Words, Meaning and Vocabulary

An introduction to modern English lexicology

Second edition

STUDOVNA
ústřední knihovny PdF MU Brno

Howard Jackson
Etienne Zé Amvela

continuum

2 Where do English words come from?

Chapter overview

2.1 The origin of English 25
How English came to England
The name of the language
2.2 Historical development of English vocabulary 27
The Old English period (450–1066)
The Middle English Period (1066–1500)
Early Modern English (1500–1800)
The Modern English period (1800–present)
2.3 Native English vocabulary 37
Anglo-Saxon words
The influence of Celtic on English
2.4 The process of borrowing 38
Latin words in English
Scandinavian loanwords in English
Greek loanwords
French loanwords
German and Dutch loans
Romance loans other than from French
Loans from the East
Loanwords from other sources
Sources of most recent loans
2.5 Creating new English words 51
Root creation
Echoic words
Ejaculations
Word formation
2.6 Characteristics of modern English vocabulary 52
The size of the vocabulary
Frequency of occurrence and use of Anglo-Saxon words
'English' vocabulary
2.7 Summary 55

This chapter covers:

- English in the context of world languages
- historical development of English vocabulary from the Old to the Modern English periods
- native English vocabulary
- foreign elements in the English lexis
- creation of new words
- characteristics of Modern English vocabulary.

2.1 The origin of English

The 5000 or so languages of the world can be grouped into about 300 language families, on the basis of similarities in their basic word stock and grammars. One of these families, the Indo-European, includes most of the languages of Europe, the Near East, and North India. One branch of the Indo-European family is called Italic, from which Latin, and later the Romance languages, developed. Another is called Germanic, which had three branches: North Germanic, East Germanic, and West Germanic. It is not certain, however, whether North, East and West Germanic corresponded to actual languages.

The North Germanic branch is the linguistic ancestor of modern Scandinavian languages, viz. Danish, Swedish, Norwegian, and Icelandic. The East Germanic branch developed into Gothic, but it eventually died out. The West Germanic branch in turn developed into modern German, Dutch, Frisian, and English. Frisian, spoken in the north-west Netherlands and the island nearby known as Friesland, is considered the closest relative of English. Both languages derive from Anglo-Frisian, just as High and Low German derive from German.

2.1.1 How English came to England

The first people known to have inhabited the land that was later to become England were Celts. They are presumed to have come to the island around the middle of the fifth millennium BC. Their languages were yet another branch of the Indo-European language family.

The Celtic warriors withstood the initial attempt to add their land to the Roman Empire in 55–54 BC, but a century later they were overwhelmed by a much larger Roman army. Most of the island of Britain was occupied by the Roman legions, government officials, and their households, from about AD 43 until 410. When the Empire began to crumble, the military and government officials withdrew, leaving behind many settlements built around the installations of military

government. These bore names such as *Doncaster*, *Gloucester*, *Lancaster*, and *Worcester* – all derived in part from the Latin word *castra* 'camp'.

The withdrawal of the Roman troops virtually invited the invasion of the rich lowlands by the Picts and Scots, two tribes in the north of Britain that the Romans had never conquered. The Celts appealed to bands of Germanic warriors from across the North Sea for aid in defending their land. Soon the Germanic tribes – called Angles, Saxons, Frisians, and Jutes – came in ever increasing numbers. The Celts found out too late that their new allies had become their conquerors. Although wars of resistance continued for the next 200 years, some of the Celts were pushed steadily to the fringes of the country, to Wales, Cornwall, Cumbria and the Scottish highlands. Others crossed the English Channel to join their relatives in French Brittany. Those who remained were subjected to the government of the newcomers and became assimilated to them by intermarriage. After a few generations, their identity was lost within Anglo-Saxon society.

2.1.2 The name of the language

The Germanic invaders referred to the native Celts as *Wealas* ('foreigners'), from which the name 'Welsh' is derived. The Celts called the invaders *Sassenachs* 'Saxons', regardless of their specific tribes, a practice which was followed by the early Latin writers. By the end of the sixth century, however, the term *Angli* 'Angles' was in use. For example, in 601, a king of Kent, Aethelbert, is called *rex Anglorum* (King of the Angles). During the seventh century, the usual Latin name for the country was *Angli* or *Anglia*. This became *Engle* in Old English, while the name of the language was referred to as *Englisc* (the *sc* spelling representing the sound 'sh'). The word *Englaland* (land of the Angles), which later gave *England*, did not appear until the beginning of the tenth century.

> *EXERCISE 2/1*
>
> Write a definition of the word 'English'. Then compare your version with that given in the Key to Exercises at the end of the book.

