1

HISTORY IN THE LANGUAGE: THE VOCABULARY AS A HISTORICAL REPOSITORY

Here, therefore, is the first distemper [imbalance] of learning, when men study words and not matter.

Francis Bacon, *The Advancement of Learning* (1605)

It has only just begun to dawn on us that in our own language alone, not to speak of its many companions, the past history of humanity is spread out in an imperishable map, just as the history of the mineral earth lies embedded in the layers of its outer crust.

Owen Barfield, *History in English Words* (1926)

Today we find Bacon's downright rejection of the study of language extraordinary, coming as it does from one of the greatest minds of the Renaissance. Barfield's archaeological model of language is entirely in tune with our way of thinking, despite being written over seventy years ago. We find recent echoes in George Steiner's dictum that 'History in the human sense, is a language net cast backwards' (1975, p. 70) and in Roger Fowler's observation that 'Quite simply, the vocabulary of a language, or of a variety of a language, is a map of the objects, concepts, processes and relationships about which the culture needs to communicate' (1991, p. 80). Has this lexical 'map' changed, and if so how? In exploring the evolution of the English vocabulary, we shall find that words are both fossils in which the culture of the past is stored and vital organisms responsive to the pressures of the present.

Words surround us in their myriad multiplicity, the common and the rare, the local and the alien, the ancient and the new, the philosophical and the technical, the private and the political, the sacred and the profane. Where have they all come from? How have they arrived in these categories? How does one analyse, how make sense of this lexicon, so vast, eccentric and copious, which at the last count amounted to more than half a million words?

A valuable starting-point was achieved by the great pioneering editor of the monumental *Oxford English Dictionary on Historical Principles*, the indefatigable and scholarly Sir James Murray, who articulated the problem with admirable clarity in his preface a little over a hundred years ago. Introducing this huge work of historical reconstruction, Murray used the memorable image of the vocabulary being like a galaxy:

> That vast aggregate of words and phrases which constitutes the Vocabulary of English-speaking men presents, to the mind that endeavours to grasp it as a definite whole, the aspect of one of those nebulous masses familiar to the astronomer, in which a clear and unmistakable nucleus shades off on all sides, through zones of decreasing brightness, to a dim marginal film that seems to end nowhere, but lose itself imperceptibly in the surrounding darkness. (1884, p. xvii)

Murray's astronomical image is arresting, but also daunting. It gives us the big picture, but we seem to be lost in it. Given that lexis is generally taken to mean the aggregate of words, phrases, idioms and meaningful units in the vocabulary, how do we divide up this 'nebulous mass'? The dictionary format, being based on the arbitrary sequence of the alphabet, not on the logical connections of associated meaning, encourages us to consider words as atoms or individuals, like disparate people in the telephone directory, with their special place, function and use. This format also imposes on words a misleading equality, since every word, no matter how central, rare or insignificant, has its entry. The comprehensive dictionary also complicates things by bringing in the historical dimension, showing that individual words have biographies or semantic histories which are long and complicated, and sometimes quite bizarre. But, as with individuals, words also belong to families and nations, though their family resemblances and national affiliations may in time become obscured by assimilation.

Developing his image, Murray starts to focus on what distinguishes the core of the galaxy from the perimeter:

> the English Vocabulary contains a nucleus or central mass of many thousand words whose 'Anglicity' ['Englishness'] is unquestioned . . . they are the *Common Words* of the language. But they are linked on every side with other words which are less and less entitled to this appellation, and which pertain ever more and more distinctly to the domain of local dialect, of the slang and cant of 'sets' and classes, of the peculiar technicalities of trade and processes, of the scientific terminology common to all civilized nations, of the actual languages of other lands and peoples. And there is absolutely no defining line in any direction: the circle of the English language has a well-defined centre but no discernible circumference. (1884, p. xvii)

Figure 1.1   Murray's lexical configuration

Murray illustrated the structural relationship between the core and the diversity of other words by means of a diagram (figure 1.1).

Murray's diagram or plan of the constellation is profoundly valuable and illuminating. (It was also an innovation in its time.) It gives us a framework within which to categorize different kinds of words, and it makes the point with simple clarity (by means of the arrows) that words are 'in' the language to differing degrees. It also establishes a fundamental hierarchy of usage, with 'Literary', 'Common', 'Colloquial' and 'Slang' categorized in descending order. In discussing it we shall develop the main lexical concepts and categories and also introduce the principal periods in the time-frame of the development of English.

Register

Murray's diagram or word-map depicts in essence an important linguistic concept, that of register, namely a particular variety of diction or choice of word appropriate to a given social situation or literary context. The term **register** is traditionally associated with music and is comparatively recent in

its application to linguistic usage (being first recorded in this sense by T. B. W. Reid in 1956). However, the core concept of a particular choice of diction is deeply traditional. We are familiar from the earliest stages of our language acquisition of the difference between the personal terms *pa* or *dad* as against the more formal *father*, or *kitty* and *pussy* as against *cat*, and so on. Register is based on the classical notion of decorum, whereby certain levels of usage are considered appropriate (or inappropriate) to particular topics and social situations. Register is typically conceived of (as it is in Murray's configuration) on a hierarchical scale, ranging from high words to low, so that in literature a high action properly requires elevated language (typically producing the genre of the epic) while a low action is more appropriately couched in humbler or even obscene language (typically apparent in the farce or 'dirty story').

The main emphasis is traditionally put on differing degrees of formality (for example, *chuck* as against *throw* and *row* as against *disagreement* or *confrontation*). A useful set of criteria which governs the choice is explained in these observations:

> This degree of formality/informality depends upon four variable factors, in increasing order of importance: sex, age, status and intimacy . . . Students talking among themselves would use a different type of vocabulary and even different grammatical structures from those they would use in addressing their teacher, or when being interviewed for a job, or when talking to a young child – or a dog. (Batchelor and Offord, 1982, pp. 1, 3)

This gives an important different emphasis, namely that register reflects the role of a speaker in a particular context. Words may be 'out there' in large numbers, but there is not a completely free choice between them. On this point Shakespeare's friend, the dramatist Ben Jonson, observed in his *Discoveries* (1641) that words

> are to be chose[n] according to the persons wee make speake, or the things wee speake of. Some are of the Campe, some of the Councell-board, some of the Shop, some of the Sheep-coat, some of the Barre, &c. (1923, p. 73)

However, register is rich and variegated in its manifestations. It can also be demonstrated in a variety of alternatives in word-choice: old or new; concrete or abstract; blunt or polite; coarse or refined; direct or euphemistic; common or recherché. The motives behind the choice of different registers can be equally various, such as social, literary, professional, commercial and political.

The noted scholar M. A. K. Halliday refined the concept further by distinguishing between the language variety of the user (termed *dialects*) and the
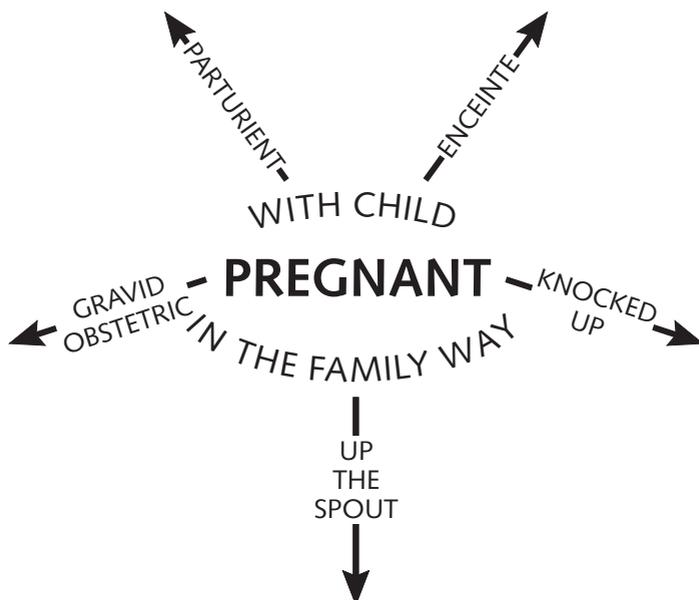
Figure 1.2 The semantic field of pregnancy

variety required by the situation (termed *registers*). He further distinguished between registers according to *field* (subject matter), *mode* (speech, writing, format) and *manner* (the social relations we have just been discussing). Of these categories, *field* is perhaps the most important, since it introduces a diversity of word-fields, for example, those of morality, law, business, politics and sport. These all have their various vocabularies, but a common word like *right* will be found in the first four categories, albeit in different senses, *goal* will be found in the last three, and so on.

We shall be returning to the central concept of register, but let us now illustrate it by applying Murray's diagram to a particular notion. In so doing we shall be constructing a semantic field, or a set of synonyms which applies to a particular notion. We shall start with a slighty unusual topic, that of being pregnant, since it contains a few highly diverse terms, which are set out in figure 1.2.

Most of us would find it difficult to construct a complete semantic field on the basis of our own knowledge. We shall gain useful assistance from a major work built on this principle, namely Roget's *Thesaurus*, first published in 1852 and updated in many editions since. Roget's enterprise was revolutionary in that it worked from concepts to words and phrases, collected into semantic

fields. Roget was truly the pioneer in English **lexicology**, meaning 'the study of the structure of the vocabulary', a term coined in 1828 by Noah Webster. Murray, on the other hand was part of the great tradition of **lexicography**, the writing of dictionaries, which goes back to the late sixteenth century.

When we distribute or allocate the various terms and phrases referring to pregnancy, we discover that there is little difficulty in matching Murray's categories, since the field is fairly rich and varied. We have, indeed, other terms which might be used: *expecting* and *expectant*, for example, would be placed slightly above *pregnant*, while *having a bun in the oven* would be lower down the scale, between *in the family way* and the coarse phrase *up the spout*.

This exercise is valuable in various ways. First, it shows us that although **synonym** is a useful term for practical purposes, there are, in fact, few exact synonyms, especially in this sensitive area. Indeed, most authorities agree that, strictly speaking, there are no exact synonyms in the sense of terms which are semantically interchangeable in all contexts. See, in this respect, David Crystal (1995, p. 164), John Lyons (1968, pp. 447–8) and Leonhard Lipka (1990, p. 142). Rather, in this field there are marked differences of directness and nuance: some words are blunt, some more discreet, others slightly mystifying. *Enceinte*, for example, from French, where it means literally 'unbound', is quite rarefied and delicate. *Parturient* and *obstetric*, both from Latin, are technical and used by specialists, but are unfamiliar to the layman. As is generally the case with technical terms, they are quite specific and cannot be used in any other sense, as *pregnant* can be in, say, the phrase *a pregnant pause*.

Secondly, the exercise makes us focus in greater detail on the categories. What is the real difference between 'scientific' and 'technical'? Where, for example, would we place *impregnated*? It is more of a 'written' word than a 'spoken' word, so on that basis it would go under 'scientific'. Furthermore, what is really meant by 'literary'? The example chosen here is the archaic biblical phrase *with child*. Many languages have areas of the lexis which are chiefly literary: as we shall see, a considerable portion of the Anglo-Saxon vocabulary was exclusively poetic. This is still true of a small sample of Modern English: witness the use of *isle* (as against *island*), *mount* (as against *mountain*), *weeds* (as against *clothes*), *dulcet* (as against *sweet*), *fount* (as against *fountain*), *steed* (as against *horse*) and *serpent* (as against *snake*). As these examples show, there is commonly an overlap between 'literary' and 'archaic', a category which, curiously, does not figure in Murray's configuration: words such as *henceforth*, *foe*, *sans*, *multitude* and *damsel* are examples. In Murray's time the euphemistic phrase *in an interesting condition* was current and would then apply: now it seems old-fashioned and prissy.

Writing a hundred years ago, Murray did not find it necessary to define 'literary' language. It meant, self-evidently, the language used by the authors who

made up the 'canon' of English literature, including the great Victorian novelists and poets of his own time. While this was a largely unproblematic concept, some might have pointed out that there are difficulties of obscenity in Chaucer and Rochester, and of **idiolect** or individual dialect and features of personal expression in Dickens. But these objections were not raised. Today 'literary' definitely is a problematic concept, in that many books are 'literary' in the traditional sense of being 'well written' in a formal fashion, while others are classed as literature but use a great deal of slang and obscenity. We shall return to this topic in our concluding chapter.

Finally, when we come to 'dialectal', we can see that this category is illuminating because it depends on where the centre of the word-field is conceived as being. This in turn depends on the definition of the speech-community. Murray would probably have confined the field to the British Isles, so that 'dialectal' would have yielded terms like *wi bairn* or *boukun*, the Scots equivalents of 'with child' or 'pregnant'. (These are found in *The New Testament in Scots*, translated by W. L. Lorimer at St Matthew 1:5 and St Luke 2:5.) The phrase chosen at present for the category (*knocked up*) is, however, an American colloquialism.

Dialect in its modern academic usage includes both the traditional meaning of a regional form of speech as well as a class usage. In this word-field the form *preggers* could be used, since it is a distinctively upper-class usage, marked by its suffix, as is *champers* for 'champagne'. *Dialect* as a category has also risen in status since Murray's time, when it implied 'back-woods', 'rustic' or 'outlandish' speech. The first comprehensive study of English dialects was carried out by Joseph Wright about a century ago and published in six volumes between 1898 and 1905. Today, with regional forms of speech under threat, they are regarded with more sympathy and respect. However, because of these traditional negative connotations, *dialect* is often replaced by the more neutral term *variety*. Nevertheless, a comment in a recent study on English dialects is noteworthy:

> The subject which Wright did so much to make popular and academically respectable, now has followers studying, for example, 'traditional' regional dialects . . . the dialects of the cities, the dialects of ethnic minorities, occupational dialects and the relationship between dialect and social class or gender. (Upton and Widdowson, 1995, p. ix)

In leaving 'dialect' as a category, it is important to stress that the distinction between spoken and written varieties carries many implications. Although we can state that *pregnant* is the central or general term in the semantic field and *boukun* is a 'dialect' word, this is a perception which comes largely from the

Figure 1.3 Dialect terms for gym-shoes (from Peter Trudgill, *The Dialects of England*, Blackwell, 1990, p. 102)

written language. In oral usage there are many cases where a whole range of regional forms are the norm for speakers on the ground. Thus a basic word like *grandfather* has the regional variants of *granda*, *granfer* and *gramp*, while *scarecrow* has the surprising range of *flay-crow* in the north, *mawkin* in the Midlands and East Anglia, *gally-bagger* in Hampshire and *mommet* (from Mohamet) in Somerset and adjoining counties. Figure 1.3 shows a similar regional range of

words for gym-shoes. A sense of the extraordinary diversity of dialect words, some of which we would not regard as being obviously English, is shown in figure 1.4, a page from Joseph Wright's *English Dialect Dictionary*.

Before we leave Murray's diagram, we should note that a slightly modified version was used in the preface to the *Shorter Oxford English Dictionary* (*SOED*), first published in 1933 (p. viii). The two substantial changes were that 'archaic' was now admitted as one of the sub-categories of 'literary' usage, and 'vulgar' was introduced as part of 'colloquial'. Also noteworthy is the point that 'slang' had a meaning for both Murray and the editors of the *SOED* which differs from ours. Murray, we recall, referred to the 'slang and cant of "sets" and classes', and elaborated: ' "Slang" touches the technical terminology of trades and occupations, as in "nautical slang", "Public School Slang", "the slang of the Stock Exchange".' (Eric Partridge's excellent study *Slang* (1933 and later editions) uses similar categories.) In other words 'slang' was then closer to what we would now call 'jargon', as well as to other 'in-group' languages, categorized by the *SOED* as 'of lower or less dignified status'. We can thus see that some categories which were clearly defined and stable for Murray and his immediate successors subsequently turn out to be fairly fluid.

The word-field of pregnancy is a fruitful example, since it consists of a few, clearly differentiated terms from a diversity of origins. Other word-fields show a different balance of registers and concentration of terms. Those for madness, for example, set out in figure 1.5 show a remarkable range and size. That for drunkenness, like that for money, shows a distinctly 'bottom heavy' imbalance, with hundreds of slang terms.[1] A discussion of 'Words of War' is to be found in chapter 8. Other fields, such as those covering economics, politics and xenophobia, are contained in my *Words in Time*, while euphemisms and the categorization of women and homosexuals are discussed in my *Swearing*.

We shall be returning at intervals to Murray's constellation. It is a most valuable word-map and a remarkable innovation, considering that it was devised over a hundred years ago. We must now return to the core of the galaxy, the 'well-defined centre' of common words.

The Three Word-Stocks

As we have already gathered from our word-fields, the English vocabulary does not originate in one language, but is a fascinatingly hybrid conglomeration, as we shall see in the course of this book. In essence and in detail, the structure of the vocabulary reflects the history of the English-speaking peoples. But even fairly common words do not all come from the same source. This point is exemplified even in the words which make up the title of this book. They

YICKIE-YAWKIE, *sb.* Sc. Also in form **yeekie-yakie**. A wooden tool, blunted like a wedge, with which shoemakers polish the edges and bottoms of soles.

Dmf. (JAM.) Gall. MACTAGGART *Encycl.* (1824). Kcb. Tam . . . gathered up his knife, ' yeekie yakie,' lap stone and rags, ARM-STRONG *Kirkiebrae* (1896) 12.

YIELD, *sb.*[1] Ags. (JAM.) See below.

When the ice melts, although there be no proper thaw, it is said to be owing to the yield of the day.

YIELD, *v.* and *sb.*[2] Sc. Nhb. Yks. Chs. Brks. Suf. Ken. Sus. Also in forms eeld Nhb.[1]; yeild Chs. Sus.[1]; yeld Ken.[1]; yelt Suf.[1]; *pret.* yald Sc.; *pp.* youden Sc. (JAM.) [jïld.] **1.** *v.* To reward.

w.Yks. THORESBY *Lett.* (1703). Chs. God yeild you (K.); Chs.[123] **2.** To produce a crop ; to be fruitful. See Eald, Ild.

Sc. (A.W.) n.Yks. T'cooarn yields weel(I.W.). Chs.[1] Thus we speak of a good crop of wheat as ' yielding well,' or peas which have many seeds in a pod as ' yielding well ' ; Chs.[3] Suf.[1] Ta yelt a matter a[1] tew coom an acre. Ken.[1] 'Tis a very good yelding field though it is so cledgy.

3. Of a cow : to give milk.

Sc. (A.W.) Nhb.[1] Hoo much is the coo eeldin? w.Yks. (J.W.) **4.** To give up ; to relinquish. Sc. (JAM.), s.Cy. (HALL.), Sus.[1] **5.** To give way.

Fif. The kirk-yard's coffins yald and broke Aneath the press o' livin' fock, TENNANT *Papistry* (1827) 168. Abd. When the effects of a thaw begin to be felt, it is common to say ' the ice is yowden,' i.e. it has begun to give way (JAM.). **6.** To admit, confess.

Rnf. Priests may preach and scribes may jaw—And sodgers shoot . . . Or [ere] ye wad shrink or yield a flaw, WEBSTER *Rhymes* (1835) 8. **7.** In *comb.* Yield-yow, a violent pressure of the thumb under the lobe of the ear. S. & Ork.[1] **8.** *sb.* Produce.

