Commission, FTC Federal Trade Commission), to the point where the practice became respectable and started a trend that is now enormously productive in all areas of life. In the US, we pay taxes to the IRS (Internal Revenue Service), our driver’s licenses are issued by the DMV (Division of Motor Vehicles), we watch NBC (National Broadcasting Company), ABC (American Broadcasting Company), and CBS (Columbia Broadcasting System). It would be unfair any longer to think of the trend as American: the BBC (British Broadcasting Corporation) can be heard all over the world, the ICA (Institute of Contemporary Art) Café in London is known to locals and visitors alike, and Dubliners ride their DART, while the people in Berkeley and San Francisco ride their BART (Bay Area Rapid Transit).

In more recent times, the proliferation of initialisms and acronyms has been much aggravated by the ubiquity of computer abbreviations: e.g., HTTP Hypertext Transfer Protocol, DRAM dynamic random-access memory, CPU central processing unit, as well as further government agency naming (DOD Department of Defense, DOE Department of Energy, HEW Health, Education and Welfare). The word acronym itself came into being in 1943, near the end of FDR’s

life. Al Smith, the New York City mayor who ran for president in 1928 with FDR as his vice-presidential candidate, referred to the trend to create more and more initialisms as “making alphabet soup.” Al Smith could not have known it, but in the *Gale Dictionary of Acronyms, Initialisms & Abbreviations*, the initialism AAAAAA is recorded as the name of an organization the Mayor would have joined: The Association for the Alleviation of Asinine Abbreviations and Absurd Acronyms. (This is also an example of a reverse acronym: see below.)

4.2 Reverse acronyms

An interesting phenomenon in recent years, a sort of political offshoot of normal acronymic coinage, has been the rise of reverse acronyms – the creators start with a word they want as their name, say, for example, CORE, and then they work from those four letters to find four words which represent something like the idea they want to be associated with. CORE is the acronym for Congress of Racial Equality, NOW is the acronym of the National Organization of Women, MADD is the acronym of Mothers Against Drunk Drivers, CARE is the acronym for Cooperative for American Remittances to Europe. Organization names such as AID (Agency for International Development), AIM (American Indian Movement), HOPE (Health Opportunity for People Everywhere), PUSH (People United to Serve Humanity) have instant appeal and are easy to remember. Recently the Microsoft Corporation announced a new program which it calls DNA, for Windows Distributed interNet Architecture. It is obviously a reverse acronym in two ways: it picks up and capitalizes on a familiar acronym, namely DNA (deoxyribonucleic acid) and it has to fudge a bit to get the three letters DNA out of the actual phrase – it ignores the W of Windows, and the I of Internet. No doubt the benefits of appearing to be familiar, famous, and scientifically distinguished are worth the fudge. A rather nice case of the opposite motivation, namely to poke fun at oneself, appears in the acronym of an investment group which is called the University Park Investment Group – UPIG, naturally. A similar jest, which at the same time pokes fun at a super-secret agency of the Federal government, is to be heard in the phrase “A CYA operation.” A small hint: the first two words are “Cover Your . . .”

Another widespread recent phenomenon is acronyms based simply on some popular phrase. People can produce acronyms or initialisms from any common phrase and from just about any string of words, most of them used only within a business or a shop. A popular restaurant chain on the West Coast of the US calls itself TGIF (Thank God it's Friday), memos start with FYI (for your information), individuals

are referred to as DEWMs pronounced [DOOMs] (dead European white males).

The frequency of alphabet soup is such as to justify the production of numerous editions of the Gale dictionary, with over 400,000 entries in its eleventh edition (1987). On the other hand, alphabet soup easily and quickly disappears from the language: among the many examples above, it is a fairly good bet that not every reader knew NRA, WPA, or CCC, and those who are not into computers would have been unfamiliar with several more of the above examples. Of the 400,000 in Gale's dictionary, an ordinary person would be unlikely to know more than two or three hundred. Very large numbers of them are abbreviations for technical terms. For instance, no one but a medical expert would be likely to recognize TMJ as an initialism for temporomandibular joint.