2.2 Historical development of English vocabulary

We shall discuss in turn the Old, Middle, Early Modern, and Modern English periods, with special emphasis on their respective characteristics at the lexical level. We shall also highlight the major contributions of each period to the development of English lexis as a whole.

2.2.1 The Old English period (450–1066)

The first Old English (OE) manuscripts were simply a few scattered inscriptions written around the fifth and sixth centuries in the runic alphabet brought in by the Anglo-Saxons. These scattered inscriptions give very little information on the language. The literary age began only after the arrival of the Christian missionaries from Rome in AD 597. The first OE manuscripts, dating from around 700, are glossaries of Latin words translated into Old English, and a few early inscriptions and poems. Unfortunately, very little material remains from this period. The most important literary work, which survived in a single copy, was the heroic poem *Beowulf*, written around 1000. But there were also a number of shorter poems concerned with Christian subjects or reflecting Germanic traditions and dealing with topics such as war, patriotism, travelling and celebration.

It is generally acknowledged that most OE texts were written in the period following the reign of King Alfred (849–899), who arranged for the many Latin works to be translated, including Bede's *Ecclesiastical History* (Bede 731). But the total corpus is still considered relatively small. Crystal (1995: 10) points out that 'the number of words in the corpus of Old English compiled at the University of Toronto, which contains all the texts, is only 3.5 million – the equivalent of about 30 medium-sized novels'.

Since we are dealing with written records, it is important to say a word on OE letters and spelling. The OE alphabet was very similar to the one still in use today, but the absence of capital letters was a distinctive feature. Furthermore, a few letters had different shapes, while some modern letters were absent, e.g. *j*, *v*, *f*, *q*, *x* and *z*. Numbers were written only in Roman symbols. Several of the letters were used in combination (digraphs) to represent single sound units, just like modern English *th* and *ea*, as in 'tru*th*' and 'm*ea*t'. There was a great deal of variation in spelling, to such an extent that even with a single scribe in a single place at a single time, there could be variation. The same word could be spelt differently on the same page.

In most general terms, if we allow for the unfamiliar spelling and the unexpected inflections, we still notice a marked difference

between the words used in prose and those used in poetic texts. While the majority of words in prose are very close to Modern English, words in poetic texts are different. For example, most of the prepositions and pronouns are identical in form (though not always in meaning): *for, from, in, he, him, his.* Modern English speakers can easily recognise *singan* as 'sing' or *stod* as 'stood': *onslepte* is quite close to 'asleep', and *geleornode* to 'learned'. Omitting the *ge-* prefix makes *-seted* more like 'seated', *-seah* like 'saw', and *-hyrde* like 'heard'. On the other hand, some of the words look very strange, because they have since disappeared from the language, e.g. *glimplice*, 'suitable', *swefn*, 'dream', *beboden*, 'entrusted', and some grammatical words such as *se*, 'the'.

The OE lexicon is characterized by its readiness to build up words from a number of parts, a feature that has stayed with English (see Chapter 4). As we might expect, some words may look familiar, but have a different meaning in Modern English; e.g. *wif* refers to any woman, married or not; *sona* meant 'immediately', rather than 'in a little while'; and *fæst* (fast) meant 'firm' or 'fixed', rather than 'rapidly'. These are 'false friends' when translating from Old into Modern English.

Old English is characterized by the frequent use of coinages known as 'kennings', a term from Old Norse poetic treatises referring to vivid figurative descriptions often involving compounds. Sometimes the interpretation is straightforward, sometimes it is obscure and a source of critical debate. Famous kennings include *hronrad*, 'whale-road' for the sea, *banhus*, 'bone-house' for a person's body. Often, phrases and compound words are used. *God*, for example, is described as *heofonrinces weard*, 'guardian of heaven's kingdom', and as *moncynnes weard*, 'guardian of mankind'.

Kennings are sometimes a problem to interpret because the frequency of synonyms in OE makes it difficult to distinguish shades of meaning. For example, there are some twenty terms for 'man' in the herioc poem *Beowulf*, including *rinc*, *guma*, *secg* and *boern*; and it is not always easy to see why one is used instead of another. Naturally, when these words are used in compounds, the complications increase. Kennings were often chosen to satisfy the need for alliteration in a line, or to help metrical structure. In their compound forms, they also allowed a considerable compression of meanings, although many of these words, and of the elements they contain, are not known outside of poetry.