Brks.[1] Whate maaykes poor yield this crap.

[1. God yelde thee, freend, CHAUCER *Tr. & Cr.* I. 1055.] **YIELD**, see Geld, *adj.*, Yeld.

YIELDY, *adj.* Sc. [jï·ldi.] Yielding, giving way.

Gall. Time's shingly sands I see Turn dry an' yieldy 'neath my feet, SCOTT *Gleanings* (1881) 31.

YIFFER, see Yewfir.

YIFF-YAFF, *sb.* *Obs.* Rxb. (JAM.) A small person who talks a great deal to little purpose.

YIGGA, *sb.* Cum. [ji·gə.] A dial. form of 'ague.' Cum.[1] (s.v. Ayga), Cum.[4]

YIGLET, **YIK**, see Aglet, Oak.

YIKE, *sb.* Sur. Hmp. [jaik.] The call of the wood-pecker.

Sur. Here you could listen to the yike, yike, yike of the green woodpecker, and watch him at his work, SON OF MARSHES *On Sur. Hills* (1891) 136 ; Then he [the yaffle] yells his loudest, making the woods ring with his maniacal yikes, *ib. London Town* (ed. 1894) 193.

Hence **Yikeing**, *ppl. adj.* making the sound of ' yike.'

Sur. The yikeing laugh of the yaffle, *Forest Tithes* (1893) 30. Hmp. *Cornh. Mag.* (June 1893) 595.

YIKKA, *v.* Sh.I. [ji·kə.] To snarl.

He yikkas an growls, BURGESS *Rasmie* (1892) 17.

YILL, *sb.* and *v.* Sc. n.Cy. Also in forms yell N.Cy.[1]; yuill Ayr. [jil.] **1.** *sb.* A dial. form of ' ale.'

Sc. Send down for bread and for yill, tobacco, &c., SCOTT *Bride of Lam.* (1819) xxiii. Frf. [She] loot them pree her yill an' kebbuck, WATT *Poet. Sketches* (1880) 108. w.Sc., s.Sc. (JAM.) Ayr. A howff . . . whaur you get your yuill oot of fine auld siller communion cups, SERVICE *Dr. Duguid* (ed. 1887) 108. N.Cy.[1]

Hence **Heather-yill**, *sb.*, see below.

Wgt. They say that the Picts brewed some awful grand kind of drink they ca't Heather Yill, out of heather and some unknown kind of fogg ; but they kept the secret of making of it to themselves, and it is now lost, SAXON *Gall. Gossip* (1878) 295.

2. *Comb.* (1) Yill-boat, an ale-barrel ; see Boat, *sb.* 3 ; (2) -cap or -cup, a horn or wooden vessel from which ale is drunk ; (3) -cup eye, a large or ' saucer ' eye ; (4) -house, a house where ale is sold ; (5) -seasoned, seasoned with ale ; (6) -seller, a person who sells ale ; (7) -shop, see (4) ; (8) -wife, a woman who brews or sells ale.

(1) Bwk. (JAM.) (2) Sc. (JAM.) Ayr. The Change-house fills,

Wi' yill-caup commentators, BURNS *Holy Fair* (1785) st. 18. Gall. MACTAGGART *Encycl.* (1824). (3) Gall. (JAM.) Kcb. Chiels wi' sooty skins, and yill-caup een, DAVIDSON *Seasons* (1789) 13 (*ib.*). (4) Sc. (JAM.) ; I never gang to the yillhouse—that is unless ony neighbour was to gie me a pint, SCOTT *Rob Roy* (1817) xiv. Edb. The coarsest night that cou'd hae blawn, I at the yill-house door bid staun', A' shiverin', R. WILSON *Poems* (1822) 35. N.Cy.[1] (5) Ayr. Yill-season'd haivers Are no worth a plack, WHITE *Jottings* (1879) 290. (6) Ayr. Yuill-sellers shouldna be story tellers, ye ken, SERVICE *Dr. Duguid* (ed. 1887) 235. (7) Lnk. Oot an' intil yill-shops they gaed, COCHILL *Poems* (1890) 77. (8) Sc. Cld. (JAM.) Ayr. Ye're welcome, neighbour yill wives. here, FISHER *Poems* (1790) 59. Edb. Yill-wives licker brisk decantet for drinkers' food, CRAWFORD *Poems* (1798) 46. Gall. MACTAGGART *Encycl.* (1824) 166, ed. 1876. N.Cy.[1]

3. *v.* To treat to ale.

w.Sc. (JAM.) Ayr. Langsyne it was the fashion . . . for lads tae tak their lasses intae the public hoose on the Sabbath day, at twull-oors, tae get a bake an' a hue o' porter, or like tat ye ken, an' that was ca'd ' yuillin',' SERVICE *Dr. Duguid* (ed. 1887) 206.

YILL, **YILLOCH**, see Yeld, Yelloch.

YILLYART, *adj.* Sc. Stubborn ; ill-conditioned ; 'thrawn.'

Per. Then ye can get, ye yillyart tyke, Home rule or ony rule ye like, FERGUSSON *Vill. Poet* (1897) 72.

YILP, **YILT**, **YILV**(E, see Yelp, *v.*, Gilt, *sb.*[1], Yelve.

YIM, *sb.* and *v.* Sc. [jim.] **1.** *sb.* A particle ; an atom ; the smallest portion of anything. See Nyim.

Ags. (JAM.) Rxb. Nor leaves in creation a yim to afford A bite to a beast or a bield to a bird, RIDDELL *Poet. Wks.* (1871) II. 204. Gall. MACTAGGART *Encycl.* (1824) 500, ed. 1876.

2. A very thin film of condensed vapour or fat. Bnff.[1] **3.** *v.* To break into fragments. Kcd. (JAM.) **4.** To become covered with a thin film. Bnff.[1]

YIMMER, see Yammer.

YIMMET, *sb.* *Obs.* Sc. A lunch ; a ' piece.' Gall. MACTAGGART *Encycl.* (1824). See Yim.

YIMOST, **YIN**, see Eemost, One, Yon.

YINDER, see Yonder.

YINK, *sb.*[1] Sh.I. [jiŋk.] An indefinite quantity.

A yink o' sheep (J.S.).

YINK, *sb.*[2] Sh.I. [jiŋk.] A lover ; a sweetheart. (A.W.) ; S. & Ork.[1]

YINK, *v.* Sh.I. [jiŋk.] To set aside.

Whin hit wis lamb'd we haed da midder o' hit hame, an' Girzzie . . . yinkit hit ta wir Gibbie, as shüne as shü saw hit, *Sh. News* (Aug. 7, 1897) ; S. & Ork.[1]

YINST, see Once.

YIP, see Gip, *v.*[2], Yap, *sb.*[2]

YIPPER, *adj.* e.An. Also in form yepper Suf. [ji·pə.] Brisk, active. in good spirits. Cf. yap, *adj.*

e.An.[1] Nrf. She is right yipper, COZENS-HARDY *Broad Nrf.* (1893) 60. Suf. (HALL., s.v. Yep).

YIP-YAP, *sb.* Chs. Also in form yip-yop s.Chs.[1] [ji·p-jap, -jop.] An upstart ; a young, scatterbrained person. See Yap, *sb.*[2]

Chs.[13] s.Chs.[1] Wo)doo ahy ky'ae·r für ü lit·l skwuu·rtin yip·-yop lahyk dhee? Wot aat·)i bür ü gau·ki wop·-strau üv ü laad·, wen au·)z sed?

YIRB, **YIRD**, see Herb, Earth, *sb.*[1]

YIRDIE-BEE, *sb.* *Obs.* Sc. A bee which burrows in the ground.

Slg. No honey dug from yirdie bees, WYSE *Poems* (1829) 19.

YIRDLINS, **YIRK**, see Earthlins, Yark, *v.*[1]

YIRKIN, *sb.*[1] Irel. [jə·rkin.] The place where a shoe is tied. Ant. *Ballymena Obs.* (1892).

YIRKIN, *sb.*[2] N.I.[1] [jə·rkin.] The side of a boat.

YIRLICH, *adj.* *Obs.* Sc. Also in form yirlisch. Wild, unnatural. Cf. eldritch.

Slk. (JAM.) ; Sett up sic ane yirlich skrighe, HOGG *Tales* (1838) 110, ed. 1866.

Hence **Yirlischly**, *adv.* wildly.

They yellit and youtit soe yirlischly, *ib. Poems* (ed. 1865) 315.

YIRLIN, see Yoldring.

YIRLING, *sb.* Not. [jə·lin.] A thatcher's handful of straw. MORTON *Cyclo. Agric.* (1863).

YIRM, *v.* Sc. Irel. Also written yerm Sc. ; yurm Sc.

Figure 1.4 The richness of regional dialect: a page from Joseph Wright's *English Dialect Dictionary* (1898–1905)

NEUROTIC
PSYCHOTIC
SCHIZOPHRENIC
SCHIZOID
MANIC
CATATONIC
MELANCHOLIC
HYPERCHONDRIAC

NON COMPOS MENTIS
MESHUGA
WACKY
FLAKY

BEREFT OF REASON
DEMENTED
POSSESSED
INSANE

MAD

DAFT
BATTY
BARMY

MENTAL
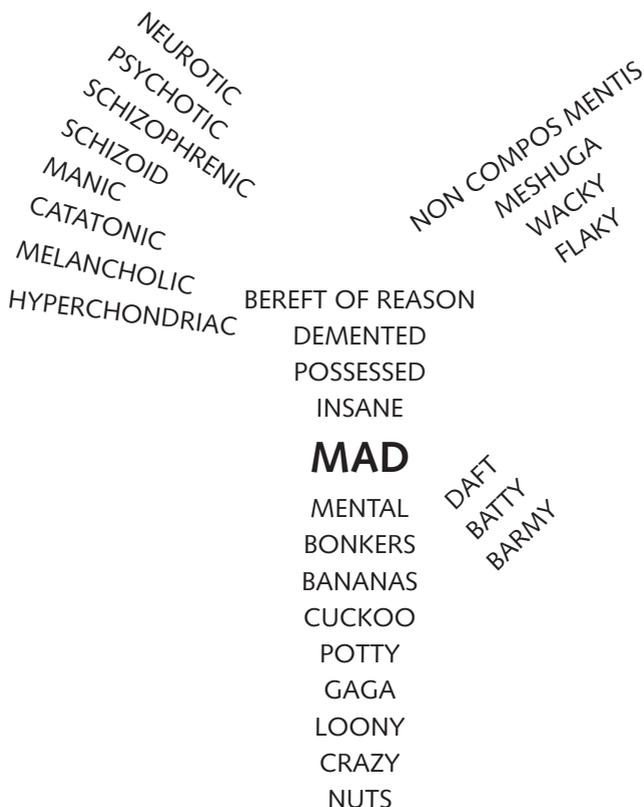BONKERS
BANANAS
CUCKOO
POTTY
GAGA
LOONY
CRAZY
NUTS

Figure 1.5   The semantic field of madness

encapsulate its theme, but they also embody it in a miniature fashion, for they come from the three seminal languages which, over some fifteen centuries, have amalgamated into English. The basic roots and core vocabulary of our language are Anglo-Saxon (also known as Old English), deriving from the ancient Germanic word-hoard brought to England by the Angles, Saxons and Jutes in the fifth century. To this skeleton or foundation there have been two further additions from the continent of Europe. The first is a Romance element, the legacy of the conquering Norman-French elite who took control of the land after the Battle of Hastings in 1066. The second element is classical, taking the form of a more bookish, learned, abstract and technical vocabulary of Latin and Greek terms steadily accumulated by authors and scholars from late medieval times and given increasing impetus by the development of printing from the late fifteenth century.

There have been thousands of subsequent borrowings, from virtually every language in the world. Those from the continental languages are especially numerous, but the tripartite character of the lexical core remains the dominant feature of the vocabulary. If we were to assign and organize the component terms of this volume's 'working' title on the basis of their origins in a scheme of linguistic archaeology, with the oldest terms at the most remote level, we would arrive at the following diagrammatic representation of the 'dig':

Latin and Greek: lexical history
Norman French: guide   language
Anglo-Saxon: a   to the   of the English

We can see here a separation of registers, as we did in the word-field for being pregnant. The common words are to be found in the Anglo-Saxon and Norman French origins, as were 'with child' and 'in the family way', while the Latin and Greek are less familiar, as were *parturient* and *obstetric*. This is a pattern we shall see repeated many times.

If, however, we were to translate our title back into the ancient terms used by scholars of Anglo-Saxon times, such as Bede or Aldhelm or Ælfric, then it would be very different. The title would then be on the same level because other, native words would be needed for *guide*, *lexical*, *history* and *language*. It would then read:

an handboc ealdgesegena wordhordes tunge angelcynnes
('a handbook of the ancient traditions of the vocabulary of the tongue of the English race')

Viewed across the mists of time and the accidents of linguistic change, the Anglo-Saxon does not now seem to us at first sight even recognizable as English, appearing as alien as the markings on some primitive heathen stone. It seems wise, therefore, to defer the deciphering of the ancient forms of the language until the following chapter.

Historical Overview: A Road Map through the Past

The language itself contains evidence of the major demographic movements which occurred in the first millennium of what we may call the English-speaking peoples. It may therefore be useful at the outset to trace in an overview a broad outline of these waves of invaders who have left parts of their language, large and small, in the linguistic amalgam of English.

The 'original' inhabitants of Britain (within the historical period) were the Celts, who some 2000 years ago occupied most of Western Europe. From them are derived the ancient languages of Britain, namely Welsh, Irish Gaelic, Scots Gaelic, Manx and Cornish (both now extinct), as well as Breton, still spoken in Brittany in France. Having been sufficiently dominant to threaten Rome in 390 BC, the Celts thereafter found themselves retreating before that expanding Empire. Within a century of Julius Caesar's assassination in 44 BC, Britain, also known as Albion, had suffered a full-scale Roman invasion by an army of 40,000 men. However, when the Roman legions were in turn withdrawn in 410 Britain became a power vacuum, into which came the Angles, Saxons and Jutes.

What must have seemed a minor turbulence of the times turned out to be the most significant event in British history. Unlike the previous Roman occupation, which had made Britain a colonial outpost of Rome, this was the permanent settlement of a whole people who were not going to be dislodged if the natives got restless. The Celts were subjected to servitude and driven into the mountain fastnesses of Wales and Cornwall, where their language retained a precarious foothold. The next influx of this early period was the dissemination of Christianity, brought initially by St Augustine, who landed in Kent in 597. The linguistic significance of this diaspora was the introduction of Latin scholarship.

From the late eighth century the Anglo-Saxons became in turn the victims of slaughter and rapine, fire and the sword, at the hands of their northern relatives, the Vikings. The depredations of the *Dene*, as they were generically called, became so devastating and so widespread that when they were eventually defeated, King Alfred was able to sue for peace only by ceding to them about half of England, to the north-east, an area which was appropriately called the Danelaw. Old Norse (ON), the language of the Vikings, was a Germanic language related to Old English (OE), but as different as are, say, Modern Dutch and German. These early developments are treated more fully in chapter 2.

The Norman Conquest of 1066 was fundamentally different in nature from the preceding invasions in that the ruling caste spoke an entirely alien language, Norman French, by which they defined themselves and imposed their rule. Consequently Norman French became the language of power and prestige, reinforcing social distance between the elite and the masses. These distinctive qualities lasted for centuries, creating a stratified linguistic separation between the rulers and the ruled. To a certain extent this linguistic class-division has continued between 'us' and 'them', 'upstairs' and 'downstairs', 'U' and 'non-U'. Looking back, what is in many ways more remarkable, is that English not only survived, but re-established itself as the national

tongue in the fourteenth century after ceasing to be the official language of the land for three centuries. This major influx of a dominant, foreign language was the basic cause for the emergence of what historians of the language term Middle English, which is thought of as extending from *c*.1100 to *c*.1500, by which time the more standardized form of Early Modern English was developing.

In the course of the Middle English period, as scholars once more absorbed the fruits of classical learning, so the third major strand of the English vocabulary was slowly intertwined with the Germanic and French. This Latin and Greek influx was also different in character from the previous one in that it came more directly from the source, linguistically speaking. The Celts and the Germanic tribes had, of course, encountered Roman merchants and legions on the continent, and had borrowed from them basic terms of food and measurement; the missionaries from Rome had brought their special religious vocabulary; the Normans had established their rule in their Romance dialect; now literary Latin came 'direct' in the form of manuscripts and charters. So it comes about that many a basic word may be borrowed more than once: Latin *discus* was first absorbed as a trading borrowing as A-S *disc*, yielding Modern *dish*; it acquired a Late Latin sense of 'table', which developed as Medieval Latin *desca*, borrowed as Middle English *deske*, Modern *desk*; the sense of 'table' also emerged in Old and Middle French as *deis*, which has become Modern *dais*; it generated *disc* in its various senses, including the compact computer device; finally, *discus* in the athletic sense was borrowed 'direct' in the sevententh century. The lexical developments of the Norman Conquest and Middle English are dealt with in chapter 3.

Thus, in the millennium and a half between the coming of the Roman legions and the absorption of Classical Latin, a whole series of linguistic and lexical layers had been brought to England. In terms of our archaeological metaphor they would appear thus:

Late Middle English ⇒	**Classical: Latin and Greek**
1066 ⇒	**Norman French**
787 ⇒	**Scandinavian**
597 ⇒	**Religious Latin**
455 ⇒	**Anglo-Saxon**
55 AD ⇒	**Colonial Latin**
2000 BC ⇒	**Celtic**

By the end of the Middle English period the vocabulary was heterogeneous and diversified, having acquired three registers, reflecting the differing status of the component elements, namely a Germanic base of common, basic words,
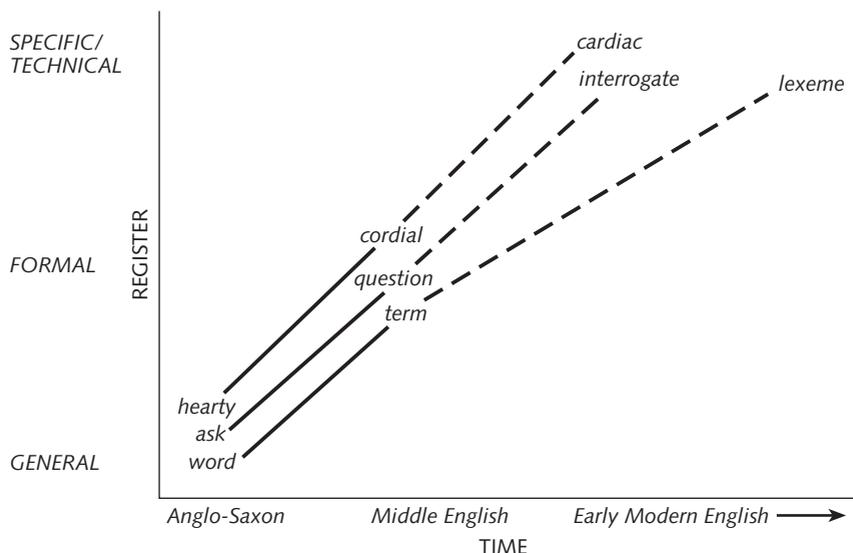
Figure 1.6   The correlation between lexical origin and register

a courtly, formal register from Norman French and an intellectual, abstract and bookish register from Latin and Greek. As a cursory study of the thesaurus will show, virtually any word-field reflects this structure. For example, in the following lists, the differing registers are clearly apparent:

Anglo-Saxon	Norman French	Latin/Greek
ask	question	interrogate
hearty	cordial	cardiac
folk	people	population
go	depart	exit
holy	sacred	consecrated
lively	vivacious	animated
guts	entrails	intestines
gift	present	donation
word-hoard	vocabulary	lexicon
word	term	lexeme

(In this study, the general terms *word* or *term* are preferred to *lexeme*,[2] the comparatively recent term used in linguistics for a lexical unit.) The relationship between lexical origin and register is shown in figure 1.6.