5 Creation by shortening

Shortening may take any part of a word, usually a single syllable, and throw away the rest, like *quiz* from *inquisitive*, *phone* from *telephone*, *plane* from *airplane*, *flu* from *influenza*. Shortening is sometimes called "clipping." The process often applies not just to an existing word, but to a whole phrase. Thus *mob* is shortened from *mobile vulgus* "fickle rabble." *Zoo* is from *zoological gardens*. *Ad* and British *advert* are transparently based on *advertisement*. In many cases it is apparent that they are deliberate shortenings to save time and space in lists. Many shortenings have entered the language and speakers have lost track of where they came from. How many people would recognize *gin* as in *gin and tonic* as coming from *Genève*? Look up *whiskey* to discover what it is shortened from: the form will be completely unfamiliar to you.

Much less commonly we find what are called back formations like *edit* from *editor*, where the final *-or* is wrongly analyzed as a suffix (like the *-er* of *worker*, *employer*, *builder*) and is therefore treated as removable. To *burgle*, from *burglar*, is formed in the same way. Most examples of back formations are no longer transparent. One does not ordinarily realize, for instance, that *cherry* is a back formation from *ciris*, with the final *-s* having been wrongly analyzed as a plural suffix. The verb *grovel* is a misanalysis of *groveling*, which was originally *grufe* "face down" plus *-ling* "one who." There are not many of these, and except for very recent ones like *burgle* they are always opaque. They came into the language, after all, because the form they came from was itself opaque and open to the wrong analysis.

6 Derivation

6.1 Derivation by affixation

Up to this point, this chapter has described ways of creating new words which are not immediately transparent to the native speaker. The processes of what is called **derivational morphology** are, in many instances, so obvious that significant numbers of derivations are not even treated by dictionaries as separate entries. Since most of this book is about the complexities of derivational morphology, we do not want to anticipate details here. Roughly, derivation consists in making up new words by adding endings to more basic forms of the word. Mostly these derivations require no special definition or explanation because they follow regular rules. For example, from the *Chambers Dictionary*, under the headword *active*, we find these derived words: *activate*, *activation*, *actively*, *activeness*, *activity*, *activism*, *activist*. Four of them are given no further explanation at all, two of them are given only the very briefest explanation because the meaning has become slightly specialized, and one – *activate* – is treated at more length because it has a technical sense that requires explanation. The question is, when is a derived form merely that, predictable and comprehensible by general rules of the language, and when does the derived form require treatment as a separate word? The line is not really clear, and different decisions are found in different dictionaries. But the basic principle is this: if the new word can be fully comprehended given a knowledge of the meaning of the base and also of the endings, then it is not a new word and should not receive independent dictionary treatment, because just by knowing the parts you also know the whole. But if the new word is not transparent in that way, then it requires full definition. Examine each of these pairs of words. The members of each pair obviously have a historically based derivational relationship:

graceful	disgraceful	spectacle	spectacles
hard	hardly	late	latter
new	news	custom	customs
civic	civics	sweat	sweater

The word on the right comes from the one on the left, but the relationship is obscured because some sort of change has occurred in the meaning of the derived form (on the right) which cannot be understood by general rules of the language. Under these conditions we must then say that the derived form is a new word (in the new meaning).

6.2 Derivation without affixation

Consider the following pairs of sentences in which the same words appear in different functions (e.g., as a noun and as a verb):

This is a **major** oversight.
She graduated with a **major** in geography.
She **majored** in geography.

My **account** is overdrawn.
I can't **account** for where the money went.

They weighed **anchor** at 6:00 a.m.
Tom Brokaw **anchored** the news at 6:00 p.m.

They gave aid and **comfort** to the enemy.
They **comforted** the enemy.

We don't have any **doubt** it's correct.
We don't **doubt** that it's correct.

It's no **trouble** at all.
Don't **trouble** yourself.

In all these cases the verb or adjective and noun look alike and sound alike. There is reason to believe that the verbs are derived from the nouns. They are called “denominal verbs” for that reason, and they are said to be derived by a process of **conversion** – the noun is converted into a verb. In one sense such converted words are not new items in the lexicon. They are already there in another function (they are nouns, in these cases; but there are also adjective/adverb–verb pairs like *near*, *idle*, *clear*, *smooth*, *obscure*, and many more). The process of conversion is, furthermore, extremely productive today: we can *chair a meeting*, *air our opinions*, *panel the walls*, *weather the storm*, *storm the gates*, *e-mail the students*, *floor our enemies*, *polish the car*, *try to fish in troubled waters*, and so on. Conversions that have been around long enough are normally shown with a single entry in the dictionary, with the identification n., a., v., meaning that the form occurs as noun, adjective, and verb all three. Recent, or surprising, conversions often get separate entries in the dictionaries.