This discussion of the OE period shows that there are many differences between the way vocabulary was used in OE and the way it is used today. We shall comment on five of the most important differences. First, the Anglo-Saxon preference for expressions that are

synonymous, or nearly so, far exceeds that found in Modern English, as does their ingenuity in the construction of compounds. Secondly, the absence of a wide-ranging vocabulary of loanwords also forces them to rely more on word-formation processes based on native elements. As a consequence, OE displays much larger 'families' of morphologically related words (1.3.3) than are typical of modern English. Thirdly, the latter period of OE was characterized by the introduction of a number of 'loan translations' (or 'calques'). Calques are lexical items that are translated part-by-part into another language, e.g. as *superman* was translated from German *Übermensch*. Calques are common in late OE, as can be seen from the following examples of loan translations from Latin:

Latin	*Old English*	*Modern English*
praepositio	foresetnys	preposition
conjunctio	gedeodnys	joining
unicornis	anhorn	unicorn
aspergere	onstregdan	sprinkle

Fourthly, grammatical relationships in OE were expressed mainly by the use of inflectional endings. In Middle English, they came to be expressed, as they are today, mainly by word order. This fundamental change in the structure of the language took place in the eleventh and twelfth centuries. The most plausible explanation for the disappearance of OE inflections is that it became increasingly difficult to hear them, because of the way that words had come to be stressed during the evolution of Germanic languages. By placing the main stress at the beginning of a word, it became difficult to hear the different endings, especially when they were phonetically similar, e.g. *en, on* and *an* as in *faren, faron, faran*.

Finally, the OE corpus is believed to contain about 24,000 different lexical items. However, this lexicon is fundamentally different from the one that we find in Modern English. About 85 per cent of OE words are no longer in use. Furthermore, only about 3 per cent of words in OE are loanwords, compared with over 70 per cent in Modern English. OE vocabulary was predominantly Germanic, which is no longer the case for Modern English.

Before we close this section, we must mention the Viking invasions of the eighth and ninth centuries. These invasions had a significant impact on the development of the English vocabulary (see 2.4.2. below).

> *EXERCISE 2/2*
>
> Examine the following lines from the Old English poem *Beowulf*. How many words can you recognize as still part of modern English vocabulary?
>
> Panon untydras ealle onwocon,
> Eotenas ond ylfe ond orcneas,
> Swylce gigantas, þa wið Gode wunnon
> Lange þrage; he him ðæs lean forgeald.

2.2.2 The Middle English period (1066–1500)

The Middle English period has a much richer documentation than is found in Old English. As Crystal (1995: 34) explains, this is partly due to the fact that the newly centralized monarchy commissioned national and local surveys which resulted in a marked increase in the number of public and private documents. However, the early material is of limited value, because it is largely written in Latin or French, and the only English data that can be extracted relates to places and personal names. Material in English began to appear in the thirteenth century, and during the fourteenth century there was a marked increase in the number of translated writings from Latin and French, and of texts for teaching these languages. This output increased considerably up to the 1430s.

Middle English poetry was influenced by French literary tradition, both in content and style. Much of the earlier Middle English literature is of unknown authorship, but by the end of the period the situation had changed. Among the prominent names that emerged in the latter part of the fourteenth century are John Gower, William Langland, John Wycliff, Geoffrey Chaucer, and later on poets who are collectively known as the 'Scottish Chaucerians'. It is this body of 'literature', in the modern sense of the word, that provides the final part of the bridge between Middle English and early Modern English.

The diversity in spelling was far greater than that found even in Old English. Even in an edited text, we still find variant spellings, e.g. *naure*, *noeure*, *ner*, *neure*, all standing for *neuer*, 'never'. This situation may be accounted for by a combination of historical, linguistic and social factors. Because of the spelling, several words look stranger than they really are. For example, *cyrceiaerd* would be close to the modern pronunciation of *churchyard*, if we understand that the two *c* spellings represent a 'ch' sound, and that *i* stood for the same sound as modern *y*. Similarly, *altegaedere* is not far from *altogether*, and *laeiden* from *laid*. As the period progressed, so the spellings changed to approximate those of Modern English.

Middle English is particularly characterized by intensive and extensive borrowing from other languages. In particular the Norman conquest of 1066, which introduced French-English bilingualism into England, paved the way for a massive borrowing of French words into the English vocabulary (see 2.4.4 below). The effect of the borrowings on the balance of the vocabulary was unprecedented. In early Middle English, over 90 per cent of the lexicon was of native English (Anglo-Saxon) origin. By the end of the Middle English period, this proportion had fallen to around 75 per cent. However, loanwords were by no means the only way in which the vocabulary of Middle English increased. The processes of word formation, such as compounding and affixation, which were already established in Old English, continued to be used, and were extended in various ways.