The invention of printing in the late fifteenth century had two contrary influences on the language. Although the press acted as a stabilizing force on

spelling, it accelerated the rate of lexical change. Words no longer travelled at the rate of a migrating community; they sped across continents in the new editions of the Renaissance. As the ancient works of the classics were translated into the vernaculars, so the influx of borrowings became a flood. Some scholars saw the process as an enrichment and embellishment of English, others as a corruption of the 'purity' of the language. These opposing attitudes, one endorsing the high status of the foreign, recherché word, the other showing chauvinist hostility to the alien interloper, consolidated into what was known as the Inkhorn Controversy. Although historically it proved to be something of a storm in an inkwell, the underlying attitudes have proved surprisingly resilient, as we shall see.

The Renaissance, which is discussed fully in chapter 4, saw a great efflorescence in the growth of the language as a fine literary and dramatic instrument. In this respect Shakespeare stands pre-eminent as the author who has made the greatest single-handed contribution to English. Furthermore, the resonant yet simple use of the language in the Book of Common Prayer (1549) and the translation of the Bible, especially the Authorized Version of 1611 (commissioned by King James) clearly gave the language increasing authority. Surprisingly, this was also the period when the argot of the underworld first appeared in the published form of canting dictionaries, which gave permanence to a whole lower register.

The period of Early Modern English (1500–1700) also saw an important change in the global status of the language as England became a major colonial and trading power. This remarkable change occurred in a mere two centuries: around 1400 English was still in the process of re-establishing itself as the official language of England, whereas around 1600 it was starting to be transplanted across the world. The early explorers and colonists also brought back with them increasing numbers of exotic words which were steadily absorbed in the language. Thus from the so-called New World came *potato*, *tomato*, *barbecue*, *buccaneer*, *cannibal*, *canoe*, *hammock*, *hurricane*, *maize*, *tobacco* and *chocolate*. This process has steadily continued as a consequence of colonialism and the spread of English. More significantly, colonial outposts were established in the Carolinas (1584), Virginia (1607) and New England (1620) which were to grow into the largest English-speaking community in the world, with distinctive features of pronunciation, lexis and idiom which have increasingly differentiated the offspring from its parent stock. The same is true of Indian English, Australian English, West Indian English, and those varieties which make up the mosaic of modern world English. The borrowings from colonialism and imperialism supply the material for chapter 6.

Words as Mobile Forms

Up to now we have been considering the general expansion of English lexis mainly as a consequence of demographic change. But there were important developments going on within the language itself which accelerated lexical change. In order to understand them we need to consider briefly the relationship between grammatical form and function.

The fundamental change in the development of English grammar has been in the direction of simplicity of form. Put more technically, there has been a **reduction of inflections**, an **inflection** being a word-ending or **suffix** which has a grammatical function, such as *-s* to indicate plurality in nouns or *-ed* to indicate the past tense in verbs. For various reasons the English inflectional system has become simpler over time, as the number of forms has been reduced. Consequently increased flexibility of function has been acquired by the individual form of a word, since the reduced number of forms has been required to take on a greater range of functions. If we ask the question 'What part of speech is *love?*' the answer now depends not on the form, but on the context, as we can see from the following statements:

I *love* you. (verb)
Love is a many-splendoured thing. (noun)
Isolde drank the *love* potion. (adjective)

Such flexibility was not possible in the earlier stages of the language, since different functions were indicated by specific inflections. In Anglo-Saxon the forms required for the contexts listed above would be quite distinct, namely *lufie*, *lufa* and *lufe* respectively. Though we do not spend much time thinking about it, our present **conjugation** of the verb *love* consists of but four forms: *love*, *loves*, *loved* and *loving*. These are the 'survivors', so to speak, of eleven different forms in Anglo-Saxon.

The major point of all this is that from late Middle English, words had largely shed their grammatical inflections and were thus no longer limited to particular functions. Words were set free, becoming what is called in modern grammar 'free forms'. This meant that they could be used in all sorts of new ways which had not been possible before. For instance, *dog*, an ancient noun (originally OE *docga*) was used as a verb for the first time (meaning 'to follow like a dog') in *c*.1519. Likewise, *hound* first developed a verbal capacity *c*.1518, *fox c*.1567 and *ape c*.1632. Simeon Potter points out that 'Names of many parts of the human body – *eye*, *nose*, *mouth*, *arm*, *breast*, *shoulder*, *elbow*, *hand*, *knuckle*, *thumb*, *stomach*, *leg*, *foot*, *heel* and *toe* – have come to be used as verbs'

(1975, p. 169). This process, called **conversion**, applied both to native words and to classical borrowings (which we shall discuss in chapter 4).

It became possible for individuals to exploit this new flexibility in a creative fashion. Shakespeare, for example, frequently extended the grammatical functions of words in new ways. Cleopatra, one of his most poetically liberated characters, says of Antony: 'He words me, girls' (i.e. he's 'chatting me up'), using *word* as a verb for the first recorded time in 1608 (V. ii. 190). After her final defeat, she anticipates the humiliation of being *windowed* ('shown off') in Rome and of being made the topic of theatrical spectacle in which some youth will '*boy* my greatness' (V. ii. 219). Here we can see very ancient words being given a new grammatical and lexical lease of life.

It is hard to overstate the significance of this development, since there was and continues to be a great semantic expansion from the same lexical base. Thus *tough*, originally an adjective, has expanded to being a noun ('a tough') and (in US usage) a verb, as in 'to tough it out'. Similarly, *clock*, originally a noun, has acquired at least three senses as a verb: to hit, to measure in time and 'to clock in' for work. In contemporary English this has become commonplace, with nouns like *showcase*, *mothball*, *host*, *flight*, *fuel*, *target* and *mushroom* being increasingly used as verbs. This theme is taken up in chapter 7.

Modern English: establishing the lexicon

The seventeenth and eighteenth centuries saw the steady coalescence of a notional Standard English. The desire of the Enlightenment for the imposition of rational order was apparent in attempts to 'fix' the language, i.e. formalize it on a rational and scientific basis. This took the form of a number of proposals for an English Academy to regulate the language, and more concretely, the first major attempts at comprehensive dictionaries by Nathaniel Bailey (in 1721 and 1730) and, most famously, by Samuel Johnson (in 1755). These established the word-stock more authoritatively, giving rise to the influential notion that for a word properly to exist, it should be 'in the dictionary'.

As we shall see in chapter 5, which deals with the evolution of the dictionary, the latter part of the nineteenth century was a period of intense lexicographical activity. This was evidenced principally by the compilation of the great *Oxford English Dictionary* or *OED* (1884–1928), with a first Supplement (1933), the major source-work for the reconstruction of the lexical history of the language. Other remarkable lexicographical achievements focusing on less central areas of the lexis were Joseph Wright's *English Dialect Dictionary* (six volumes, 1898–1905), Bosworth and Toller's *Anglo-Saxon Dictionary* (1894) and J. S. Farmer and W. E. Henley's *Slang and its Analogues: Past and Present* (seven volumes, 1890–1904).

Recent decades have seen the publication of the four-volume *Supplement* to the *OED* (1972–86). In 1989 the two sequences were integrated, producing the second edition, consisting of twenty volumes. In addition to the third edition of *Webster* (1961), definitive works on Australian, South African and New Zealand English have been produced, in 1988, 1995 and 1998 respectively. For a variety of reasons, contemporary English shows a diminution of previous attitudes of suspicion or haughtiness about borrowings, and has become omnivorously receptive of words of all varieties, new words, exotic words, 'buzz' words. Consequently, the major dictionaries have difficulty keeping up, and compilations of new words appear almost on an annual basis. Dictionaries on historical principles are giving way to those based on corpora or collections of quotations showing written or spoken language. New lexical varieties are discussed in chapter 7, while the concluding chapter considers changes in the lexical structure, as well as the implications of these developments for the future of the language as English or Englishes.

Semantics and Lexis:
Synonyms, Antonyms and Hyponyms

As we can see from this overview, there are many kinds of words, or lexical varieties. Apart from the variations of register, there are terms for geographical features or place-name elements, such as *fell*, *cwm* and *by*. These are, in the nature of things, often regional or dialectal. However, as we shall see, place-names are also reflective of social dominance in various ways: originally they reflect demographic movement (*Denby*, for example, meaning 'the town of the Danes'); subsequently they reflect political dominance, as in the use of *Victoria* and *Wellington* all over the former British Empire.[3] There are general names for things, such as *book* and *guts*, and specific terms, such as *lexicon* and *intestine*. There are names for qualities, such as *hearty* and *cordial*, and for abstractions, such as *beauty* and *idea*, as well as for whole areas of study, such as *philosophy* and *physics*. There are terms for political ideologies, such as *democracy* and *communism*. There are even words for imaginary beings like *unicorn*, *elf*, *banshee* and *gnome*. We can see here a range from basic words for everyday communication to crucial keywords by which a whole nation's values may be defined or altered. These may extend to phrases, such as *for king and country*, *it's not cricket* and *free enterprise*, which encapsulate certain ethical notions. In a useful phrase, the literary critic William Empson called them 'compacted doctrines' (1977, p. 21).

Most languages contain such key cultural terms: *cunha* in Portuguese defines a sense of obligation to respond to a call for help; Spanish *mañana* conveys an unhurried, lackadaisical attitude; *Deutsche ordnung* sums up a German passion for efficiency; *noblesse oblige* is a Gallic formulation of upper-class obligations to behave nobly. This said, it would be naive to assume that these terms define national characteristics in any strict sense. Still more erroneous is it to construct supposed national characteristics on the basis of vocabulary. Consider these chauvinist sentiments on the German word *Schadenfreude* ('malicious pleasure at the misfortune of others') in the *Spectator* of 24 July 1926: 'There is no English word for *Schadenfreude*, because there is no such feeling here.' Leaving aside the highly questionable assumption that Germans are more guilty of *Schadenfreude* than other nations, the word English *gloat* conveys, albeit in a broader sense, the essential meaning of *Schadenfreude*.

Leaving these broad cultural matters of language, the essential point is that the lexical richness of English makes it possible to articulate a great variety of shades of meaning or semantic nuances, and it is to the general interrelationship between semantics and lexis that we now turn.

Semantics is the study of meaning, which is a complex matter in that it involves the relationship between words, ideas and things as well as the relationship between words of similar meaning. A distinction is often made in this respect between **reference**, or the relations between language and the world, and **sense**, or the relationship between words of similar meaning. Semantics also examines how sets of words are used to divide up experience: thus in English *black* and *blue* designate different colours, whereas in Old Norse the term *blá* served for both; in Russian, however, there are two distinct terms for blue, *goluboy* for light blue and *siniy* for dark blue, like *azure* and *violet*. The prism of colour terms tends to open up with time and cultural contact: historically *black*, *white*, *red*, *yellow* and *green* are Anglo-Saxon in origin, but *blue*, *brown* and *orange* entered the vocabulary via Norman French (as did *azure* and *violet*).[4]

Up to now we have been discussing words and meanings largely in terms of **synonyms** in particular word-fields, showing how these reflect in their structure the evolution of the language. And we have noted that exact synonyms are seldom to be found, even in a multitude of similar terms. However, meaning is also conceived of in terms of opposites and negations, generating the category of **antonyms**. We note, as a prime example, that Roget's *Thesaurus* is fundamentally structured by means of synonymic and antonymic categories. Some of these derive from the physical world, for example:

> light/darkness
> heat/cold
> summer/winter

These we may think of as being mutually exclusive, i.e. one category denies the other, or as the classical grammarians put it, *tertium non datur*, i.e. there is no third possibility (Lipka, 1990, p. 146). Others may derive from the physical world or nature, but are more complex:

male/female
human/animal

The male/female categorization seems absolute, but is complicated by *hermaphrodite* (having characteristics of both sexes) and *neuter* (having neither). Likewise, the human/animal dichotomy, thought of as absolute by medieval philosophers, is complicated by terms like *brutal* and *bestial*, which mean broadly having the characteristics of an animal, but can be applied only to humans.

Other antonyms derive from social categories, for example:

single/married

Once again, this pair has traditionally been regarded as mutually exclusive, but is complicated by the modern categories of *separated* and *living together* or being *a common law husband/wife*. One can refer, somewhat unkindly, to a person as being 'much married'. Other examples, philosophical or moral, are:

good/evil
physical/spiritual
kind/unkind
fair/unfair

Most of these are categorized by authorities under the heading of 'complementarity' (Lipka, 1990, p. 145). It is perhaps a sad reflection on human nature that the 'negative' word-fields in the thesaurus are so much larger than the 'positive'.

English has developed the simple category of the antonym in a variety of nuanced ways, mainly by the subtle use of complex negatives. Thus the plain antonymic pair of *talented/untalented* is complemented by the double negative *not untalented* which, despite its literal formulation, implies 'having considerable talent'. Similarly *able* and *lacking in ability* are complemented by *not without ability*, and so on. These *not un-* categories show the fallacy of applying to semantics the simple arithmetical or logical notion that 'two negatives make a positive.' They indicate a more guarded, noncommittal assessment on the part of the speaker or writer, who avoids the baldly negative or positive.

Although negative forms like *uninterested*, *disinterested*, *unmake*, *non-intervention* and the like have been growing by the thousand as the language has developed, new kinds of negatives have grown up in recent decades. Thus a positive term like *charming* has acquired a whole variety of antonyms ranging from *unattractive* to *repulsive*; to these has been added the equally critical coinage of *uncharming*, first used by Dryden in 1687 but currently fashionable. On the same model are the modern formations *unchic* (1960), *unglamorous* (1960), *unprestigious* (1968), *unsexy* (1959), *un-with-it* (1965) and many more. Bram Stoker resuscitated the form *undead* (recorded in Anglo-Saxon) in his *Dracula* (1897), while *ungreen* has similarly been revived to mean 'environmentally unacceptable'. As we shall see in chapter 2, Anglo-Saxon had an interesting lexical category, the 'intensive antonym'.

One of the curiosities of the field is the category of **pseudo-antonyms**, the considerable number of negative forms which either lack a positive or use a negative form in a way unrelated to the positive. In the first category are *inclement*, *uncouth* and *disgruntled*, which lack a positive *clement*, *couth* and *gruntled* (although D. G. Wodehome used 'gruntled' (facetiously in 1938). Similar 'negative half-pairs' are *unspeakable* and *malcontent*. The second category includes forms like *disaffected*, *disagreeable* and *indifferent*, the last of which has developed *two* senses ('mediocre' and 'unconcerned'), which are quite unrelated to *different*. 'New' negatives coined in recent years include *unbundle*, *unban*, *unpick*. Less numerous are the new **pseudo-positives** like *forgettable*, used ironically of books, films and events to which *unforgettable* has been too commonly applied.

The **hyponym** is a generally less well-known term, having been coined only in 1963 by John Lyons (Lipka, 1990, p. 141). It describes an important notion which is central to the way we define meanings in terms of lexical structures. Thus we would define *crimson* as 'a brilliant red colour', i.e. it is a particular variety of the categories 'colour' and 'red', and *camembert* as 'a kind of cheese'. In other words, even though we may not be aware of it, we construct meaning in terms of a hierarchy of categories, ranging from the generic, technically called the **hypernym** (literally 'the word above') to the subsidiary or **hyponym** (literally 'the word below').

Roget's *Thesaurus* works on this lexicological structure, dividing the world and the map of human knowledge into six vast general categories, namely Abstract Relations, Space, Matter, Intellect, Volition and Emotion, into which are subsumed Religion and Morality. Each category has numerous subdivisions, ranged hierarchically through hypernyms to hyponyms. The great merit of Roget's scheme is its flexibility and its capacity to accommodate new terms and concepts as the speech community using English has evolved. Each revision (and there have been over thirty since 1852) shows further expansion. As

civilization and technology have developed (not always in concert) so previous hyponyms have become hypernyms. For example, *atom* in the age of Newtonian physics was a hyponym, meaning the smallest indivisible form of matter, from its Greek etymology 'that which cannot be cut'; with the growth of modern atomic physics and the splitting of the atom, it has become a hypernym. A contrary example would be *soul*, which in the Middle Ages was a crucial spiritual hypernym, but which has largely fallen out of use in modern secularized society, having developed a sense in modern music as a hyponym. A useful discussion of schematic arrangements of knowledge is to be found in Tom McArthur's *Worlds of Reference* (1986).

In a profound observation, George Santayana perceived that a split in the register of the English vocabulary between the Germanic and the Classical induces a parallel division of consciousness:

> In French, Italian and Spanish, as in Latin itself . . . the reader passes without any sense of incongruity or anti-climax from passion to reflection, from sentiment to satire, from flights of fancy to homely details. . . . As the Latin Languages are not composed of two diverse elements, as English is of Latin and German, so the Latin mind does not have two spheres of sentiment, one vulgar, the other sublime. All changes are variations on a single key, the key of intelligence. (1916, pp. 131–2)

Before we proceed further with our discussion of semantics and lexical variation, in which Santayana's observation can be assayed, let us consider the essential problems of definition and the role of the dictionary.

Agreed Meanings: Usage and the Dictionary

How do we learn or know what a word means? Essentially we rely on two models, namely usage and dictionary definitions, both of which have their strengths and limitations. In chapter 5 the main focus will be on the evolution of the dictionary as a form, starting in the late sixteenth century. But at this stage we need some discussion showing that words do not have absolute values like numbers, and that meanings are conventional, according to global, regional and social contexts.