7 Compounding

This is the largest, and therefore the most important, source of new words. To produce new words by compounding, what we do is put together two words in a perfectly transparent way, and then various changes take place which cause the compound to lose its transparency.

A clear example from very early English is the word *Lord*, which is an opaque form of *loaf* “bread” (you can see the “l” and the “o” still) and *warden* “guardian” (you can see the “rd” still). A less extreme example, without the phonetic complication, is a word like *hoe-down* “noisy dance associated with harvests and weddings in the old South and West.” The *Oxford English Dictionary* gives it as the equivalent of an earlier sense of *breakdown*, now obsolete in the relevant meaning. In neither case can one infer the meaning from knowing the meaning of the constituent parts. It is therefore an opaque compound. Other examples of the “Lord” type which were once compounds and are now recognizable only as fully assimilated single words include *woman* from *wife* + *mon* (“female” + “person”), *good-bye* from *God be with you*, *holiday* from *holy day*, *bonfire* from *bone fire*, *hussy* from *house wife*, *nothing* from *no thing*.

A full description of compounds is far beyond our scope, but because it is the largest and most important source of new words in the English vocabulary, outside of borrowing, we shall try to convey some sense of the variety of words that have come into English through the process of compounding. We will not include those compounds that are now totally opaque, like *Lord* – which of course is no longer felt to be a compound at all – but will include examples of those that are transparently composed of two familiar elements that have taken on a unique new meaning that cannot be inferred totally from the meaning of the elements, like *airship* or *frogman* or *icebox* or *hovercraft*. By unique new meaning we mean that *airships* are not ships, *frogmen* are not frogs, an *icebox* is not a box made of ice, and *hovercraft* do not hover.

We begin by distinguishing between syntactic compounds and lexical compounds.⁴ One can always figure out what a syntactic compound means. Such compounds are formed by regular rules of grammar, like sentences, and they are not, therefore, listed in a dictionary. So if someone were to say,

“Playing quartets is fun.”

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

“Quartet playing is fun.”

Quartet playing is therefore a syntactic compound. Other transparent syntactic compounds are *shoemaker* (someone who makes shoes), *bookkeeper* (someone who keeps the books in order), *washing machine* (we wash things with the machine), *candlelight* (light provided by

⁴ We have drawn examples freely from a truly great piece of scholarship, *The Categories and Types of Present-Day English Word-Formation*, by Hans Marchand, 2nd edn. 1969, C. H. Beck'she: Munich.

candles), *birdcage* (a cage for birds), *playgoer* (someone who goes to plays regularly). In fact the majority of compounds we use on a daily basis are the transparent syntactic ones.

On the other hand, we cannot figure out what *ice cream* or *iced cream* means just from the rules of grammar. We cannot compute the sense of *ice cream* from something like,

They iced the cream.

Therefore *ice cream* is a lexical compound which (if we don't know the meaning already) has to be looked up in a dictionary like a totally novel word. *Crybaby* must also be treated as a lexical compound, because it refers not to babies that cry but to people who act like babies that cry, i.e., who complain when anything makes them unhappy. Similarly, *girl friend* is not just a girl who is a friend, nor is *boy friend* just a boy who is a friend. Both of these compounds actually can mean what they appear to mean on the surface, but usually they mean more than that. *Sweetheart* is not a "sweet heart," whatever that would be, but it is an opaque compound that has been in the language since the thirteenth century. *Highlight*, as in "the highlight of my day," is opaque from the seventeenth century. One can see how such a compound becomes opaque: it starts its life as a transparent description of lighting which causes some object to stand out, and then it is generalized or extended to refer to anything which stands out in one's memory or experience. As soon as this extension of the meaning is taken, then – at least in this meaning – the compound is opaque. *Bull's-eye*, which most speakers of modern English would associate with the center of a target as the primary sense, originally referred to the central protuberance formed in making a sheet of blown glass. Its earliest occurrence is a slang name for a British coin, the crown, from the beginning of the eighteenth century. The transfer of meaning to "center of a target" is simply an extension of the notion "center" which is a function of the way glass is blown, starting as a hot glob and gradually expanded outward in all directions from the center.