> *EXERCISE 2/3*
>
> Make a modern English 'translation' of the following lines from William Langland's *Piers Plowman* (late fourteenth century):
>
> Nau awaked Wrathe, wiþ two white eyes
> And wiþ a nevelyng nose, and nypped his lippys.
> 'I am Wraþ,' quod þat weye, 'wolde gladely smyte,
> Boþ wiþ stone and wiþ staf, and stel apon my enemy
> To sle hym sleyliest slezthes Y þynke.'

2.2.3 Early Modern English (1500–1800)

The transition from Middle English to Modern English would be too abrupt without the recognition of an Early Modern English period. However, there is no consensus about the beginning of this transitional period. Some consider an early date, around 1400 or 1450, others a later date, around 1500, to mark its beginning. But many consider the advent of the printing revolution as a determining factor. The year 1476, when William Caxton set up his press in Westminster, may therefore be a safe beginning point for the period. Printing played a major role in fostering the norms of spelling and pronunciation, in providing more opportunities for people to write, and in giving published works much wider circulation. As a result, more books were published, providing reliable evidence on the development of the

language. Furthermore, in the sixteenth century, scholars began seriously to talk about their language, making observations on grammar, vocabulary, the writing system and style.

The period encompasses the 'Renaissance', which runs from the middle of the fifteenth century until around 1650. It is characterized by a renewed interest in the classical languages and literatures, and by major developments in the sciences and arts. It includes the Protestant Reformation, important scientific discoveries, and the exploration of Africa, Asia and the Americas. All these factors had a major impact on the English language, especially on its vocabulary. Writers began to borrow from other European languages to express the new concepts, techniques and inventions that first came from Europe. But as exploration developed worldwide, words came into English from North America, Africa and Asia. Some came directly, while others came indirectly via other European languages. Furthermore, thousands of Latin and Greek words were introduced, as translators of texts from these languages could not find precise equivalents in English, especially in fields such as medicine and theology.

The influx of foreign vocabulary attracted bitter criticism from purists, who opposed the new terms on the grounds that they were obscure and were interfering with the development of native English vocabulary. Some attempted to revive little known words from English dialects; others, such as Thomas Elyot, went out of their way to find new words in order to 'enrich' and 'credit' the language. Their objective was to bring the new learning within the reach of the English public, so as to make the English language fit not only for the street but also for the library. However, their efforts were not enough to stem the influx of new words. In fact, it is generally acknowledged that the increase in foreign borrowings is the most distinctive feature of the Renaissance for English (see Crystal 1995: 60).

It is generally agreed that the two most important influences on the development of the English language during the last decades of the Renaissance are the works of William Shakespeare (1564–1616) and the *King James Bible* of 1611. Shakespeare's work provides numerous instances of the way in which the language was developing at the time. It also gives abundant and reliable information on areas such as pronunciation, word formation, syntax, and language use. However, the Shakespearean impact on English was primarily in the area of the lexicon, because his poems and plays introduced or popularized thousands of new words in the language.

The *King James Bible* was appointed to be read in churches throughout the kingdom. Its influence on the population and on the language as a whole was therefore far-reaching. As the translators point out in the preface, they aimed for a dignified, not a popular style, opting for older forms of the language even when modern alternatives were available. Consequently, their style was more conservative than that found in Shakespeare, and unlike Shakespeare, they did not feel the need to introduce large numbers of new words. However, there are many phrases in the *King James Bible* that have entered the language as idioms, though sometimes with minor changes in grammar or emphasis, e.g. can the leopard change his spots, an eye for an eye, fight the good fight, if the blind lead the blind, a wolf in sheep's clothing, in the twinkling of an eye, money is the root of all evil, new wine in old bottles, the salt of the earth, the skin of my teeth, the straight and narrow, a thorn in the flesh. The frequency of occurrence of such phrases in both literary and everyday language is a clear indication of the impact that the *King James Bible* continues to have on contemporary English.

The period between 1530 and the Restoration in 1660 (of the monarchy, after the period of the Commonwealth) witnessed the fastest lexical growth in the history of the language. In addition to massive borrowing and the application of different types of word formation using native sources, there were also a great many semantic changes, as old words acquired new meanings. By the end of the seventeenth century, many critics felt that English was changing too rapidly and randomly. They considered the resulting language 'unruly', 'corrupt', 'unrefined' and 'barbarous'. It was therefore felt to be necessary to 'stabilize' the language, so that their works would be intelligible to future generations. However, neither Britain nor the United States chose the Academy solution for the stabilization of the language, as had the French and the Italians. The only part of the English-speaking world that has ever set up an Academy is South Africa. The debate on language corruption in the seventeenth century had the merit of drawing public attention to the need for 'preserving the consistency and stability' of the English language. Since the solution was not by means of an Academy, efforts were made to produce grammars, spelling guides, pronunciation manuals and dictionaries, from which it was expected that standards of correctness would emerge.