As the early users of the language were for the most part illiterate, and as we are all illiterate at the first stages of our acquisition of language, usage has been the predominant force, historically speaking. Isolated speech-communities, such as those of Tristan da Cunha and some boarding schools, have their own vocabulary.[5] But most groups use words current in global English in their own quite distinctive ways. Thus soldiers in the US Army prefer

the euphemism *wasted* to the blunt 'killed', the Sloane Ranger set in London uses the term *wrinkly* for a 'middle-aged person', while in Black American teenage street talk *bad* means 'good' and *vicious* means 'excellent'. The users of Valspeak, in the vicinity of the San Fernando valley of Los Angeles, use *radical* positively and unpolitically to mean 'good', 'wonderful' or 'challenging', while *satisfactory* will have very different meanings in a school report and in a medical bulletin. Likewise, a basic word like *grass* has the special meaning of 'informer' in the criminal underworld and 'cannabis' or 'marijuana' in the drug culture. Several of these meanings are, of course, unfamiliar to outsiders and, like most 'in-group' language, prone to fashion. Outsiders can make considerable gaffes when venturing into unfamiliar semantic terrain: perhaps the most hilarious was that of Robert Browning, who wrote innocently of 'cowls and twats' in *Pippa Passes* (IV. ii. 96) 'under the impression that it [twat] denoted some part of a nun's attire'. It is usually only when one travels outside one's community, either physically or via the written word, that one becomes aware of the same word covering different semantic areas. Thus *pond*, *vest*, *trailer* and *pants* have different meanings in American and British English, while *bagarap* is the general word for 'destroy' in Pidgin English, being an erosion of *bugger up*, but lacking any sense of impropriety. *Bastard* has a strongly critical sense in America, a broad range of tones in Britain, ranging from hostility to sympathy, but is a fairly mild term in Australia, where it is often preceded by the adjective 'good'.

To the members of these various speech-communities, their particular usage is predominant, and if they are illiterate, their meaning is for them the only meaning. It is worth observing here that *speech-community* can be a slightly misleading term, since it implies an unrealistically circumscribed area of relationship. It may have been a valid notion in the past, when society was comparatively static, and can still be applied meaningfully to isolated and 'primitive' communities, but in modern, commuting, socially mobile, mass-mediated society, most people belong to more than one speech-community. They will consequently use different language conventions at work, at home, and with the various professional and social sets to which they may belong. Therefore, in a profound sense, 'learning the language' involves becoming aware of, and discriminating between, the usage of these different speech-communities. As the editor of the *Oxford English Dictionary*, Sir James Murray, observed in the preface to the great work, 'No one man's English is *all* English' (1884, p. vii).

The Uses and Abuses of Etymology

Etymology is the study of the root or origin of a word: it derives from the Greek root *ɛtymos*, meaning 'true'. The importance and the implications of etymol-

ogy are considerable, as we shall see. Generally speaking, there are two contradictory processes at work in the relation between etymology and meaning. The first is a gradual erosion of the original link, discussed more fully in the following section on semantic change: words tend to move steadily away from their original meanings. Contrary to this is a desire to revive the link, to get words 'to make sense' with their past, an attitude which has various consequences, affecting not just semantic change, but almost everything to do with semantics. First, people prefer memorable or logical origins for words, and even invent them if they do not exist. Some words do indeed have such striking origins. Few of us ever forget (once we are told) that the *sandwich* derives from the Earl of Sandwich, a compulsive gambler who, in order not to leave the gaming table during a twenty-four-hour bout, sustained himself in part with slices of cold beef between slices of toast. Thus was born the *sandwich*, first recorded in 1762. Similarly engaging origins are discussed later in this chapter.

The two basic words used in greeting and parting provide a useful pair of examples demonstrating the erosion of the link between etymology and meaning. *Hello* is a later form of *hallo*, *halloo* and *hollo*. In these earlier forms it was not a greeting but a shout to call attention or to express surprise, which was still the sense up to about a century ago. (Interestingly, this was the original sense of *Hi!*, which has now adopted a less challenging meaning and is a standard greeting, especially in the USA.) In the forms *halloo* and *hollo* it was a hunting cry used to urge the hounds on in the chase, recorded as far back as the sixteenth century. *Goodbye* shows more dramatic erosion from its origins, which were in the phrase *God be with you*. Many intermediary forms such as *God be wi ye* and *God bye* are recorded from the sixteenth century. The change from *God* to *good* started around 1700, possibly as a result of confusion with the other formulas of parting, such as *good morrow*, *good day*, *good evening* and so on.

But most words have dull, obvious or unclear origins. Nevertheless, plausible, colourful explanations are proffered. In my own experience, I recall over a dozen people informing me that the word *fuck* (which has complex and uncertain roots) 'in fact' originates in a coded acronym: one group insisted that the term derived from the words 'fornicate under command of the King', supposedly a royal edict issued during the time of plague. Another group insisted (with equal certainty) that the word was a police acronym for the phrase 'for unlawful carnal knowledge'. These are examples of **folk etymology**, the positing of plausible but inaccurate explanations for the origin of a word.

Folk etymologies, as their name suggests, are collective and spontaneous. The origin of the word *woman*, for example, lies in A-S *wifmann*, which transliterates as 'wife-man', since A-S *mann* meant both 'man' and 'mankind' and *wif* meant 'woman'. However, as the *OED* notes, the word was used 'in the 16th

and 17th centuries frequently with play on a pseudo-etymological association with *woe*'. The first instance is given as *c.*1500 from the Chester plays, while even that noted humanist Sir Thomas More could write in 1534: 'Man him-selfe borne of a woman is in deede a wo man, that is full of wo and miserie.' Simeon Potter records Ruskin's preference for another pseudo-etymology of the same word: he 'found pleasure in reminding the married women in his audience that since *wife* means "she who weaves", their place was in the home' (Potter, 1961, p. 106). In recent times pressure groups have taken to invent-ing origins for ideological purposes: the form 'wimmin' has been created by some feminists who stated: 'We want to spell women in a way which does not spell men' (*Observer*, 13 March 1983). This remains a minority and specialist use. In other cases the speech community as a whole will bend the shape of a word to suit the meaning: thus the fanciful form *sparrowgrass* was created in the seventeenth century from *asparagus* (which has been in the language since Anglo-Saxon times). (Dr Johnson lists only *sparrowgrass*.) There are many cases of innocent misnomers: thus the 'white' rhinoceros is so called, not from its colour, but from a corruption of '*wide*', referring to the distinctive shape of its mouth. Once the misnomer was established, it became logically necessary to extend the confusion by referring to the other variety as 'black', although it is indistinguishable in colour.

Instances of folk etymology affecting spelling (to suggest the origin of the word) are surprisingly numerous. For example in *bridegroom*, the second element turns out to be a confusion of A-S *guma*, 'a man'. The Anglo-Saxon form *bryd-guma* survived up to the fourteenth century as *bride-gome*, but is later superseded by *bridegroom*, probably because *gome* had become obsolete. The *Jerusalem artichoke* has nothing to do with Jerusalem, but is so called through a distortion of Italian *girasole* meaning 'sun flower' (both plants belong to the same botanical genus).

In similar fashion *humble pie* derives from the *numbles*, originally a loin of veal, but subsequently downgraded to 'certain inward parts of an animal as used for food'. *Belfry* comes from OF *berfroi*, 'a watch-tower' before the form changed from popular associations with *bell* in the fifteenth century. In the phrase 'to *curry favour*', the second element derives from Favel, the name of a famous medieval French horse. *Crayfish* derives from OF *crevice*, Modern French *écrevisse*, more related (etymologically) to the crab than the fish. In similar fashion *penthouse* is an understandable distortion of ME *pentice*, 'a small sacred building dependent upon a larger church', while *shamefaced* comes from the OE form *sceamfæst*, meaning 'bashful or modest'. The *cellar* in a *salt-cellar* is a corruption of OF *saler*, meaning a container for salt.[6] The oddly named *Welsh rarebit* comes about from an etymologizing alteration of what was originally and facetiously called *Welsh rabbit*. Two foreign terms showing the

same process are *mongoose* (from Marathi *mangus*) and *cockroach* (from Spanish *cucuracha*). In all these cases we can see that as the original form ceases to be understood (through time or foreignness), so it is changed to something recognizable or probable.

The study of semantic change necessarily shows us that the *etymology* or original root meaning of a word has little subsequent status, despite the etymology of *etymology*. Nevertheless the 'argument from etymology' is often resorted to, in forms such as this: 'The word *aggravate* is derived from Latin *gravis*, "heavy or serious" and therefore means "to make worse"; it should not be used to mean "to annoy".' On such a basis one could supposedly insist that a *climax* is a 'ladder', that a *marshal* is 'a boy who looks after horses' and that a *candidate* is a person dressed in white.[7] Manifestly this is not so; one cannot turn back the semantic clock so drastically. Nevertheless, this kind of argument tends to be used by older, usually more educated people to resist the development of a new meaning. The shrewd comments of C. S. Lewis are apposite here:

> Statements that *honour*, or *freedom*, or *humour*, or *wealth* 'do not mean' this or that are proof that it was beginning to mean, or even had long meant, precisely this or that. We tell our pupils that *deprecate* does not mean *depreciate* or that *immorality* does not mean simply *lechery* because these words are beginning to mean just those things. We are in fact resisting the growth of a new sense. (1960, p. 18)

The 'argument from etymology' can also be shown to be fallacious through the study of **doublets** or words which derive from a common origin. some of these are quite remarkably diverse. For example, *lobster* and *locust* are doublets, as are *glamour* and *grammar*, *cretin* and *Christian*, as well as *zero* and *cypher*. In other cases the root word can ramify and mutate to the point that the descendants no longer resemble the parent. Thus Latin *panis*, meaning 'bread' is the root of words as diverse as *pannier*, *companion*, *pantry*, *pastille* and *marzipan*. Similarly, salt used to be a valuable commodity, as is evidenced in such phrases as 'the salt of the earth' or to be 'worth one's salt', and so on. Yet the root notion is no longer obvious in such derivatives as *salary*, *salad*, *sauce*, *saucer*, *sausage*, *silt* and the verb to *souse*.

Yet etymological roots often reveal obscure and tantalizing connections which words retain through their semantic changes. Thus *custom* and *habit*, now essentially mental in their senses, are both rooted in the concept of clothes; likewise the verbs *flounce* and *bustle*, both meaning to move with agitation, derive from parts of a woman's dress.

The importance of etymology, for all its fascination, should not be over-rated. In the eighteenth century, when serious lexicography started to evolve,

the assumption developed that 'proper' words, like people of 'good' family, came from clearly defined origins, namely Latin, Greek and Germanic (or 'Teutonic', to use the contemporary term). If they did not, then they were regarded as linguistic bastards. Dr Johnson was particularly vexed by contemporary slang terms, which seemed to have come from nowhere. He could not ignore them; neither could he explain them. So he cast them into outer linguistic darkness. He thus denounced *to banter* as 'a barbarous word, without etymology'; nevertheless, the word had been in use for nearly a century, having been recorded by Samuel Pepys in his *Diary* on 24 December 1667. It has continued to thrive, despite Johnson's stricture. More significantly, we do not know the origins of basic words like *boy*, *bird*, *child*, *dog* and *smell*, but ignorance of their origins casts no aspersions on their legitimacy.

The Problems of Definition

While the dictionary will obviously seek to give a clear meaning for a term or a quality, much of life remains semantically elusive. As a character in Tom Stoppard's play, *Jumpers*, pronounces, 'Language is a finite instrument crudely applied to an infinity of ideas' (1972, p. 63). Persuasive though this view undoubtedly is, the dictionary contains an astonishing diversity of notions and terms. For example, a *gammerstang* is a tall awkward person, usually a woman, a *battologist* is 'one who needlessly repeats the same thing', to *slonk* is 'to swallow greedily', to *yamph* is 'to bark, especially of a small dog', a *taisch* is in Gaelic folklore 'the apparition of a living person who is about to die', while to *digitate* means 'to speak with the fingers', an activity which is becoming increasingly common, when people gesture that a particular word should be put into inverted commas or what are called 'scare quotes'. All of these are to be found in the *OED*. However, there are words in the dictionary like *unicorn* and *mermaid* for which there are no factual referents. The odd word *muggle* the *OED* classifies perplexingly as having 'origin and meaning obscure', while the odder entry *sooterkin* is defined with equal sobriety as 'an imaginary kind of afterbirth formerly attributed to Dutch women'.

Even in defining the physical world the dictionary is not entirely satisfactory. An Eskimo looking up *fallow* (adj.) in the *OED* would not be much helped: 'Of a pale brownish or reddish yellow colour, as withered grass or leaves'. In Anglo-Saxon times the meaning was even more unspecific, *fealwe* being applied, variously, to the colours of gold, the sea and horses, in which last sense it is still used of *fallow* deer. We may ascribe this lack of clarity quite fairly to the basic instability of colour terms. For example, 'A gem of a sky-blue to apple-green colour' is the *OED*'s definition of *turquoise*, while *azure* (from lapis lazuli) is 'a bright blue pigment or dye . . . Prussian blue'. How many people would

recognize the definition 'a very pale blue with a trace of red' as describing the colour of *lavender?* Emotive terms also present problems. The *COD* (sixth and seventh editions) defines *chuffed* as 'Pleased; displeased', while the simple ejaculation *shit*! covers a range of emotions from anger, disappointment, surprise, pleasure and exasperation. The context will usually determine which meaning is appropriate.

The simplest matters are often the most difficult to define. It takes most of us some time to decode this: 'A perennial plant with single woody self-supporting stem, usually unbranched for some distance above ground'. It is the *COD*'s definition of *tree*. Dr Johnson's *tree* has, for a modern reader, some affinities with science fiction: 'A large vegetable rising, with one woody stem, to a considerable height'. Here our difficulty arises from his use of *vegetable* in its old broad sense. Sometimes words may retain in certain contexts an anachronistic disguise. Consider the following statement: 'A rocket, having more thrust than a jet engine, is the ideal mechanism for launching a satellite.' To us, this has the appearance of an unmistakably modern passage. But this notion derives solely from the technology it describes. On a strictly lexical basis, the passage could have been written as far back as *c.*1700, since by that time the key words *rocket*, *jet*, *engine*, *launch* and *satellite*, had all acquired meanings which would make broad technical sense in this context, even though most of them started with quite different meanings.

We may briefly consider the problems posed by old words used in some special technical sense which is no longer clear. For example, the noun *stole* carried a number of senses, from 'robe', to 'vestment or narrow strip of silk or linen worn by an ecclesiastic', before developing (about a century ago) the present meaning of 'a woman's fur or feather garment' of a similar shape. However, there is a second sense, *stole*[2], which the *OED* notes:

> Commonly identified with *stole*[1], to which the unauthenticated sense of 'royal robe' is commonly assigned. But there seems to be little doubt that the 'stole chamber', served by the Groom or Yeoman of the Stole was originally the room containing the king's close-stool [privy or lavatory] and that the word is properly a variant of STOOL.

This amusing example records a process which is reasonably common, whereby an erroneous meaning may become institutionalized, often aided by misleading spelling.

An apparently comprehensive definition may, on the other hand, be equally inadequate, as Dickens shows in his bitter satire on utilitarian education in *Hard Times* (1854). When the classroom tyrant, Thomas Gradgrind, demands of his class the definition of 'horse', his lackey Bitzer responds with a tissue of opaque equine information in the bloodless style of an encyclopaedic dictionary:

> Quadruped. Graminivorous. Forty teeth, namely twenty-four grinders, four eye-teeth, and twelve incisive. Sheds coat in the spring; in marshy countries, sheds hoofs too. Hoofs hard but requiring to be shod with iron. Age known by marks in mouth. (chapter 2)

Dr Johnson was more succinct and practical: 'A neighing quadruped, used in war, and draught and carriage'. Often dictionary definitions depend on other dictionary definitions: the horse, according to the *OED* is 'A solid-hoofed perissodactyl quadruped (*Equus caballus*), having a flowing mane and tail, whose voice is a neigh'. (*Perissodactyl* is defined as 'having an odd number of toes on each foot'; for *Equus caballus* one would need a Latin dictionary or a guide to zoological nomenclature.)

Most people get through life without needing a definition of a horse. They use the dictionary to discover or to verify and, in fairness, the best definitions capture the essence of what are often very subtle qualities, as in Dr Johnson's definition of *pedant* 'A man vain of low knowledge, a man awkwardly ostentatious of his reading' and *to sit*: 'to rest upon the buttocks'.

One essential problem which compilers of dictionaries increasingly face is what level of usage is to be assumed. The older dictionaries tended to assume a written standard; the more recent, being increasingly based on corpora of usage, incorporate spoken idioms and a great variety of oral usage. These issues are discussed more fully in chapter 5.

Semantics

A historical study such as this necessarily involves some coverage of semantic change, namely the remarkable and fascinating changes of meaning undergone by words over time. The *OED* supplies copious details of such changes, showing that few words have maintained a stable meaning through their histories. Here are some examples:

wan	A-S *wann*, 'dark'
worm	A-S *wyrm*, 'dragon'
free	A-S *freo*, 'noble'
fiend	A-S *feond*, 'enemy'
silly	A-S *sælig*, 'blessed'[8]

This process differs from lexical change, which concerns changes in the structure of the vocabulary. As a consequence of the Norman Conquest, many Anglo-Saxon words were replaced by French equivalents. Some were central terms, such as *uncle*, which displaced A-S *sweostorsunu* 'sister's son' and

vegetable, which displaced A-S *wyrt*. However, the two processes are often related, since the arrival of new words through lexical change necessarily affects existing word-fields, causing semantic changes. Thus the modern verb *starve* used to mean' to die' in the form of A-S *steorfan*; our word *die* derives from Old Norse *deyja*, which displaced it. *Starve* has survived, but in a different sense. We shall see, especially in chapters 2 and 3, how this process is repeated many times.

In studying the changing structure of the vocabulary, a valuable lexicological distinction was made by the French scholar Georges Matoré between 'witness words' (*mots témoins*), reflecting material progress, and 'key words' (*mots clés*), reflecting ethical change (1953, pp. 65–8). Matoré's own example of the first category was *coke*, the by-product of coal, developed in the eighteenth century, and first recorded in 1669. We can usefully apply Matoré's distinction to more modern developments, where there are many examples. Witness words from recent decades are *sputnik* (1957), *video* (1958), *laser* (1960) and hundreds more reflecting technical advances. *Mots clés* are generally fewer in number and less easy to detect, but Matoré chose the emergence of *gentilhomme* in the course of the nineteenth century. From the English past there are the terms for such social types as the *rake*, 'a man of loose habits and immoral character; an idle, dissipated man of fashion', recorded from *c.*1653, the *beau*, 'a man who attends excessively to dress, mien, and social etiquette, a fop, a dandy', recorded from *c.*1687, and the *scab*, 'a strike-breaker, from *c.*1777. From recent times we can point to similar terms for such types as the *yuppie* (*c.*1982) and the *couch potato* (*c.*1976). A general term in this category is *weekend*, recorded from 1879 and clearly affecting the lifestyles of whole populations.

One must be cautious in assuming too direct a correlation between lexical change and social change. Not even all technical words are reliable winesses: thus *railway* is first recorded in 1776 (in Act 16 of George III), but the first railway (from Stockton to Darlington in the English Midlands) opened nearly fifty years later in 1825. *Helicopter* is first recorded in 1872, but the first helicopter flight took place only in 1907. *Baseball* is mentioned in chapter I of Jane Austen's novel, *Northanger Abbey* (1818), but it is obviously not the modern American game. Sometimes a word formulates an ancient practice: *contraception* is recorded from 1886, but there are oblique references to the practice as far back as the *Ancrene Riwle* ('The Rule for Nuns') in the twelfth century. A similar time-lag may be assumed to apply in terms such as *sadism* (1880), *masochism* (1893) and *security blanket* (1956).