All of the compounds exemplified above have two parts, and their meaning is a function of the interaction of these parts plus the context of use that may gradually change them from transparent to opaque. Are there also phrasal compounds made up of more than two words? Is *maid of honor* or *good-for-nothing* or *man of the world* or *jack-of-all-trades* a phrase or a compound, and do we care? There is, unfortunately, no easy answer. Where the meaning is not obviously computable, some dictionaries list them as lexical compounds: e.g., the *Oxford English Dictionary* does not list *jack-of-all-trades*, but the much smaller *Webster's Collegiate* does. *Maid of honor* is listed by both, whereas *good-for-nothing* is not listed by any, nor is *man of the world*,

though in both these instances there would seem to be good reason to single them out as having special properties: one can know about men and about the world without knowing what man of the world really means, and good-for-nothing refers to a special kind of worthlessness, usually laziness.

8 Eponyms

These are new words based on names (*epi-* “upon” *onym* “name”). All eponyms necessarily involve some degree of change in the meaning of the word: *watt*, for example, refers to a unit of electrical power, not to the individual who invented the steam engine. The number of new words of this type in fields like biology, physics, and medicine is very large, since new discoveries are very often named for their discoverers. Quite often we take the name of an individual, a character familiar from mythology, history, or folklore, a place name, a brand name, and so on and extend its scope beyond the original individual reference, thereby turning what is called a proper noun, i.e. somebody’s name, into a common noun, i.e. a word like boy, girl, doctor, house, town that does not refer to a particular individual but to a class of individuals sharing relevant defining properties. Even proper nouns, of course, can be of several types: those which are associated with real people, those that are associated with imaginary creatures or mythological figures, those that are associated with places. All three types have provided words in English based on their names. Some examples:

8.1 Based on personal names

boycott (Charles Boycott, an English land agent in Ireland)

dahlia (developed by Anders Dahl, a Swedish botanist)

cardigan (Earl of Cardigan, nineteenth century; a style of waistcoat that he favored)

derrick (the name of a hangman at a London prison in the time of Shakespeare and Queen Elizabeth I)

guy (In Britain, Guy Fawkes Day, November 5; for the Catholic conspirator, member of the Gunpowder Plot in Great Britain, 1606. Since he was held up to ridicule, and in Britain the word still means “a person of odd or grotesque appearance,” it is apparent that American English has generalized and neutralized the word.)

lynch (Capt. William Lynch, a planter in colonial Virginia, originated lynch law in 1780)

nicotine (Jacques Nicot introduced tobacco into France in 1560)

ohm (unit of electrical resistance, named for nineteenth-century German physicist, Georg Simon Ohm)

sadistic (eighteenth-century Marquis de Sade, infamous for crimes of sexual perversion)

sandwich (eighteenth-century British nobleman, the Earl of Sandwich, who brought bread and meat together to the gambling table to provide sustenance for himself, and started the fast food industry)

8.2 Based on geographical names

bikini (the islands where the atom bomb was tested; presumably gets its meaning from the style of female native costumes encountered there)

cheddar (a village in Somerset whence the cheese first came)

china (short for chinaware, from china-clay, employed in the manufacture of porcelain, originally made in China)

denim (cotton cloth now, originally serge, made in the town of Nîmes, southern France, hence serge de Nim)

hamburger (the word is an Americanism; from Hamburg steak, some form of pounded beef, found in Hamburg in the nineteenth century and brought to the US by German immigrants, though the word and specific concept of the hamburger originated in the US)

jean (from the Italian city of Genoa, where the cloth was first made, as in blue jeans)

port (shortened from Oporto, the chief port for exporting wines from Portugal)

sardonic (should be sardinic, coming from the island of Sardinia; the vowel change is based on the Greek form; refers to a type of sarcastic laughter supposed to resemble the grotesque effects of eating a certain Sardinian plant)

sherry (a white wine from, originally, Xeres, now Jerez de la Frontera, in Spain; the final <s> was deleted on the mistaken view that it was the plural suffix, an instance of what is known as morphological reanalysis)

spartan (from the ancient Doric state of Laconia, in the south of Greece; the meaning comes from their chosen lifestyle, which eschewed luxuries)

turkey (an American bird, confused in America at first with an African Guinea-bird, brought into Europe through Turkey, whence the name: but certainly a confusing sequence of borrowing and renaming!!!)