The importance of dictionaries had been felt long before the debate on language corruption. For example, Richard Mulcaster wrote the following in 1582: 'It were a thing verie praiseworthie if someone well learned and as laborious a man, wold gather all the words which be vse in our English tung ... into one dictionarie' (quoted in Crystal 1995: 73). In 1604, Robert Cawdrey published the first 'dictionary of hard words', which had about 3000 entries of 'hard vsuall English

wordes', mostly borrowings, such as *abbettors*, glossed as 'counsellors', and *abbruiat*, glossed as 'to shorten, or make short'. It was in fact the first synonym dictionary, and it was followed by other compilations on similar lines. In 1721, Nathaniel Bailey published his *Universal Etymological English Dictionary*, which was a marked improvement on the previous 'hard word' dictionaries. Bailey gives more and fuller entries, but his definitions still lack illustrative support, and he gives little guidance on usage.

The English lexicon received its first really authoritative treatment in 1755, when Samuel Johnson published his *Dictionary of the English Language*. Although he has fewer entries than Bailey, his selection is considered more wide-ranging and his lexicological treatment more discriminating and sophisticated (Crystal 1995: 74). Another important feature of Johnson's *Dictionary* is that the alphabetical section is preceded by a preface in which he outlines his aims and procedures. The preliminaries also include a short history of the language, and a grammar, with sections on orthography and prosody. Johnson made a major departure from the prevailing prescriptive, to a descriptive approach in lexicography. In his preface, he stressed that his aim was to 'not form, but register the language'; and it is this principle that introduced a new era in lexicography. This work dominated the dictionary market for decades and appeared in several editions for much of the next century, when it began to receive a great deal of criticism. However, as Crystal (1995: 75) puts it: 'the fact remains that Johnson's *Dictionary* was the first attempt at a truly principled lexicography'. Among other things it was the first accurate description of the complexity of the lexicon and of word usage.

EXERCISE 2/4

Examine the following extract from George Fox's *Journal* (mid-seventeenth century). How does it differ in vocabulary from modern English?

> & before I was brought in before him ye garde saide It was well if ye Justice was not drunke before wee came to him for hee used to bee drunke very early: & when I was brought before him because I did not putt off my hatt & saide thou to him hee askt ye man whether I was not Mased or fonde: & I saide noe: Itt was my principle . . .

2.2.4 The Modern English period (1800–present)

The Modern English period runs from the beginning of the nineteenth century to the present day. As in the case of the preceding periods, the change from Early Modern to Modern was equally gradual and imperceptible. For example, while it may be observed that many of the words used today had a different meaning in the eighteenth century, it would be very difficult to say exactly when the change of meaning occurred.

From the point of view of lexis, Modern English may be characterized by three main features: the unprecedented growth of scientific vocabulary, the assertion of American English as a dominant variety of the language, and the emergence of other varieties known as 'New Englishes' (see 6.2).

English scientific and technical vocabulary has been growing steadily since the Renaissance. But the nineteenth century saw an unprecedented growth in this domain, as a consequence of the industrial revolution and the subsequent period of scientific exploration and discovery. The steady increase in the level of education made the general public more and more curious about and interested in science and technology. Consequently, most discoveries, such as Faraday's, and theories, such as Darwin's, received widespread publicity. By the end of the nineteenth century, one could actually speak of 'scientific English' as a variety of the language (see 6.3). The characteristics of this variety are made explicit in grammar books and style guides of academic journals. Some sciences, such as chemistry, physics and biology, made spectacular lexical developments during this period.

The strong linguistic influence of American English as a dominant variety of the language can be explained, at least in part, by the emergence of the USA as the leading economic power of the twentieth century. This influence is felt directly in the areas of pronunciation and grammar, but more especially at the lexical level. Although there are still marked differences between American and British English (see 6.2.1), it must be acknowledged that the two varieties are becoming more and more alike, for several reasons: communication systems have improved greatly since the beginning of the twentieth century, the USA has become more and more involved in world affairs, the UK, just like the rest of Europe, is increasingly open to American culture, and finally the development of the mass media is playing an important role in this process. The assertion of American English is made even stronger by the Americans' numerical strength. In fact, the USA contains nearly four times as many speakers of English as a first language

as the UK, and according to Crystal (1995: 106) these two countries comprise 70 per cent of all speakers of English as a first language in the world. Such dominance, with its political and economic correlates, makes American English a force to reckon with in any discussion of how the English language is likely to develop.