Semantics itself is a comparatively new term, having been coined around the turn of the last century, derived from the Greek roots *sema*, 'a sign' and *semaino*, 'to mean', which have also yielded *semaphore* and *semiotics*. The term was given

31

special prominence when the French scholar, Michel Bréal, produced his classic study, the *Essai de Sémantique* in 1897, translated into English as *Semantics: Studies in the Science of Meaning* in 1900. However, the study of meaning obviously preceded the emergence of the term *semantics* by centuries, major contributions having been made by many philosophers and poets. Although the word originally carried the prestige of its classical origins, it has, regrettably, acquired increasingly negative overtones (outside professional use) in recent decades, becoming virtually a 'dirty word' implying linguistic obfuscation or dishonesty. An early recorded instance of this debased usage dates from 1944: 'The technique of character-assassination instead of argument is . . . standard totalitarian semantics.' Even within the profession, several major studies, such as those by Ogden and Richards (1923), Stern (1931), Lewis (1960), Waldron (1967) and Williams (1976), prefer the general term *meaning.*

Meaning is a highly complex subject, since it involves tacit understandings between users as well as overtly defined relationships between words and referents, and a symbiotic contract between individuals and groups within a given speech community. In the case of a world language like English, this relationship becomes more complex, involving a global community of users with regional conventions. Thus a term like *fanny* has quite different meanings and degrees of taboo in the United Kingdom, where it is an impolite slang term for a woman's genitals, whereas in America it is a fairly common slang term for the buttocks.

Despite these problems, it is remarkable how precisely speakers and listeners are able to isolate the intended meaning out of hundreds of available options, for instance, those attaching to the word *lost* in the following statements: 'He *lost* his pen'; 'He *lost* his life'; 'England *lost* the match'; 'He *lost* his temper and his way'; 'She *lost* her mind'. To these can be added the more American idioms 'He *lost* his cool', 'He's just *lost* it' and 'Tell him to get *lost*'. Just as diverse are the meanings of *just* in 'He's a *just* man'; 'He's *just* a man'; 'He's *just* the man'; '*just* as I was leaving, it started to rain'; '*just* listen to him!'; 'wouldn't you *just* like to give it all up?' Quantifying the meanings of common words, the American scholar G. K. Zipf arrived at the alarming statistic that 'Different meanings of a word will tend to be equal to the square root of its relative frequency' (1945, p. 255).

The examples just given are comparatively straightforward, but communication often seeks more indirect modes of euphemism or vagueness: people are described as 'financially embarrassed' rather than 'poor'; or 'experiencing some discomfort' rather than 'in pain', and so on. (We notice in both examples a marked change in lexis.) Delicacy and vagueness are often used of romantic or sexual matters, e.g. *affair*, *understanding*, *courtship*, *scene*, sometimes with disastrous results. In L. P. Hartley's novel *The Go-Between* (1953), the meaning of

spooning is withheld from the sexually innocent central character until the traumatic dénouement. Sometimes deliberate ambiguity is resorted to: when Benjamin Disraeli, a notable combination of Prime Minister and novelist, was sent unsolicited manuscripts, he allegedly replied: 'I shall waste no time in reading your manuscript.' This had one meaning for him but another for the would-be novelist.

The argument that meaning is not absolute or eternal, but simply conventional (philosophically termed *nominalism*) is highly plausible, particularly in an age such as ours in which semantic engineering by oligarchies is widely manifest. It has a long tradition, especially among philosophers. Ferdinand de Saussure, the great Swiss scholar, regarded as the founder of linguistics, made an axiomatic statement on 'the arbitrary nature of the sign' (1966, p. 67). However, there are semantic areas where the need for agreed and stable meanings is vital. Matters of law, especially those concerning constitutions and treaties, come immediately to mind. The same is true of scientific or technical language. These semantic areas require what is termed **referential** language, i.e. that which is neutral, factual and primarily **denotative**, the **denotation** of a word being its central or essential area of meaning.

In other domains, such as, say, political rhetoric and advertising copy, a higher degree of semantic licence, even of 'legitimate puffery', is condoned. Here a far more **emotive** kind of language is common, exploiting the favourable or negative **connotations** of words. **Connotations** are the associations, overtones and implications which exist in addition to a word's primary meaning. For example, among the synonyms for *thin* are *skinny* and *scrawny*, which have unfavourable connotations, as opposed to *slim* and *svelte*, which are favourable alternatives. Commenting on the current use of *gay*, Kingsley Amis observes: 'The word *gay* is cheerful and hopeful, half a world away from the dismal clinical and punitive associations of *homosexual*' (1997, p. 84). Connotations are commonly exploited for particular purposes. Thus a government report may note 'a shortage of *houses* and *electricity*', using referential and denotative language, whereas a bank or building society will prefer emotive language in persuading the public to purchase connotatively attractive *homes*. Likewise, an electricity company will claim to be selling, not referentially neutral *watts* and *ampères*, but favourable qualities such as *warmth*, *efficiency* and *comfort*.

The Historical Perspective

Looked at from a historical point of view, there emerges a clear relationship between parts of the word-stock and their denotations and connotations. A

notable feature of the English vocabulary is that the native register has a more emotive quality than does the borrowed classical element, which is more referential. Thus the following piece of enticing copy is pure Anglo-Saxon: 'Warm, rich and full of golden-goodness, Fido dog food will give your furry friend health, strength and get-up-and-go.' However, marketing aimed at giving a more scientific 'image' to a product uses more referential vocabulary, most of which came into the language after the Renaissance. In the following example, from a packet of breakfast cereal, the classical terms are in bold type:

> **Nutritionists estimate** that the body needs 30 **grams** of **fibre** each day to **assist** the **normal** working of the **intestine** by speeding up the **passage** of **waste** so that harmful **materials** are **eliminated** quickly. **Fibre** plays a **definite role** in the **prevention** of **obesity**. **Fibre** can also be **instrumental** in the **prevention** of heart **disease**.

In this example 20 words out of 55, or 36 per cent, are of classical derivation. Technical matters have the highest concentration of classical vocabulary, as is seen in the following definition of *plastic* (taken from the *OED*):

> Any of a **large** and **varied class** of **substances** which are **polymers** of a high **molecular** weight **based** on **synthetic resins** or **modified natural polymers** and may be **obtained** in a **permanent** or **rigid form** following **moulding**, **extrusion** or **similar treatment** at a **stage during manufacture** or **processing** when they are **mouldable** or **liquid**.

Here 26 words out of 54, or 48 per cent, are of classical derivation. These 'content analyses' show how clearly the native core makes up the basic 'nuts and bolts' of the language.

The distinction between referential and emotive language is valuable and illuminating. However, it is not absolute, but a matter of degree, since no variety of the language can be completely referential nor entirely emotive. The context of usage is often a primary defining factor. It is even possible for the same form of words to be referential or emotive: hence the statements 'Mr Jones is a bastard' or 'Mr Smith is a bugger' would be referential when used by a judge, but emotive when used by a layman. Contrariwise, classical terms which were originally neutral and academic, such as *phenomenal*, *sensational*, *categorical* and *absolute*, continue to be used in this fashion in philosophical textbooks, but can also be effectively exploited in an emotive fashion: 'Jones must be categorically condemned as a phenomenal liar and an absolute fraud.' The same word can have different denotations in the same statement: 'Granny and Grandpa used to have a beautiful home down in Sussex, but after Grandpa died, Granny

couldn't cope, so she's now moved into a very comfortable home not far from here.' The second 'home' is obviously a retirement or old folks' 'home'.

As has partly become apparent, the difference between native and classical terms is also shown in their degree of comprehensibility. Native terms are, generally speaking, semantically *transparent.* If one focuses on the core of the vocabulary, on words like *light*, *darkness*, *good*, *evil*, *strong*, *weak*, *hand* and *heart*, the basic meanings are obvious, even though they may be greatly diversified by metaphorical extension. Native terms form the language of first resort. In Spoken Word Counts, or analyses of natural conversation, the native content is very high. One such analysis (Jones and Wepman, 1961) yielded the following results: of the 200 most commonly used words, 83.5 per cent were Anglo-Saxon, 4.5 per cent were Old Norse (the closest Germanic relative), 10 per cent were from Latin via Old French and the remaining 1.5 per cent were from post-medieval Latin borrowings. As Dickens remarked fulsomely in an essay called 'Saxon English' in *Household Words*: 'When a man has anything of his own to say, and is really in earnest that it should be understood, he does not usually make cavalry regiments of his sentences, and seek abroad for sesquipedalian words' (1858, vol. 18). (*Sesquipedalian* is an example of itself, a ponderous polysyllabic alien formation. It was used by the Roman poet Horace to refer to words which were 'a foot and a half long', metrically speaking.)

The essentiality of the simple but flexible roots of the language was demonstrated artificially but dramatically through the invention of Basic English by C. K. Ogden in 1928. Using only the condensed core of the language, Ogden showed that with only eighteen verbs (curiously called 'operators') suitably combined with prepositions, a great diversity of meanings could be conveyed. Thus the simple combination *give up* can cover the meanings of a whole range of classical terms, such as *abandon, abdicate, abjure, cease, cede, desert, desist, discontinue, relinquish, renounce, resign, sacrifice, succumb, surrender, vacate, yield* and others. Without labouring the point, let us consider the remarkable range of common phrases which can be generated from the simple verb *do*, such as *do away with*, *do down*, *do for*, *do-gooder*, *do in*, *do-it-yourself*, *do or die*, *do up*, *do with* and *do without* (We should note, even-handedly, that while native terms have the advantage of flexibility, the classical have that of precision). The eighteen verbs of Basic English were: *come, get, give, go, keep, let, make, put, seem, take, be, do, have, say, see, send, may* and *will*. Of these all are Anglo-Saxon, except for *take*, from Old Norse.

The resonant clarity of the native word stock is apparent in many contexts, notably in the marriage service, formulated in the sixteenth century:

I take thee to my wedded wife/husband, to have and to hold, from this day forward, for better for worse, for richer for poorer, in sickness and in health, to

love and to cherish, till death us do part, according to God's holy ordinance, and thereto I plight thee my troth.

Here the words which stand out as being not native are the Latin *ordinance* and *according*; *poorer*, *part* and *cherish* are actually from Norman French, but have become part of the core.

The only unfamiliar native terms are those which have become archaic or regional. Thus *housecarle*, *folkmoot* and *wapentake*, originally central terms concerning the social structure of Saxon times, are now obsolete, as is *swain*. Similarly, *thole* 'suffer', *bairn* 'child', *urchin* 'hedgehog' and *ken* 'know' are now regional survivals of what were originally central words.[9]

As we have seen, historically the varieties of register clearly reflect the class distinctions of their origins: the Anglo-Saxon terms tend to be those of the common people; the Norman-French overlay came from the Norman overlord, while the Latin and Greek terminology derived from a scholarly elite. The native terms tend to be short, blunt, emotive and direct; the Norman French tend to be imbued with courtliness and refinement, while the Latin or Greek are abstract and bookish, or technical and precise.

Although this broad categorization is sound, there are interesting historical complications, whereby words may change register over time, especially if they change shape or spelling. Consider the following high register terms *vis-à-vis* their subsequent *déclassé* descendants:

physiognomy (ME)	phiz (1688) ⇒ fizz (face)
lunatic (ME)	looney (1872)
fanatic (1553)	fan (1889)
obstreperous (1600)	stroppy (1951)
perquisites (1565)	perks (1869)
acute (1570)	cute (1868)
demonstration (1668)	demo (1963)
pornography (1864)	porno (1970)

Furthermore, the relationship between register and class is complex in its historical development. The amusing distinction that 'Horses sweat, men perspire, but ladies only glow' rests on a prescriptive decorum of style, rather than fact. The notion that the upper classes use exclusively high-register language is simplistic. Queen Elizabeth I 'swore like a man', we are told, and there is a great tradition of aristocratic swearing, vituperation and insult: this mode has even been developed into the art form of ritual insult known as flyting.[10]

The distinction between 'U' and 'Non-U', initially formulated by A. S. C. Ross in a scholarly journal in Finland in 1954, is an aspect of register which has received much publicity in the past few decades.[11] Ross demonstrated, albeit in

a tellingly impressionistic fashion rather than one based on documented academic research, that 'U' (upper-class speakers) prefer plain low register rather the supposedly genteel evasions or euphemisms preferred by the bourgeoisie. Among the examples he gave ('U' terms first) were: *bog/toilet*, *mad/insane*, *rich/wealthy*, *false teeth/dentures*, *die/pass on*, *sweat/perspire* and *what?/pardon*. Others are simply conventional, namely *vegetables/greens*, *salt and pepper/cruet*, *napkin/serviette* and the naming of meals. Only *breakfast* is shared by all classes: thereafter the sequence for 'U' eaters is lunch ⇒ tea ⇒ dinner, while for the non-'U' it is dinner ⇒ tea/supper ⇒ high tea.

Ross's insights were popularized in a humorous collection of essays edited by Nancy Mitford under the title of *Noblesse Oblige* in 1956 and Ross himself edited a simiar symposium called *What are U?* in 1969. In a subsequent collection *U and Non-U Revisited*, edited by Richard Buckle, he observed that 'the antitheses between U and Non-U have *not* changed' (1980, p. 28). Linguistic class-consciousness was a necessary aspect of *The Official Sloane Ranger Handbook* (1982), by Ann Barr and Peter York. This glossary indicates the same preference in this upper-class type for the blunt or direct word in preference to the vaguer euphemism, commonly derived from the higher registers. They listed *bumph* for 'paper', *fuck-up* for 'organizational disaster', *pissed* for 'drunk', *bollock* for 'ball or social gathering', *poncy* for 'effeminate or aesthetic', *thunder box* for 'lavatory', *pong* for 'an unpleasant smell', *spastic* or *thick* for 'stupid' and *bin* for 'mental hospital'. In this milieu there is also a tendency for baby language to be used, as in: 'Mummy is dotty about this silly old colonel: he's a real sweetie, but he's always losing lolly on the geegees.'

It is a notable feature of sub-culture languages that they are made up of specialized use of low-register general terms rather than specific words. For instance, underworld argot includes *pig*, *mole*, *mule*, *hit*, *pinch (vb)*, *lift (vb)*, *shop (vb)* and *heat*, used in specialized senses. Similarly, the drug 'scene' uses *coke*, *pot*, *acid* and *high*. A single word may be used in both milieux: thus *grass* means both 'an informer' and 'cannabis', while *crack* means both 'to break in' and 'cocaine'. Among homosexuals a special currency of the words *gay*, *pink*, *cottage*, *drag*, *cruise* and *queen* is prevalent. In Black street slang in the United States, *bad*, *mean* and *wicked* have precisely opposite meanings.

We must distinguish between what is the natural evolution of registers within semantic fields through social conventions on the one hand, and the exploitation of particular registers for various motives. Variation in register is frequently manipulated to establish authority. 'If you do that again, I shall hit you' is an example of low-register, direct personal style. A public warning, on the other hand, would not normally be framed in such terms, but would employ high-register abstractions and an impersonal mode: 'Infringement of these regulations will result in prosecution.' In recent times much has been

written, notably by George Orwell, on the deliberate exploitation of opaque, high-register Latinization for the purposes of evasion, deceit and propaganda. Terms like *liquidation*, *operation*, *incident* and *elements* are useful as 'anaesthetic' variants of *murder*, *invasion*, *riot* and *people*. C. S. Lewis appositely observed in 1942 'Once we killed bad men; now we liquidate undesirable social elements.' This theme forms a substantial part of chapter 7.

Register and Specialization

In various technical fields, there is a division of register clearly reflected between the general and the specialist term. Even in what we may call the general vocabulary, the base-term will commonly be Saxon, but the descriptive adjective will be classical:

finger	digital
ear	aural
mouth	oral
hand	manual
tooth	dental

Professional language is especially marked by the use of the higher register. For instance, in legal language:

theft	larceny
beat	assault
burn down	arson
crime	felony/delict

Thus the demotic idiom of being 'caught red-handed' is translated in legal jargon as *in flagrante delicto*. The same division applies in the fields of science:

hole	orifice/cavity
speed	velocity
force	intensity
size	volume

It is particularly apparent in medicine:

bleeding	haemorrhage
wound	laceration
skull	cranium

Anglo-Saxon	Middle English	Renaissance	Augustan	Victorian	Modern
shit(n) turd	**ordure**	**excrement**		crap*	**defecation**
	piss (v)	**urinate**	**micturate**[1]		pee
sleep with swive	fuck*	**copulate** screw	make love	shag	bonk
	pollution frig[2]	onanism **digitation** / **self-abuse masturbation**			wank
arse	bum* buttocks	**fundament anus** / **posterior**(s)	bottom		
	cunt thing[3]	**coney pudendum**	twat*	**vagina**	quim*
weapon[4] cock	yard	tool	prick	**penis**	(**privy**) **member**

Notes: **Bold** type indicates Romance origin
  * Origin uncertain
  1 'The sense is incorrect as well as the form' (*OED*)
  2 *Frig* overlapped with *fuck* in the seventeenth and eighteenth centuries
  3 *Thing* has served for both male and female genitalia since Middle English
  4 OE *wæpened* ('weaponed' or 'armed') has the basic sense of 'male' in many compounds including gender in children and plants

Figure 1.7   The semantic field of 'rude' words

sweat	perspiration
heart attack	cardiac arrest

Thus in medical jargon the opaque, high-register statement that 'The patient is experiencing a potentially fatal haemorrhaging episode' means basically: 'The patient is bleeding to death.'

In essence, we now have two vocabularies for our bodies, flora and fauna: an 'outer' vocabulary made up of common words, and 'inner' semantic fields comprising classical anatomical and botanical terminology. The development of medical and legal professional language is discussed in more detail in chapter 4.

In the area of sexual vocabulary, the separation of registers is so strong that there are no neutral, generally acceptable terms in English for the genital organs and sexual activity. As C. S. Lewis put it trenchantly, 'as soon as you deal with [sex] explicitly, you have to choose between the language of the nursery, the gutter and the anatomy class' (in Tynan, 1975, p. 154). Figure 1.7 depicting 'the semantic field of rude words' shows the evolution of the different registers. The field shows two important shifts over time, first a move from plain native terms such as *shit* and *arse* (from Anglo-Saxon) to abstract classicism in the Renaissance and Augustan periods, and then a reversion to plain terms like *bonk* and *wank* in recent times. However, the notion that terms now regarded as grossly impolite, coarse or obscene have always been taboo is

not valid. The first recorded usage of the word *cunt* is in a medieval London street-name, *Gropecuntlane*, which would not pass muster in a modern borough. Many ancient proverbs are full of racy, naughty terms, to be discussed in chapter 3.