8.3 Based on names from literature, folklore, and mythology

atlas (he was condemned by Zeus, the leader of the Greek gods [called Jupiter by the Romans], to support the earth on his shoulders; the

name was assigned by an imaginative early anatomist to the top vertebra of the neck, the one which supports the head; it came to refer to a collection of maps because many early publications of world geography showed drawings of Atlas holding the world up on his shoulders)

casanova (Giovanni Jacopo Casanova de Seingalt. He wrote vividly about his sexual adventures throughout most of Europe)

chimera (a mythological Greek monster, purely a creature of the imagination)

morphine (Morpheus was the son of the Greek god of sleep)

nemesis (after the name of a Greek goddess who punished violations of all forms of rightful order and proper behavior)

panic (noises which caused fear in the flocks by night were attributed in ancient Greece to Pan, who was the God of misdeeds; a panic is irrational behavior in the herd)

platonic (Plato was an early Greek philosopher; the word originally referred to the kind of interest in young men that Socrates, the first great Greek philosopher, is supposed to have had. As originally used, it had no reference to women, though now its main reference is to a non-sexual relationship between men and women)

saturnine (as the *OED* says, “sluggish, cold, and gloomy in temperament”; one wonders why a car should be named after it. Presumably the sense of saturnine is based on the fact that Saturn was the most remote of the seven planets known to ancient astronomers)

satirical (a satyr was a creature with a mixture of human and animal properties, and supposed to be gifted with a prodigious sexual appetite; the word satire refers originally to theatrical pieces which hold these qualities, and others, up to ridicule)

8.4 Based on commercial brand names

Band-aid[®] is commonly generalized to refer to any small bandage for a cut or scratch, and it has moved out into general use in metaphors like “The IRS needs major reforms; we’ve had enough of these taxation band-aids!”

Jello[®] a particular brand of jellied emulsion, is generalized to refer to any edible substance of the same type.

Levis[®] a brand of canvas trousers, now refers to any denim-like, rough and ready, trousers.

Tampax[®] is one of many brands of feminine hygiene devices, generalized to them all.

Xerox[®] especially as a verb (“to xerox something”), has come to mean “to copy by any dry process.”

Zipper[®], based on the echoic word (see below) *zip*, which imitates the sound of speeding objects. The verb is from 1852, the noun 1926.

9 Other sources

It is part of the common mythology about language that many words must have come from efforts to imitate the sounds that the words represent. There are in fact only a few legitimate instances of this sort, and they are called **echoic** words. Bloomfield⁵ distinguished between those words that are actually imitative, like *oh!*, *ah!*, *ouch!*, those that are coined to sound like a noise made by some object or creature, such as *bang*, *blah*, *buzz*, *burp*, *splash*, *tinkle*, *ping*, *cock-a-doodle-doo*, *meow*, *moo*, *baa*, *cuckoo*, *bob-white*, *whip-poor-will*, and those that have the property that “to the speaker it seems as if the sounds were especially suited to the meaning.” His examples are *flip*, *flap*, *flop*, *flitter*, *flimmer*, *flicker*, *flutter*, *flash*, *flush*, *flare*, *glare*, *glitter*, *flow*, *gloat*, *glimmer*, *bang*, *bump*, *lump*, *thump*, *thwack*, *whack*, *sniff*, *sniffle*, *snuff*, *sizzle*, *wheeze*. The total number of any of these types of words that may be called roughly echoic is very small, in English or any other language. It is not a major resource for expanding the vocabulary.

Another rather unimportant, though often amusing, resource for expanding the vocabulary is through a process called **reduplication**, in which part or all of a word is repeated.⁶ Only a few of these examples are more than trivial expansions of the vocabulary: *dum-dum* (type of bullet), *bonbon*, *tom-tom*, *fifty-fifty*, *hula-hula*, *so-so*, *boob tube*, *brain drain*.

So much for the ways of introducing new words into English without borrowing them. Since well over 80 percent of the total vocabulary of English is borrowed, we turn now to the rest of the book to study many aspects of the history of borrowed words in English.

⁵ Leonard Bloomfield, *Language*, New York: Henry Holt, 1933: 156.

⁶ A recent study, from which our examples are taken, is John M. Dienhart, “Stress in Reduplicative Compounds: Mish-Mash or Hocus-Pocus,” *American Speech* 74.1: 3–37 (1999).