The concept of 'New Englishes' refers to new varieties of the language that have become localized not only through the influence of the other languages of the regions where they are used, but also through being adapted to the life and culture of their speakers. They are considered varieties of English in their own right, just like the 'older' Englishes, such as British and American English. These new Englishes have their origin in the colonial era. Well-known examples include Indian English, Philippine English, Singapore English, and African Englishes of such nations as Cameroon, Ghana and Nigeria. Vocabulary is the area in which these new Englishes best assert themselves.

In addition to new Englishes that are associated with a given geographical area, we must also mention the varieties of the language based on subject matter (see 6.3). These have also known an accelerated development during the Modern English period. Some, such as telecommunications and computing, are relatively new; others, such as religious and legal English, find their origins in earlier periods.

As a closing remark, and following Kachru (1983), quoted in Fennell (2004: 255–56), English–speaking countries may be placed in three concentric circles depending on the status of the language:

(1) The **inner circle** whose countries have English as the primary language, e.g. Great Britain, Ireland, USA, Canada, Australia and New Zealand;

(2) The **outer** or **extended circle** includes countries affected by colonization, where English is used in the leading institutions and plays a second language role in a multilingual setting, e.g. India, Malawi, Nigeria, Singapore, and about fifty other territories;

(3) The **expanding circle** includes countries which have no history of colonization and where English has no special administrative status; but where it is recognised as an important international language, e.g. China, Japan, Poland, and a growing number of other states. In her discussion of the use of English throughout the world, Fennell (2004: 255) estimates that a total of 75 territories give a special place to the English language in their respective societies (see 6.2 for more on 'national and regional vocabularies' of English).

EXERCISE 2/5

Look up the history of the following words, preferably in the *(Shorter) Oxford English Dictionary*, otherwise in a dictionary with good coverage of etymology (e.g. *Collins English Dictionary, Oxford Dictionary of English*). How has their meaning changed since their first entry into the language?

cousin, hose, knight, nice, span, spell, starve, train

2.3 Native English vocabulary

Native English vocabulary is made up of Anglo-Saxon words; but we shall also discuss the influence of Celtic on English.

2.3.1 Anglo-Saxon words

This category consists of words that arrived with the Germanic invaders and are still used in modern English. Some are grammatical words, such as *be, in, that*, while others are lexical words: *father, love, name*. Most of them are the common words of the language. They constitute a nucleus or central mass of many thousand words with an undisputed claim to 'Anglicity' or Englishness. Some of them are only literary, some only colloquial, while the great majority are both literary and colloquial.

Anglo-Saxon words are generally short and concrete; e.g. parts of the body (arm, bone, chest, ear, eye, foot, hand, heart), the natural landscape (field, hedge, hill, land, meadow, wood), domestic life (door, floor, home, house), the calendar (day, month, moon, sun, year), animals (cow, dog, fish, goat, hen, sheep, swine), common adjectives (black, dark, good, long, white, wide), and common verbs (become, do, eat, fly, go, help, kiss, live, love, say, see, sell, send, think) (see Crystal 1995: 124).

2.3.2 The influence of Celtic on English

The conditions of cultural contact between the Celts and the invading Anglo-Saxons were such that the Celtic language could not have any serious impact on Old English. Only a handful of Celtic words were borrowed at the time, and just a few have survived into modern English, sometimes in regional dialect use: *crag, cumb*, 'deep valley', *binn*, 'bin', *carr*, 'rock', *torr*, 'peak' and *luh* (Scots *loch*, Irish *lough*), 'lake'. A

few Celtic words of this period derive ultimately from Latin, brought in by the Irish missionaries, e.g. *assen*, 'ass', *ancor*, 'hermit', and possibly *cross*. There are a few Celtic-based placenames in what is now southern and eastern England. They include river names such as *Avon*, 'river', *Don, Exe, Ouse, Severn, Thames, Trent, Usk* and *Wye*. Town names include: *Bray*, 'hill', *Dover*, 'water', *Eccles*, 'church', *Kent, Leeds, London* (a tribal name), *York*, and the use of *caer*, 'fortified place' (as in *Carlisle*) and of *pen*, 'head, top, hill' (as in *Pendle*).

In more recent times, a few more Celtic words were introduced into English from Irish Gaelic in the seventeenth century – *brogue, galore, shamrock, tory* – and later on: *banshee, blarney, colleen*. From Scots Gaelic come: *clan, loch, bog, slogan, whisky*. And from Welsh: *crag*, occurring first in Middle English. Probably there are no more than two dozen Celtic loanwords in all. Consequently, Celtic has had a rather negligible influence on English. English is, thus, basically a Germanic language.