Today, the general division between Saxon and Classical is not absolute. In many semantic fields both options are available. Thus we have *equine* for horses, but *horsey* for people, *catty* for people but *feline* for cats; but both *piggish* and *porcine* are for people. The context is paramount, a point made in a memorable moment in W. H. Auden's poem, 'Musée des Beaux Arts', on a classic painting of the fall of Icarus by Brueghel. Commenting on the irony that Icarus drowns close to land but quite unnoticed, Auden writes:

> The dogs go on with their doggy life

deliberately using the low-register word, so unphilosophical and untragic, evocative of their purely physical existence of dozing, sniffing and scratching for fleas. Obviously *canine* would be unsuitably scientific. A wonderful contrary instance is the use of *liquefaction*, an apparently scientific term, in a sensual lyric by Robert Herrick ('Upon Julia's clothes', 1648):

> Whenas in silks my Julia goes,
> Then, then, how sweetly flows
> The liquefaction of her clothes.

Herrick uses *liquefaction* in a literal sense of 'flowing like water'.

Registers can be exploited or transposed for humorous or facetious effects. Thus the familiar nursery rhyme 'Twinkle, Twinkle, Little Star' is, expectedly, in almost pure native register, with *diamond* the sole classical term in this simple ditty. However, when facetiously transposed into an artificial classical idiom, the text comes out as:

> Scintillate, scintillate, globule vivific
> Fain would I fathom thy nature specific
> Loftily poised in the ether capacious
> Greatly resembling a gem carbonaceous

Now fourteen out of twenty-two words are of classical origin, making the piece abstract and opaque. The alien quality is also accentuated by the inversions and the contrast between the archaic idiom of 'Fain would I fathom thy' and the scientific terminology of 'nature specific'. In similar vein, Sir Arthur Quiller-Couch wrote a fine parody of Hamlet's soliloquy 'To be or not to be' in ponderous officialese:

To be or the contrary? Whether the former or the latter would be preferable would seem to admit of some difference of opinion; the answer in the present case being of an affirmative or of a negative character according as to whether one elects on the one hand to mentally suffer the disfavour of fortune, albeit in an extreme degree, or on the other to boldly envisage adverse conditions in the prospect of bringing them to a conclusion. (In Nash, 1993, p. 94)

Turning to more familiar publicly mediated language, the registers employed in modern journalism vary according to the sector of the market a particular publication is aiming to reach. The popular and tabloid press tends to use short, highly emotive, low-register terms like *slam*, *slate*, *blast* and *ban*, while the 'quality' press, the serious or responsible press, uses a more sober style. Curiously, even the popular sector occasionally resorts to archaisms such as *slain*, *yule*, *agog* and *scribe*.

Classical terms are often semantically *opaque* to native speakers, especially to those unfamiliar with their roots. Thus *otorhinolaryngology* is meaningless to a person who does not perceive that it is made up of four Greek elements, namely *ot-*, 'ear', *rhin-*'nose', *laryng-*'throat' and *logos* 'word' used in the sense of 'the study of'. The point of opacity in classical terms is easily demonstrated by their concentration in that area of comic semantic error called **malapropism**. This term derives from French *mal à propos* ('inappropriate'), but became part of the language through a humorous stage character created by Sheridan in his play *The Rivals* (1775). Mrs Malaprop blundered into such verbal gaffes as '*allegories* on the banks of the Nile' (for *alligators*) and *hydrostatics* (for *hysterics*). However, this mistaken use of classical terms had been previously exploited centuries earlier by Chaucer, Langland and Shakespeare.[12] While their malapropisms are often contrived as a comic device, there is a genuine core of incomprehension at the heart of the phenomenon. This has been explored in studies such as *The Lexical Bar* (1985) by David Corson and 'Elaborate and Restricted Codes' (1971) by Basil Bernstein.

Because of their alien and often mystifying quality, classically derived terms have invoked hostility from writers and commentators on the language. We shall see in chapter 4 that one such period of opposition was the Inkhorn Controversy in the sixteenth century, when classical borrowings were seen by some as bankrupting and weakening the healthy native stock of the language. However, in the eighteenth century, classical borrowings were regarded more favourably, and figure largely in the vocabularies of Gibbon and Dr Johnson, who showed such a penchant for them that he included in his *Dictionary* such oddities as *excubation* ('the act of watching all night'), *ebriosity* ('habitual drunkenness'), *effosion* ('the act of digging up from the ground') and *enecate* ('to kill; to destroy'). Others can be seen in figure 5.2. In recent decades the old

hostility has revived. It is well exemplified in George Orwell's much-quoted prescriptions: 'Never use a long word where a short one will do; never use a foreign phrase, a scientific word or a jargon word if you can think of an everyday English equivalent' (1958, p. 88).

Amidst lexical diversity, the concentration of Anglo-Saxon in the core vocabulary is equally dramatic, bearing out Murray's observation quoted earlier. It is this central core 'whose Anglicity is unquestioned' which forms the nucleus of everyday speech, as is demonstrated in analyses such as the Spoken Word Count, previously mentioned. More remarkably, a similar analysis applied to the diverse vocabulary of literary authors (by Professor O. F. Emerson about a century ago) yields similar results:

King James Bible	–	94%	Pope	–	80%
Shakespeare	–	90%	Johnson	–	72%
Spenser	– –	86%	Hume	– –	73%
Milton	– –	81%	Gibbon	–	70%
Addison	– –	82%	Macaulay	–	75%
Swift	– –	75%	Tennyson	–	88%

<div align="center">(In Wood, 1969, p. 47)</div>

The percentages are remarkably high, especially in writers like Milton, Gibbon and Dr Johnson, who had a penchant for the recherché or esoteric term. Of course, there will be variations within an author's work depending on theme or topic. But the durability, indeed the tenacity of the Anglo-Saxon core thus remains a permanent feature of the language, written as well as spoken, no doubt because it is the register of immediate recourse. The concluding chapter contains further discussion and more recent analysis of this aspect.

However, as we have already seen, classically derived words have their place, notably in technical and abstract semantic areas. Where would we be without words like *quantity*, *quality*, *procrastination*, *maturity?* Native equivalents, like *muchness*, *suchness* and *ripeness* are either quaint, unmeaningful or already bespoken. *Procrastination* has no true native synonym and requires a cumbersome paraphrase or translation such as: 'the deferment of action'. The same is true of *simultaneous*, *synchronize* and *atmosphere*. One thinks, too, of the power of the classical register evidenced in Disraeli's devastating denunciation of his rival, Gladstone as 'a sophistical rhetorician inebriated with the exuberance of his own verbosity'. A translation into the native equivalents, 'A dishonest public speaker drunk with the sound of his own longwindedness' manifestly does not have the same effect. A high proportion of classical terms occurs in recently coined scientific vocabulary, which has generated terms such as *spectrophotometer*, *teleroentgenography*, *transpepidation*, *tropomyosin*,
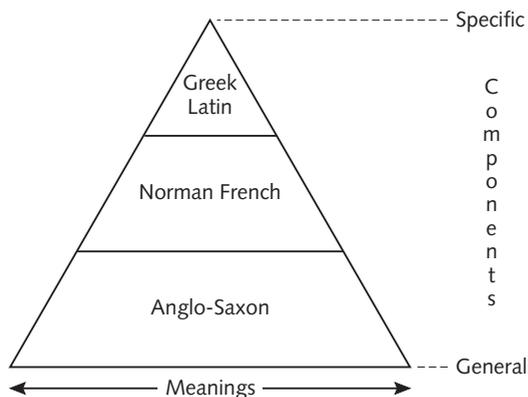
Figure showing a triangle/pyramid divided into three horizontal sections labeled from top to bottom: "Greek / Latin", "Norman French", "Anglo-Saxon". To the right, vertical text reads "Components" (spelled vertically). The apex is marked "Specific" (dashed line) and the base "General" (dashed line). The base has a horizontal double arrow labeled "Meanings".

Figure 1.8   Lexical origins and range of meaning

zochlorella and thousands more. The most most astonishingly gargantuan specimen is *pneumonoultramicroscopicsilicovolcanoconiosis*, a facetious word coined by F. Scully in 1936. While these are undoubtedly opaque and alien to the non-specialist, we should see them not simply as strangers, but as extensions of a classical word-stock of basic technical terms, such as *theory*, *method*, *pressure*, *motion* and *reaction*. The inestimable advantage that English has gained is its lexical richness, evidenced in virtually every word field. Consider the amazingly diverse (if morbid) possibilities of this range: *deadly*, *killing*, *lethal*, *mortal*, *fatal*, *deathly*, recently joined by *life-threatening*.

We may sum up the relationship between lexical origin, range of meaning and comprehensibility by means of two schematic figures, namely figure 1.8 'Lexical origins and range of meaning' and figure 1.9 'Lexical categories and comprehensibility'.

Taboos and Euphemisms

We have noted many examples of the perennially close relationship between social and lexical change. A large, complex and changing field concerns **taboos** and **euphemisms**. *Taboo* is a fairly recent borrowing, having been brought back from Polynesia by Captain Cook in 1777. It originally concerned areas of human experience which were sacred and therefore prohibited; it now refers to that which is unmentionable because it is ineffably sacred or unspeakably vile. **Euphemism** refers to the use of deliberately indirect, conventionally imprecise or 'comfortable' ways of referring to taboo or unpleasant topics.
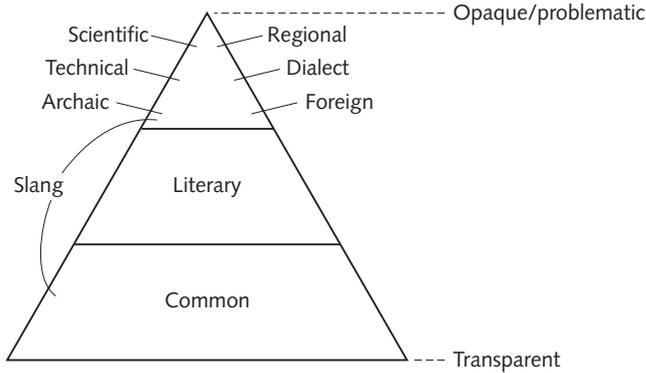
Figure 1.9   Lexical categories and comprehensibility

Euphemism shows that interference in the semantic market is not always deliberate, since taboo is frequently unconscious and collective in its effects.

In origin euphemism is profoundly concerned with word-magic, that primitive belief that there is a mystical relationship between words and things. The essential mode of euphemistic expression is thus indirect, and as the etymology of the term in Greek *eu* ('well) and *pheme* ('to speak') suggests, the motive is to describe the situation as better than it is, or to avoid the taboo area, thereby pacifying some dreaded force by managing not to offend it. Interestingly, one of the most striking examples is the ancient Greek use of the term *Eumenides*, meaning literally 'the friendly ones', to refer to the Furies. In many European languages the weasel, a bloodthirsty and ferocious creature, is called by a variety of pacifying names, such as 'little beauty' or 'little lady' (Ullmann, 1951, p. 77). Within the Christian frame of things one notes similar titles of respect for the Devil, such as *Old Nick*, *the Prince of Darkness*, *Lord of the Flies* and so on.

Despite the exotic origins of the word *taboo*, the notion of things sacred and unmentionable occurs at every level of civilization and in all kinds of environments. Feared or prohibited semantic areas vary greatly, including the name of God, reference to death, disease, madness, being crippled, as well as such common aspects of physicality as copulation, the genitalia and the varieties of excretion, even the most trivial of embarrassments, which in some societies include references to underclothes and humble occupations.

In order to avoid these areas of taboo it becomes necessary to adopt a disguise mechanism. In extreme cases the subject is avoided entirely. Thus no major English dictionary included the two most egregious 'four-letter' words, *fuck* and *cunt* between 1728 and 1963, a point which is taken up in chapter

5. The more common mode is less conscious and involves using a euphemism. Some of these allude explictly to the inexpressible quality of the taboo object. Thus in Victorian times it was taboo to mention *trousers*, with the consequence that terms such as *inexpressibles*, *indescribables*, *inexplicables* and other comic evasions were forced into service.

One mode of avoiding the embarrassingly direct is via a metaphorical route. Death and sex, being universal areas of euphemism, provide many examples. Death is referred to, not as a final shocking or peaceful state, but as a metaphorical journey, in comforting variants and traditional forms such as *passing away*, *passing on*, *going to one's Maker*, *in Abraham's bosom*, *joining the majority* and the speciality of the Salvation Army, *promoted to glory*. Similarly, we note that to *sleep with*, *go to bed with*, *make love*, *make out* are socially acceptable allusions in modern times to sexual intercourse. Of course, the degree of explicitness of these formulas has not remained constant historically. Until the early part of this century to *make love* meant roughly the same as 'to flirt'. In earlier times *lover* could mean 'friend': in Shakespeare's *Julius Caesar* (1599) Brutus refers to Caesar as 'my best lover' (III. ii. 45), with no sexual sense implied, but such an allusion has a very different meaning today.

These metaphorical modes existed as far back as Anglo-Saxon times. Thus the *Anglo-Saxon Chronicle* annal for the year 588 records: 'In þissum ჳeare King Ælle forþeode' ('In this year King Ælle died,' literally 'passed on, went forth'). Similarly, the attempted seduction of Joseph by Potiphar's wife in Genesis 39 : 7, translated in the King James Bible (1611) as 'his master's wife cast her eyes upon Joseph and said "Lie with me",' was rendered by Ælfric in the year 1000 in a more modern idiom: 'His hlæfdige lufode hine and cwæþ to him Slap mid me' ('His lady loved him and said to him "Sleep with me" '). Incidentally, in the King James version, instead of the direct verb 'loved' we find the phrase 'cast her eyes upon Joseph', and note the similarity to the modern idiom of seduction 'to make eyes at', one of several suggestive but imprecise phrases, such as 'to make a pass at'.

Metaphors, it should be noted, can all too easily lead to the opposite of euphemism, namely **dysphemism**, which is the startlingly direct, low-register and shockingly coarse violation of the taboo. Thus dysphemisms for death include *pushing up daisies*, *snuff it* and *croak* which allude, with a large measure of black humour, to the physicality of dying, including the death-rattle and reincorporation into the cycle of nature. In similar vein, in American underground slang, a corpse is a *stiff*, while to go to the electric chair is jauntily translated as 'to take a hot squat', and in Army parlance comrades may jokingly say that they will be 'back in a jiffy' i.e. in a body bag.

The major alternative euphemistic strategy is to use a high-register, abstract term. Thus in the vocabulary of death *undertaker* took on its funereal function

around 1700, but has since lost its euphemistic sense and become the stand-ard term in British English, whereas in American English the high-register equivalent *mortician* has been established for over a century. Similarly *coffin*, which originally meant in French a small basket, has become the direct term in British English, being replaced by the more elegant and precious euphemism *casket* in American English. Referring to the dead person remains problematic in British English, since *corpse* and *body* are uncomfortably direct; the Ameri-can equivalent, *the loved one*, is ideally vague and comforting, if a little senti-mental. In the sexual area, terms like *intercourse* and *intimacy* have lost their original euphemistic quality; similarly, *assault* is often preferred to *rape* despite its lack of precision, and *interfere with* often does service for *molest* or *abuse sexually*.

As we have noted in the discussion on semantics, the strong, direct and emotive terms in English are predominantly drawn from the Anglo-Saxon base register: among them are *murder*, *kill*, *steal* and *lie*, terms which would not be acceptable in diplomatic language. By far the most common register of recourse for euphemisms is the classical component, well described by Edward Gibbon in the eighteenth century as supplying 'the decent obscurity of a learned language' (1854, p. 212). From this source we derive *perspiration*, *uri-nation*, *micturition*, *expectoration*, *defecation*, *copulation* in the area of 'bodily functions' (itself something of a euphemism) and *elimination*, *extermination* and *liquidation* in the semantic disguise of death. Some areas are surprisingly well-stocked in classical variants: *bad breath*, for example, is supplemented by *halitosis* and *pyorrhea*. As the explicit terms for sexual activity became unac-ceptable and then taboo, so numerous Latinized variants were drawn in to the semantic vacuum. Among them were *rape* (1482), *consummation* (1530), *seduce* (1560), *erection* (1594), *copulation* (1632), *orgasm* (1684), *intercourse* (1798), *climax* (1918) and *ejaculation* (1927). In other cases the classical term does not have a socially acceptable native equivalent: hence *fellatio* and *cun-nilingus*. French supplies a useful array of euphemisms in *accouchement* for 'being brought to the labour bed' and *derrière* for 'backside'.

Classically taboos lead to another kind of disguise-mechanism, the use of what are called 'minced forms' of taboo words. We are familiar with self-conscious forms such as 'the f-word' for *fuck* and *bleeding* for *bloody*. However, many historical forms no longer reveal the original link with the taboo terms to the majority of users. Among them are *gosh*, *golly* and *gad* for *God*, *Jiminy*, *Jeez* and *Jeepers* for *Jesus* and *effing* and *frigging* for *fucking*. Figure 1.10 deals with the astonishing variety of euphemistic mutilations of the name of God, showing that the process has been fairly continuous since the Middle Ages, but received a marked boost in the period 1598–1602, when Puritan injunctions against profanity on the stage produced such forms as *'snails* for *God's nails* and

Term	Date	Euphemism
God	1350s	gog
	1386	cokk
	1569	cod
	1570	Jove
	1598	'sblood
	1598	'slid (God's eyelid)
	1598	'slight
	1599	'snails (God's nails)
	1600	zounds (God's wounds)
	1601	'sbody
	1602	sfoot (God's foot)
	1602	gods bodykins
	1611	gad
	1621	odsbobs
	1650s	gadzooks (God's hooks)
	1672	godsookers
	1673	egad
	1695	od
	1695	odso
	1706	ounds
	1709	odsbodikins (God's little body)
	1728	agad
	1733	ecod
	1734	goles
	1743	gosh
	1743	golly
	1749	odrabbit it
	1760s	gracious
	1820s	ye gods!
	1842	by George
	1842	s'elpe me Bob
	1844	Drat! (God rot!)
	1851	Doggone (God-damn)
	1884	Great Scott
	1900	Good grief
	1909	by Godfrey!

Figure 1.10   Euphemisms for the name of God

zounds for *God's wounds*. One of the linguistic ironies of life is that these euphemisms often become fashionable.

One of the curiosities of euphemism, indeed of the English language, is the development in the eighteenth century of the code language known as Cockney Rhyming Slang. This exemplifies a highly developed 'disguise mechanism' in witty and ingenious coded formulas where the last term rhymes with the intended word: thus *trouble and strife* alludes to *wife*, *brass tacks* to *facts* and *loaf of bread* to *head*. These are, of course, not especially euphemistic and have passed into general usage, so that their origin is often not realized. But in the grosser provenance there are such established forms as *Richard the Third* for

turd, *Bristol Cities* for *titties*, *Hampton Wick* for *prick*, *Khyber Pass* for *arse*, *cobbler's awls* for *balls* and *Berkshire Hunt* for *cunt*. Of these the abbreviated forms *wick* and *berk* have passed into common slang, while *bristols* and *cobblers* are commonly used in the original speech community. Indeed, *cobblers* is often used in the broad sense of 'rubbish'. In the television series 'Porridge' (a slang term for 'prison'), a character says 'I sees through all that Home Office cobblers' (1990, p. 19). A rarer variety of slang is back slang, of which *yob* for *boy* is the only commonly used term.