EXERCISE 2/6

(a) Take two pages from your dictionary, one from the letter 'h' and one from the letter 't', and count how many words have, according to the etymology, their immediate origin in Anglo-Saxon (A-S) or Old English (OE). What is the proportion of OE words on these two pages?

(b) Now take the first 100 words of a news article in a newspaper, and, by looking up the etymology of each word in your dictionary, calculate the proportion of OE words in the extract. If a word occurs more than once, include each occurrence in your count.

How do the two proportions differ?

2.4 The process of borrowing

By definition, when speakers imitate a word from a foreign language and, at least partly, adapt it in sound or grammar to their native language, the process is called 'borrowing', and the word thus borrowed is called a 'loanword' or 'borrowing'. There is a sense in which neither term is really appropriate, on the grounds that the receiving languages never return the borrowed words (Crystal 1995: 126). English may be considered an insatiable borrower, in the sense that, while other languages take special measures to exclude foreign words from their

lexicons, English seems to have welcomed such words throughout its history, especially from the Middle English period onwards. It is estimated that over 120 languages from all over the world have been sources of present-day English vocabulary.

The history of a loanword may be quite complex, because such words may have come to English not directly, but via another language or two. According to Crystal (1995: 285), 'loanwords have, as it were, a life of their own that cuts across the boundaries between languages'. To give a stock example, *chess* was borrowed into English from Middle French in the fourteenth century. The French word (plural *esches*, singular *eschec*) came from Arabic, which had earlier borrowed it from Persian *shah*, 'king'. Thus the etymology may be traced from Persian, through Arabic and Middle French, to English. So, the direct or immediate source of *chess* is Middle French, but its ultimate source (as far back as we can trace its history) is Persian.

We shall now discuss in turn English borrowings from a number of languages, beginning with Latin.

2.4.1 Latin words in English

Latin has been a major influence on English right from the Germanic period up to modern times. The degree of this influence has varied from one period to another.

The Anglo-Saxons must have encountered Latin as used by the continental Roman armies; but only a few words have subsequently come into English as a result of these early contacts. In England, some early borrowings were concerned with the military domain, commerce and agriculture; others dealt with refinements of living, which the Germanic people acquired from their continental contacts with the Romans, e.g. words relating to clothing (*belt*, 'belt', *cemes*, 'shirt', *sutere*, 'shoemaker'), buildings and settlements (*tigle*, 'tile', *weal*, 'wall', *caester*, 'city', *straet*, 'road'), military and legal institutions, commerce and religion (*maesse*, 'mass', *munyc*, 'monk', *mynster*, 'monastery'). It is generally established that the total number of Latin words present in English at the very beginning of the Anglo-Saxon period is less than 200 items. Most of these items indicate special spheres in which the Romans excelled, or were believed to do so by the Germanic peoples. Many of these words have survived into modern English, including: *ancor*, 'anchor' (Lat. *ancora*), *butere*, 'butter' (Lat. *butyrum*), *caelc*, 'chalk' (Lat. *calco*), *cese*, 'cheese' (Lat. *coquina*), *disc*, 'dish' (Lat. *discus*), *mile*, 'mile' (Lat. *milia*), *piper*, 'pepper' (Lat. *piper*), *pund*, 'pound' (Lat. *pondo*), *sacc*, 'sack' (Lat. *saccus*), *sicol*, 'sickle' (Lat. *secula*), *weall*, 'wall' (Lat. *vallum*).

The labels 'popular' and 'learned' are often used to refer to the main categories of words borrowed during this early period (Pyles and Algeo 1993: 288). Popular loanwords were transmitted orally and constituted part of the vocabulary used in everyday, non-specialized communication, e.g. *plante*, 'plant', *win*, 'wine', *catte*, 'cat', *straete*, 'road'. Most speakers would consider these words to be native English words and are generally not aware of their foreign origin. As for learned words, they came into English either through the church or through more or less scholarly influences. Some learned words may in time have passed into general use, e.g. *clerk*, from Latin *clericus* or Old French *clerc*.

Borrowing from Latin during the Old English period increased over that during the earlier Germanic period, but the popular/learned distinction continues to apply. In fact up to 1000 AD, many words continued to be borrowed from spoken Latin, dealing with everyday practical matters. However, after 1000, owing to renewed interest in learning encouraged by King Alfred and the tenth century Benedictine monastic revival, borrowings came from classical written sources. Most of these terms were scholarly and technical.