Even more interesting historically are the substitute terms which have no obvious relationship with the taboo word. For example, *donkey*, which one would expect to be a common word found in the earliest stages of the language, is actually first recorded only in 1785. The traditional synonym, *ass*, had been in the language since Celtic times, and was the natural term in Scripture, proverb and folklore. However, in the eighteenth century the word started to fall into disrepute through an uncomfortable proximity to *arse*, so that the lexicographer Francis Grose observed that 'a lady who affected to be extremely polite and modest would not say *ass* because it was indecent.' Thus *donkey*, a dialect word, moved into the lexical gap.

Similarly *rabbit*, recorded from the fourteenth century, replaced the old word *coney* when it started to develop too close a relationship with *cunny*, meaning *cunt*, illustrated in jaunty verses such as 'All my Delight is a Cunny in the Night' (1720). A similar syndrome is apparent in American English, where *cock* has traditionally been replaced by euphemisms and substitutions: hence *rooster* for the famyard fowl, *faucet* for *cock* in the sense of 'tap', and the emasculated form *roach* for *cockroach*. (By contrast, in British English *cock* has never been a taboo term and is found in dozens of compounds, notably *cock-horse*.) The essential distinguishing point about all these substitutions is that they were collective, and presumably unconscious developments within the speech communities, not deliberate interventions by individuals.

There is, of course, an important distinction between those euphemisms which appear to be natural, unconscious and universal currency in most cultures, and those which are contrived (usually by oligarchies) to conceal politically unpalatable truths. George Orwell is especially associated with the trenchant exposure of such cynical evasions, now captured in his enduring coinages Newspeak and Doubletalk, discussed further in chapter 7. Although we tend to think of this last category as being an aspect of modern propaganda, the Roman historian Tacitus succinctly exposed the process in the first century in the sardonic quotation: 'They make a wilderness and call it peace' (*solitudinem faciunt pacem appellant*; *Agricola*, chapter 30).

However, these particular instances have proved surprisingly enduring, since it is generally the fate of euphemisms to become too closely associated

with the taboo area, to become tainted and thus to require replacement. The considerable turnover of euphemisms naturally causes imbalance in various semantic fields as new replacements are brought in. In their classic slang thesaurus, *Slang and its Analogues* (1890–1904), Farmer and Henley show that no less than 600 bawdy phrases have been coined over the centuries for the sexual act. A small sample is given in figure 1.11, under the odd headword *Greens*. The dynamic between the taboo term *fucking* and related euphemisms, with dates of their first recorded usages, is shown in figure 1.12. As can be seen, the process has been going on for centuries.

In the underground slang of two centuries ago *commodity* was defined in an amusingly trenchant fashion as 'the private parts of a modest woman and the public parts of a prostitute'; *corinthians* were 'frequenters of bawdy houses', and a *commons* was 'a necessary house', i.e. a lavatory. These have all passed away, as euphemisms tend to do. Others prove surprisingly resilient: in the same period *hump* was 'a fashionable word for copulation' while *screw* was a plainer term for the same activity.

Western society has since added other areas of taboo, such as matters of race, financial collapse, poverty, going to prison, even trivialities which include fatness and shortness. Hence terms like *ethnic* for *racial*, *coloured folk* for *blacks*, *technical correction* for *crash*, *recession* for *slump*, *financially underprivileged* for *poor*, *choky*, *clink*, *slammer* and *nick* for *prison*, *dustman* for *rubbish collector*, *vegetable executive* for *greengrocer*, *vertically challenged* for *short*, *possessing an alternative body image* for *fat* and *substance abuse* for *drug addiction*. These are clearly more conscious, indeed highly contrived, part of the explicit agendas of political correctness, discussed in chapter 8.

Word-Formation

Most native speakers of a language are not consciously aware of word-formation. It would probably have to be pointed out to them that in the previous sentence the form *native* has a Latin root *nat-* which comes from the verb 'to be born' and a suffix *-ive* which indicates an adjective function (as in *adjective* itself); that *speakers* comes from a root *speak*, an agent suffix *-er* and a plural inflexion *-s*; that *consciously* has a Latin verbal root *scio* 'I know', with a collective prefix *con-* and an adverbial suffix *-ly*, and that *word-formation* is a compound. We are more aware of new formations like *disinformation*, *malfunction*, *unbundle* and *animatronics* than the well-worn, familiar words which we use every day. Since part of chapter 7 will deal with the more recent varieties of word-formation, this section will cover the more traditional.

GREENS, *subs.* (old).—1. Chlorosis: *i.e.*, the green sickness.

1719. DURFEY, *Pills*, etc., i., 313. The maiden takes five, too, that's vexed with her GREENS.

2. *in. pl.* (printers').—Bad or worn out rollers.

TO HAVE, GET, or GIVE ONE'S GREENS, *verb phr.* (venery).—To enjoy, procure, or confer the sexual favour. Said indifferently of both sexes.

Hence, also, ON FOR ONE'S GREENS=amorous and willing; AFTER ONE'S GREENS=in quest of the favour; GREEN-GROVE=the pubes; GREEN-GROCERY = the female *pudendum*; THE PRICE OF GREENS = the cost of an embrace; FRESH GREENS=a new PIECE (*q.v.*). [Derived by some from the old Scots' *grene*= to pine, to long for, to desire with insistence: whence GREENS = longings, desires; which words may in their turn be referred, perhaps, to Mid. Eng., *zernen*, A.S., *gyrnan*, Icelandic, *girna*=to desire, and Gothic, *gairns*=desirous. Mod. Ger., *begehren*=to desire. *See* DALZIEL, *Darker Superstitions of Scotland*, 1835, p. 106:—'He answered that he wald gif the sum Spanyie fleis callit cantarides, quhilk, gif thou suld move the said Elizabeth to drynk of, it wold mak hir out of all question to GRENE eftir the.' *Trial of Peter Hay, of Kirklands, and others, for Witchcraft, 25th May, 1601*. But in truth, the expression is a late and vulgar coinage. It would seem, indeed, to be a reminiscence of GARDEN (*q.v.*), and the set of metaphors—as KAIL, CAULIFLOWER, PARSLEY BED, and so forth (all which *see*) —suggested thereby.]

ENGLISH SYNONYMS.—TO BE all there but the most of you; in Abraham's bosom; up one's petticoats (or among one's frills); there; on the spot; into; up; up to one's balls; where uncle's doodle goes; among the cabbages.

To DANCE the blanket hornpipe; the buttock jig; the cushion dance (*see* MONO-SYLLABLE); the goat's jig; the mattress jig; the married man's cotillion; the matrimonial polka; the reels o' Bogie (Scots'); the reels of Stumpie (Scots'); to the tune of THE SHAKING OF THE SHEETS; with your arse to the ceiling, or the kipples (Scots').

To GO ballocking; beard-splitting; bed-pressing (Marston); belly-bumping (Urquhart); bitching (Marston); bum-fighting; bum-working; bum-tickling; bum-faking; bush-ranging; buttock-stirring (Urquhart); bird's-nesting; buttocking; cock-fighting; cunny-catching; doodling; drabbing; fleshing it; flesh-mongering; goosing: to Hairy-fordshire; jock-hunting; jottling; jumming (Urquhart); leather-stretching; on the loose; motting; molrowing; pile-driving; prick-scouring; quim-sticking; rumping; rump-splitting; strumming; twatting; twat-faking; vaulting (Marston, etc.); wenching; womanizing; working the dumb (or double, or hairy) oracle. twat-raking; tummy-tickling; tromboning; quim-wedging; tail-twitching; button-hole working; under-petticoating.

TO HAVE, or DO, A BIT OF beef (of women); business

Figure 1.11 The profusion of copulatory synonyms: a page from Farmer and Henley's *Slang and its Analogues* (1890–1904)

footering
c.1753

footling
1905

foutra
c.1592

FUCKING

effing
c.1929

foot
c.1600

frigging
c.1785

Figure 1.12   The dynamic between a taboo term and related euphemisms

As we have seen in the previous survey, English has a hybrid vocabulary. This applies not just to the individual words, but to the elements which are used in word-formation. Thus a basic word like *beautiful* is French + English; *musical* is Greek + Latin; *bureaucracy* is French + Greek and *tarmacadamization* is really polyglot, English + Celtic + Hebrew + Greek + Latin. In earlier times there was a purist view that linguistic hybrids of this sort should be avoided: thus *sociology*, coined in French by Auguste Comte, was condemned as being 'barbarously termed' shortly after its appearance in English in 1843 (Potter, 1975, p. 92). Today it is entirely acceptable, as are hybrids like *unscientific*, *doable*, *antifreeze* and so on.[13]

It may be of use to supply the student with a concise guide of the more common elements which indicate these etymological origins (Table 1.1). They are arranged alphabetically, with the less familiar elements being glossed.

As was stressed, this is a concise guide. When the Revd Walter W. Skeat, a great Victorian philologist, produced his *Concise Etymological Dictionary of the English Language* in 1882, he listed five pages of prefixes, but declined to do the same for suffixes, saying that 'The number of suffixes in Modern English is so great, and the forms of several, especially the words derived through the French from Latin, are so variable, that an attempt to exhibit them all would tend to confusion' (p. 630). Laurence Urdang has published a useful collection, *Suffixes* (1982), over 250 pages long.

The most obvious feature of the list is the great proportion which come from Latin and Greek, an aspect studied by Donald M. Ayers (1986) and Richard M. Krill (1990). These elements feature largely in new words like *semi-conductor*, *microchip*, *sociopath*, *superhighway* and *synergy*. A number of these prefixes have developed their own new senses as individual words, such as *retro*, *techno*,

Table 1.1   Prefixes and suffixes

Prefixes/beginnings	Suffixes/endings
acc- Latin/French	*-able*, *-ability* Latin/French
aero- Latin	*-age* French
ante- Latin (before)	*-ance* French
anthrop- Greek (mankind)	*-ant* Latin
anti- Latin (against)	*-archy* Greek (rule)
arch- Greek (chief)	*-arian* Latin
astro- Greek (star)	*-ary* Latin
audio- Latin	*-centric* Latin
auto- Greek	*-cide* Latin (kill)
be- English	*-craft* English
bi- Greek	*-cy* Greek/Latin
bio- Greek	*-dom* English
cardi- Greek (heart)	*-ectomy* Greek (cutting)
cent- Latin	*-ed* English
circum- Latin	*-ee* French
col- com- con- Latin	*-en* English
contra- Latin	*-ese* French
counter- French	*-esque* French
cross- English	*-ess* French
crypto- Greek (secret)	*-est* English
deca- Greek (ten)	*-fold* English
dys- Greek (abnormal)	*-ful* English
eco- Greek (home)	*-gon* Greek (angle)
em- en French	*-gram* Greek (written)
ex- Latin	*-graph* Greek (pictured)
for- English	*-hand* English
fore- English	*-hood* English
geo- Greek (earth)	*-ian* Latin
haem- Greek (blood)	*-iana* Latin (associated with a person)
half- English	*-ible* Latin/French
hand- English	*-ic* Latin/Greek
hetero- Greek (different)	*-ics* Greek
homo- Greek (same)	*-ing* English
home- English	*-ion* Latin
hydr- Greek (water)	*-ish* English
hyper- Greek (over)	*-ism* Greek (doctrine)
infra- Latin (below)	*-ist* Greek (follower)
inter- Latin	*-itis* Greek (illness)
kilo- Greek (thousand)	*-ity* Latin
macro- Greek (great)	*-ive* Latin
mal- Latin (bad)	*-kind* English
man- Latin (hand)	*-less* Englsih

Table 1.1   (Continued)

Prefixes/beginnings	Suffixes/endings
matri- Latin (mother)	*-let* French
mega- Greek (great)	*-like* English
micro- Greek (small)	*-logue* Greek (speech)
mid- English	*-ly* English
mille- Latin	*-made* English
mini- Latin	*-man* English
mis- English	*-mania* Latin (madness)
mono- Latin	*-ment* Latin/French
multi- Latin	*-metre* Latin
neo- Latin (new)	*-naut* Latin (sailor or ship)
neuro- Latin	*-ness* English
non- Latin	*-ocracy, -ocrat* Greek (rule)
octo- Latin	*-ology* Greek (field of study)
omni- Latin	*-osis* Greek (process or state)
out- English	*-phobia* Greek (morbid fear)
over- English	*-phone* Greek (sound)
pan- Latin (all)	*-ship* English
para- Greek (beyond)	*-side* English
photo- Greek (light)	*-some* English
physio- Greek	*-type* Greek
post- Latin	*-ular* Latin
pre- Latin	*-ure* French
pro- Latin	*-ward* English
proto- Greek (first)	*-ware* English
pseudo- Greek	*-wide* English
psycho- Greek (soul)	*-wise* English
quasi- Latin	*-woman* English
re- Latin	*-work* English
retro- Latin (back)	*-worthy* English
self- English	*-wright* English
semi- Latin	*-y* English
socio- Latin	
sub- Latin	
super- Latin (above)	
sym-, syn- Greek	
techn- Greek	
tele- Greek (distance)	
theo- Greek (god)	
thermo- Greek (heat)	
trans- Latin	
ultra- Greek (beyond)	
un- English	

physio, *psycho*, *ex* and *mega*. But although the classical variety is very great, the native forms are highly productive, albeit fewer in number. Thus the *un*-prefix made up about 4 per cent of the words in Anglo-Saxon and can still be used with great flexibility, as was shown in the section on antonyms. Similarly, consider the semantic range which can be derived from suffixes added to a basic word like *wood*, namely *woody*, *wooded*, *wooden* and *woodsy*.

Word-formation operates not only by using prefixes, suffixes and roots as building blocks on the model of *antidisestablishmentarianism*, a slightly absurd nineteenth-century formation. As we have already seen, with the reduction of grammatical inflections, new verbs could be derived from existing nouns by conversion. New verbal forms can also be derived from relatives. Thus the verb *edit* was formed about a hundred years after *editor*, a seventeenth-century agent noun, while the verb *peddle* occurs about 200 years after *pedlar*. This process is called **back formation** or **back derivation** since, as Simeon Potter puts it, 'New words are created by analogy from existing words that are assumed to be derivatives' (1975, p. 83). In a similar fashion *greed* and *difficult* are back formations from *greedy* and *difficulty*. The process is steadily on the increase, with verbs like *enthuse*, *commute* and *electrocute* already over a hundred years old, and being joined by annual harvests, like *bulldoze*, *televise*, *escalate* and *babysit*.

Lexical Varieties

Moving to more formal lexical categories, we find a considerable range of varieties. Our treatment of them, which concludes this introductory chapter, will discuss those types which are found throughout the history of English. These are **borrowings** or **loan words**, **neologisms**, **archaisms**, **eponyms**, **toponyms** and **compounds**. In chapter 7 we shall return to the topic to discuss those lexical varieties which are more a feature of Modern English, namely **portmanteaux**, **abbreviations**, **conflations**, **acronyms** and **diminutives**, as well as newer kinds of **compounding**.

Loan words

The many thousands of examples in the English vocabulary, from a great diversity of sources, form one of the major focuses of this study. (Latterly, with the growth of English as a global language, there has been the widespread converse development of 'Anglicisms', borrowings from English into other languages.) Both **borrowing** and **loan word** are misleading terms in that words

are not really borrowed, since they remain in the parent language and are not returned by the borrower.

A threefold distinction derived from German is applied by scholars to loan words on the basis of their degree of assimilation in the new host language. A **Gastwort** ('guest word') retains its original pronunciation, spelling and meaning. Examples are *passé* from French, *diva* from Italian and *leitmotiv* from German. A **Fremdwort** ('foreign word') has undergone partial assimilation, as have French *garage* and *hotel*. *Garage* has developed a secondary Anglicized pronunciation ('garrij') and can be used as a verb; *hotel*, originally pronounced with a silent 'h', as the older formulation *an hotel* shows, has for some time been pronounced like an English word, with the 'h' being sounded. Finally, a **Lehnwort** ('loan word') has become a virtual 'native' in the new language with no distinguishing characteristics. *Loan word* is thus an example of itself. Other examples abound in the earlier borrowings, such as *bishop* from Greek *episcopos* 'an overseer' and *cheese*, from Latin *caseus*, has developed the forms *cheesy*, *cheesed off* and so on. A more recent example is *waltz*, puritanically defined in 1825 as 'a riotous and indecent German dance', but now accommodated in various phrases, such as *waltz in, waltz off with* and *waltz up to*, which suggest impropriety but have nothing to do with the original dance.

A **calque** is a loan translation from another language. Examples are *superman*, the creation of George Bernard Shaw from the German original *Übermensch*, coined by Nietzsche in 1883, and English 'world-view', derived from German *Weltanschauung*. We may note the difference that whereas *superman* now seems to us a natural English formation, partly because of the character in popular fiction, 'world-view' still seems slightly unfamiliar. Calques (loan translations) were more common in the earlier stages of the language (see Crystal, 1996, p. 27). One of the most notable from the Anglo-Saxon period was *godspell*, literally 'good news', from Latin *evangelium*; although *evangelist* and *evangelical* have since become established, in recent decades *good news* has become part of religious parlance.

Neologisms and archaisms

As the language develops through time, so new words or **neologisms** make their appearance and old words or **archaisms** become obsolescent and then fall out of use. The Roman poet Horace used the charming image of a tree dropping its leaves and replacing them to illustrate the point in his *Poetics*. Historically this process is not continuous and constant. The Renaissance was a period of great lexical expansion. Of the hundreds of words which are first

recorded in the plays of Shakespeare, we may confine the list (for the time being) to ten: *assassination*, *obscene*, *critic*, *lonely*, *pedant*, *barefaced*, *puke*, *perusal*, *prodigious* and *mutiny*. By contrast, verbal innovation was less common in the eighteenth century.

There are degrees of neologism, differing extents of originality, which in turn affect the acceptability of a new term. For example, *weird*, now a common term virtually indispensable in American spoken English, was an ancient word which died out after the Anglo-Saxon period. It owes its modern currency to Shakespeare's resuscitation in the phrase 'the weird sisters' for the Witches in *Macbeth* (III. iv. 133). Other extremely original new words do not develop a life in the general lexicon. For example, in Shakespeare's *Macbeth*, consider the extraordinary use of *incarnadine* in Macbeth's terrified soliloquy (II. ii. 62–4):

> . . . this my hand will rather
> The multitudinous seas incarnadine
> Making the green one red.

Incarnadine is never encountered again, apart from some rare literary uses in the nineteenth century, unlike the ten words previously listed, which are now common. Another example would be *Amansstrength*, the unique creation of the Victorian priest-poet Gerard Manley Hopkins in his poem 'Harry Ploughman'. Such words are termed nonce words, meaning words found only once. (The term is derived from the archaic phrase *for the nonce*, meaning 'for the occasion', from Anglo-Saxon *for þæm anes*.)

Most neologisms are not literary, but witness words reflecting technical change. Examples from the past decade are: *walkman*, *teflon*, *daisy wheel* and *skateboard*. But they may derive from many sources. The word *blurb* was coined by the American journalist and wordsmith Frank Gelett Burgess in 1907, originally as a fictional name (Belinda Blurb) for a 'pulchritudinous young lady' on a book cover he had designed. The word obviously 'caught on' because it filled a semantic gap. *Serendipity*, meaning 'the capacity for making unplanned but fortunate discoveries', was coined by Horace Walpole in 1754. A rarer instance derives from the American mathematician Edward Kasner, who asked his son to coin a word for the number 1 followed by a hundred zeros: his son came up with the memorably symbolic form *googol*, which has become a standard term in the profession. In modern times neologisms have become very fashionable, especially in journalism: hence such recent forms as *airhead*, *brat-pack*, *toyboy*, *bonk*, *wannabe*, *lookalike* and *bimbo*. We shall return to the varieties of neologisms in chapters 7 and 8.

Archaisms are generally less in favour today, but one still encounters *sans* ('without'), *smitten* ('to be seriously in love') and *woo*, meaning 'to court', itself becoming obsolescent. In recent years *wondrous* has been resuscitated as something of a vogue word. Archaisms continue to exist in traditional phrases, like extinct insects in amber. For example, the formula in the death penalty runs 'You shall be hanged [not hung] by the neck until dead.' The old sense of *fond* was 'foolish', still preserved in the phrase 'fondly imagine'.

Eponyms and toponyms

Two lexical categories which tend to excite philologists particularly are words derived from the names of people, or **eponyms**, and those originating in place-names, or **toponyms**. As with other areas of the vocabulary, we should distinguish between common words like *guy* (from Guy Fawkes), *pander* (from Pandarus), *tantalize* (from Tantalus), *nicotine* (from Jean Nicot), *panic* (from Pan) and the more rarefied, such as *pompadour* 'a shade of purple and an elaborate, usually high, coiffure' (from Madame de Pompadour, the mistress of Louis XV of France) and *mithridatize*, from Mithridates, meaning 'to achieve immunity from poison by taking small doses'. A suprisingly large number of common words are eponyms in origin, as the five italicized terms in the following brief extract show:

> As it was bitterly cold and wet, I wore a *cardigan*, a *mackintosh* and *wellington boots*. An hour after trekking up the muddy path we reached the *tarmac* road and stopped to have a makeshift picnic of pork pies and *greengages*.

The *cardigan* is named after the seventh Earl of Cardigan, who led the disastrous charge of the Light Brigade in 1854 during the Crimean War, when the garment was first worn by British soldiers. Charles Mackintosh invented the waterproof material from which the *mackintosh* is made, while *wellington boots* derive from the high leather boots made fashionable by the great British general, the first Duke of Wellington. John McAdam developed the process of producing *tarmac* roads, while the *greengage* is partly named after Sir William Gage, who introduced the fruit into England from France in 1725.

Certain notable individuals have left their names in a variety of forms: thus the *caesarian* operation derives from Julius Caesar, who was delivered in this way, while his political pre-eminence is commemorated in the titles *Kaiser*, *Czar* and *Tsar*, as well as in 'Caesar's wife', meaning the consort of a great man, who should be above suspicion.

The depth of the classical tradition is apparent in many eponymous terms, such as *cereal* (from Ceres, the Roman goddess of fertility), *volcano* (from

Vulcan), *siren* (dangerously seductive mythological women), *meander* (from the river near Troy) and the more literary terms *draconian*, *procrustean* and *laconic*. The panoply of pagan gods lives on in the adjectives *jovial*, *martial*, *mercurial*, *saturnine* and *venereal* (from Jove/Jupiter, Mars, Mercury, Saturn and Venus respectively), while *Aphrodite* survives in *aphrodisiac*.

Toponyms similarly range from everyday names like *jeans* (from Genoa) *alsatian* and *spaniel* (dogs originating in Alsace and Spain respectively) to curiosities like *balaclava*, *antimacassar* and *solecism*. The first is a woollen head-covering, named after a battle in the Crimean War in 1854; the second is a covering on the back of a tall chair, originally to prevent soiling by macassar oil, used to lubricate gentlemen's hair (though there is some dispute about the actual origin of the oil) and the third is a linguistic mistake, deriving from the inhabitants of Soloi in Greece. Toponyms are also surprisingly common, being found in the following six italicized words:

> Brigitte Bardot created *bedlam* in the fashion show when she removed her plain *denims* and *paisley jersey* on stage, revealing a stylish *turqouise bikini* beneath.

Bedlam, meaning 'pandemonium' or 'confusion', comes from the name of a medieval lunatic asylum in London, the Hospital of St Mary of Bethlehem. *Denim* was originally called *serge de Nîmes*, after the city in France. The *paisley* pattern takes its name from a town in Scotland, although the design originated in India. *Jersey* originates from the Channel Island of that name, while *turquoise*, meaning 'Turkish', refers to the bluish-purple gemstone, associated with Turkey. The *bikini* derives from the name of an atoll in the Pacific where the USA exploded an atom bomb in 1946. However, the two-piece swimsuit has no direct link with the place, being a clever opportunistic fashion-launch term trading on the associations of sun, heat and powerful impact.

As these examples partly show, eponyms and toponyms frequently relate to certain categories: items of clothing, cloths or materials, inventions, wines and spirits are prominent. Thus *muslin* comes from Mosul, *worsted* from *Worthstead* in Norfolk, *gauze* from Gaza, *calico* from Calicut, while *aran*, *duffle* and *jodhpurs* similarly derive from place-names. Likewise, the hats variously known as the *trilby*, the *bowler* and the *stetson* are named after those who made them fashionable. So are *bloomers*, the *teddy bear* and the *Windsor knot*. Sometimes the names become confused in the process: *sideburns*, for example, were made fashionable by Ambrose Everett Burnside, a Union general in the American Civil War. *Tawdry*, meaning an item of clothing or jewellery which is bright, arresting but cheap, comes from a fair named after St Awdrey. The enquiring student will make many surprising discoveries in words which do not seem per-

sonal or exotic: thus the common *tabby* cat derives its name from silk with a waved pattern, with the pedigree of fifteenth-century French *tabis*, from Spanish *tabi*, from Arabic *utabi*, the name of the quarter in Baghdad where the silk was made.

Most wines are named after their grape cultivars, such as *chardonnay* and *riesling*, or their area of original production: hence *champagne*, *burgundy*, *bordeaux*, *alsace* and *sherry*, from Jerez in Spain. Spirits, on the other hand, have more interestingly complex origins. Thus *whisky*, previously spelt *whiskybae* and *usquebaugh*, derives from Gaelic *Uisgebeath*, meaning 'water of life'. This emphasis on the supposed medicinal benefits of the alcoholic beverage is, of course, a standard euphemism paralleled by *aqua vitae* and its continental variations, *akvavit* and *eau de vie*. In a similar fashion Russian *vodka* means 'little water'. *Gin*, the demon of eighteenth-century England, is an abbreviated form of *geneva*, via Dutch *genever* and Old French *genevre*, ultimately from Latin *juniperus*, 'juniper', the main flavouring agent. *Brandy* also has a Dutch origin in *brandewijn*, meaning 'burnt wine', since it is distilled in oak barrels which have been lightly burnt inside to impart the distinctive flavour of the spirit. *Rum* is a truly curious term, and a more spectacular abbreviation, being known in the eighteenth century as *rumbo*, short for *rumbowling*, properly *Rumbullion* in Barbados, named from a Devonshire word meaning 'uproar'.

Understandably, many eponyms commemorate inventors. Amongst them are: *ampère*, *biro*, *braille*, *bunsen*, *Celsius*, *derrick*, *diesel*, *Fahrenheit*, *galvanize*, *guillotine*, *joule*, *morse*, *ohm*, *pasteurize*, *pavlovian*, *richter*, *salmonella*, *saxophone*, *silhouette*, *volt*, *watt*, *zeppelin*. A sub-category here is the naming of guns, including *colt*, *gatling*, *kalashnikov*, *mauser*, *maxim gun*, *tommy gun* and *winchester*.

The diversity of human types, of what Alexander Pope called 'the glory, jest and riddle of the world' is commemorated in the names of those famous for extremes of behaviour, notably: *casanova*, *chauvinist*, *dunce*, *martinet*, *masochist*, *maverick* and *sadist*. Particular individuals remembered as criminals, victims or eccentrics are: Fanny Adams, Heath Robinson, Hobson's choice, Ned Kelly, Namby Pamby, Peeping Tom, Robin Hood, Uncle Tom, Smart Aleck, Boycott and Hooligan.

A whole vocabulary of literary eponyms has filtered into daily usage. Amongst them are: *bowdlerize*, *don juan*, *Falstaffian*, *gargantuan*, *kafkaesque*, *lilliputian*, *odyssey*, *pander*, *romeo*, *quixotic*, *rabelaisian*, *jekyll* and *hyde*, *malapropism*, *stentorian*, *yahoo*, *shylock* and *scrooge*. Some are now unrecognizable offspring of their distant parents: *zany*, meaning a buffoon or something bizarrely comical, comes from *Giovanni*, a comic character in the Italian *commedia dell'arte*.

Eponyms and toponyms provide a stock of roots which are often interesting in their own right. Most people respond with simple philological pleasure to the information they provide about human behaviour and enterprise. But students must be cautious and sceptical not to be seduced into accepting folk etymologies. Examples abound, but let us consider the case of *crap* in a standard work of eponyms:

> To *crap* is to defecate and derives from Crapper's Valveless Water Waste Preventor which was the name under which the first flush lavatory was sold in England. The inventor, Thomas Crapper, who was born in Thorne, near Doncaster, in 1837, delivered England from the miserable inconvenience of the chamber pot and the garderobe. (Boycott, 1982, p. 35)

Unfortunately, the connection between the name and the object is not supported by any major reference work. Although it is difficult to distinguish between the various senses of 'rubbish' and 'waste' which accumulate around *crap*, they seem to have solidified into the main low sense of 'excrement' by the late eighteenth century, at least fifty years before Crapper was born.

Folk etymology is an interesting phenomenon, since it shows, not so much that people are gullible, but that they expect words to have intelligible origins, as many eponyms and toponyms in fact do. This particular lexical area has attracted a number of studies, but students and philologists should be warned against taking James Cochrane's amusing study, *Stipple, Wink & Gusset* (1992) in any spirit other than that intended. *Wink* is very much the key term in the title, appearing last, and facetiously etymologized by Cochrane as follows:

> Wink, Friedrich von 1755–1811 German nobleman . . . As Bavarian ambassador to the court of George III he introduced into London society his native custom of closing one eye to indicate that a witty or ironic remark was being made.

Compounding

As **compound** forms such as *barefaced*, *weekend*, *superman* and *brat-pack*, already encountered in the discussion, show, English has considerable capacities for making up new words from its own resources. In fact compounding is one of the earliest and most extensive kinds of word-formation, and continues to be so. It is a general feature of the Germanic languages, and we shall find many examples in the sections of chapter 2 which deal with Anglo-Saxon.

There are literally hundreds of ways of making up compounds. Two nouns are commonly combined, as in *handbook* and *workshop*; some are very ancient,

such as Old Norse *wind-auga* ('eye of the wind') which has yielded modern *window*. Thousands, like *cupboard* and *saucepan*, *holiday* and *eyelid*, no longer immediately strike us as being compounds. Two adjectives may be used, as in *bitter-sweet* and *deaf-mute*. Adjectives and verbs combine in older formations, such as *blindfold*, while more recent formations, such as *see-through*, combine a verb and a preposition making a new compound adjective, alternative to *transparent*. Nouns and verbs make up *shoplift* and *babysit*, while adjectives and nouns combine in *eyesore* and *black market*. More rarely, two verbs are used, as in *hitch-hike* and *has-been*. More complex formations are *man in the street*, *dog in the manger*, *lackadaisical*, *dryasdust*, *ne'erdowell* and *good for nothing*. As we have seen in the previous section, great numbers of prefixes and suffixes, which may be native or classical, create further possibilities. The great efflorescence of compounding in Modern English will be discussed in chapter 7.

A curious mode of compounding consists of **reduplicating forms**, such as *shilly-shally*, *hugger-mugger*, *rugger bugger* and *walky-talky*. Although they seem like baby-talk, many of these formations have interesting formal origins and extensive histories. The first derives from Middle English *shill I shall I*, 'shall I do it or not?' before it took on the meaning of to vacillate around 1700. The second is much associated with the secrecy, skulduggery and intrigue of the Elizabethan stage; the third is a humorous modern description of the sport-obsessed philistine type, while the last is a piece of portable radio equipment, although the form is first recorded in Jamaica in 1774. Another enduring example is *namby-pamby*, applied to someone insipid or feeble, originally the eponymous nickname given to a minor poet of the eighteenth century, Ambrose Philips. *Hocus-pocus* is a corruption of the sacred phrase *hoc est corpus*, used in the Eucharist; it became a juggling phrase before the first word degenerated further into *hoax*. The reduplicating type thrives creatively, with hundreds of forms deriving from a great diversity of sources, such as *lovey-dovey*, *mumbo-jumbo* (from West Africa) and *Oedipus-schmoedipus* from Yiddish.

Conclusion: Words and Power

Looking back over the evolution of the English vocabulary which we have traced in this introductory outline, we can isolate and analyse the main factors affecting stability or change in lexical development. In the period up to Early Modern English the main factor was demographic change, with waves of invaders producing what Daniel Defoe amusingly described as 'Your Roman-Saxon-Danish-Norman-English' in his poem, *The True-Born Englishman* (1701,

l. 139). The main social, political and economic changes of this earlier period are reflected symbiotically in lexical changes.

As we have partly seen, these lexical elements exist in different proportions and are found in different registers. We shall be tracing their incorporation in more detail in the course of the following chapters, showing that words are the signs and markers of cultural dominance in the process of social change. We shall focus in detail on the foundations of the word-stock in Anglo-Saxon and trace the survivals from Celtic, the subsequent additions from Norse, Norman French, Latin, Greek and various foreign languages as reflections of different social structures and dynamics within the English speech-community. These social dynamics may be categorized as annihilation, cohabitation, dominance, prestige and colonialism.

Thus the sparse remnants of Celtic exemplify the annihilation mode: apart from place-names only a few words such *ass*, *bin*, *dad* and *brock* have survived. The grass-roots infiltration of Norse exemplifies the cohabitation mode: although the Vikings were territorially constrained to the Danelaw in the north-east, they interacted with the Anglo-Saxons on the same social level, so that a surprising number of common words deriving from Old Norse have penetrated the core vocabulary. These include the central grammatical forms *they*, *their*, *them* and *are*, as well as such everyday words as *law*, *egg*, *sky*, *leg*, *ugly*, *rotten* and *husband*. The dominance mode is seen in the exercise of a new language of power in Norman French. The Norman influx in fact affected every lexical field, from basic words like *face* and *vegetables* to cultural words like *art*, *paint*, *picture* and *music*, but revealingly our vocabulary of power is still largely derived from them, as is attested to by words like *power*, *reign*, *realm*, *court*, *state* and *govern*. Latin and Greek terms represent the prestige mode, and (looking further ahead) the slow absorption of foreign and exotic words signals the mode of imperialism. Finally, the growth of a large and powerful speech-community in the 'New World' can be plotted in the shift in the balance of linguistic power between British and American varieties of English.

Other social changes had lexical consequences. With the coming of printing, semantic fields became more responsive to literary and academic interventions and influences through the various forms of the print media, as new words were brought into the language by individuals rather than peoples. Furthermore, technological changes brought about by noted innovators have created many new terms or witness words. Some of these are semantic recyclings of older words, such as *satellite* (which originally meant 'a servant') and *plastic* (originally an adjective meaning 'mouldable'); others are entirely new words like *wireless*, *radio*, *transistor* and *railway*.

Perhaps the most significant change between Early Modern English and Contemporary English from a lexical point of view has been one of attitude. This is shown in two ways. First, there has been a far greater acceptance of innovation. Secondly, there has been a diminution of the earlier chauvinism and suspicion towards alien terms, regarding them as interlopers, and an increasing acceptance of the foreign and the exotic. This change, part of the recognition that English has attained the status of a global language, in which norms are no longer determined by standards within the British Isles, has brought about the tremendous lexical growth of recent decades, which we shall discuss further in chapters 7 and 8.

Notes

1 The richness of the field of slang terms for drunkenness can be gauged just from the words beginning with the letter 's', namely *sloshed*, *slammed*, *slewed*, *smashed*, *stewed* and *sozzled*. Similar terms for money are more recent, historically speaking, but include *loot*, *bread*, *moosh*, *dough*, *duff*, *lolly* and *spondulix*.

2 *Lexeme* appears to have been coined by that remarkable researcher, Benjamn Lee Whorf. He distinguished between 'lexemes (stems) and other morphemes (formatives)' (1940, p. 160).

3 St Petersburg provides the most extreme case of a place-name being changed for ideological reasons. It was renamed Petrograd (1914–24), then Leningrad from 1924 to 1995, when it resumed its original name.

4 Berlin and Kay (1969) have shown the sequence of colour terms in lexical history. The basic sequence is black and white $\Rightarrow$ red $\Rightarrow$ yellow and green $\Rightarrow$ blue $\Rightarrow$ brown, grey, orange, purple and pink.

5 Winchester College, the oldest of the English public schools (founded 1382), has a very extensive dialect called 'Notions' which every new boy is expected to learn.

6 *Salt-cellar* is thus etymologically a tautology: so is *greyhound*, in which the first element is derived from ON *grey*, 'a dog', not from OE *græg*, 'grey'.

7 Students will enjoy Kingsley Amis's retailing of the hilariously implausible explanation (in Liddell and Scott's *Greek-English Lexicon*) of the etymology of *sycophant* as 'fig-shewer' from GK *sukon* 'a fig' and *phaino* 'to show' (1997, p. 221).

8 Other notable semantic changes are *garble* from 'separate' to 'confuse', *anon* and *presently* from their original sense of 'immediately' to 'in a while', *aristocracy* from its literal sense of 'rule by the best' to its modern 'unpopular' meaning, and *purchase* from 'to acquire by force' to the modern monetary sense. Many other examples are covered in my study, *Words in Time*.

9 Studies of dialect terms, such as those by Orton and Wright (1974), Upton and Widdowson (1996) and Trudgill (1990), give further examples of regional survivals of ancient terms.

10 The curious history of flyting is discussed in more detail in my study *Swearing* (1991).

11 Ross's seminal article, which started with the premise that the aristocracy was distinguished solely by its language, was originally entitled 'Linguistic class-indicators in present-day English' in *Neuphilologische Mitteilungen*, 1954.

12 The term *malapropism* dates only from 1849, but Grose's *Dictionary of the Vulgar Tongue* (1785) includes *slip-slopping* for 'misnaming and misapplying any hard word' (from the character of Mrs Slipslop, in Fielding's *Joseph* And*rews* (1742)).

13 Although hybrids are increasingly the order of the day, there are limits. While Latin and English elements can be mixed quite freely (as in *earth science*), Greek requires more harmony. Hence *bibliophile* not 'bookphile', *philanthropy*, not 'philperson'. Similarly, words ending with *-archy* or *-ocracy* require first elements from Greek, making up combinations such as *oligarchy* and *ochlocracy*.