Among the early English borrowings, some were acquired not directly from Latin, but from the British Celts, e.g. *candel*, 'candle', *cest*, 'chest' (Lat. *cista/cesta*), *crisp*, 'curly' (Lat. *crispus*), *maegester*, 'master' (Lat. *magister*), *mynster*, 'monastery' (Lat. *monasterium*), *port*, 'harbour' (Lat. *portus*), *sealm*, 'psalm' (Lat. *psalmus*, from Greek), *tilge*, 'tile' (Lat. *tegula*). Later borrowings are characterized by the fact that the English form is closer to the Latin word, e.g. *alter*, 'altar' (Lat. *altar*), *(a)postol*, 'apostle' (Lat. *apostolus*), *balsam*, 'balsam' (Lat. *balsamum*), *circul*, 'circle' (Lat. *circulus*), *demon* (Lat. *daemon*), *martir*, 'martyr' (Lat. *martyr*). Since Latin borrowed freely from Greek, some of the loanwords are ultimately of Greek origin, e.g. apostle, balsam, demon. During this period, a number of Old English words were given new 'Christian' meanings, under missionary influence: heaven, hell, God, gospel, Easter, holy, ghost, sin.

It is generally estimated that around 500 words in total were borrowed from Latin during the entire Old English period. This is a relatively small number compared with Latin borrowings in later times. Furthermore, many Latin loanwords, especially from the later OE period, were not widely used, and some fell out of use. Some were borrowed again at a later period, perhaps with a different meaning. For instance, modern English *sign* and *giant* are not survivors from Latin borrowings into Old English (*sign*, *gigant*), but more recent borrowings from French *signe* and *géant*.

In the Middle English period, French is undoubtedly the dominant influence on the growth of Middle English vocabulary, between the Norman conquest and 1500 (see 2.4.4). Many borrowings also occurred directly from Latin during that period, though it is often difficult to specify whether a word was borrowed from French or from Latin. For instance, on the basis of their form alone, the words *nature, register, relation* and *rubric* might be from either language.

Most of the Latin borrowings in this period were professional or technical terms, belonging to fields such as religion, law and literature, scholastic activities and sciences in general. Religious terms include: collect (short prayer), mediator, redeemer. Legal terms include: client, conviction, subpoena. Connected with scholastic activities are: library, scribe, simile. And scientific words include: dissolve, equal, essence, medicine, mercury, quadrant. Several hundred such words were borrowed before 1500. A more extensive list would include verbs (admit, commit, discuss, seclude) and adjectives (complete, imaginary, instant, legitimate, populous). Latin words were borrowed by some writers in a deliberate attempt to produce literary or elevated styles, but only a few such terms entered the language (e.g. mediation, oriental, prolixity), while most disappeared soon after borrowing (e.g. abusion, sempitern, tenebrous).

The simultaneous borrowing of French and Latin words led to a highly distinctive feature of modern English vocabulary: sets of three items (triplets), all expressing the same fundamental notion but differing slightly in meaning or style, e.g. kingly, royal, regal; rise, mount, ascend; ask, question, interrogate; fast, firm, secure; holy, sacred, consecrated. The Old English word (the first in each triplet) is the most colloquial, the French (the second) is more literary, and the Latin word (the last) more learned.

Borrowing continued into the Modern English period, both from Latin, and from Greek via Latin. The Early Modern period saw an avalanche of Latin words, including: abdomen, area, digress, editor, fictitious, folio, graduate, imitate, lapse, medium, notorious, orbit, peninsula, quota, resuscitate, sinecure, superintendent, urban, vindicate.

By contrast with the Early Modern period, when exisiting Latin words were borrowed, in the later Modern period English has borrowed words or morphemes from Latin, but fashioned them into new words that Latin never knew. Such neo-Latin or neo-classical words (see 4.5.3) are used abundantly not only in the international vocabulary of science and technology, but also in other areas of modern life. Examples of such coinages are: *aleatoric*, 'dependent on chance' (from *aleator*, 'gambler'), *circadian*, 'functioning or recurring in 24-hour cycles' (from *circa diem*, 'around the day'), *pax americana*, 'peace

enforced by American power' (modelled on *pax romana*), *vexillology*, 'study of flags' (from *vexillum*, 'flag').

This discussion of loanwords from Latin shows that Latin is not only the first major contributor of loanwords to English, but also one of the most important sources for the coinage of new English words.

2.4.2 Scandinavian loanwords in English

The second major influence on English lexis came as a result of the Viking raids on Britain, which began in AD 787 and continued at intervals for some 200 years. In fact, by the mid-ninth century, the Danes controlled most of eastern England, which became known as Danelaw. Further invasion in 991 resulted in the English king being forced into exile and the Danes seizing the throne. England was then under Danish rule for 25 years. The linguistic effect of this prolonged period of contact between the Anglo-Saxons and the Danish settlers was threefold. First, a large number of settlements with Danish names appeared in England. Secondly, there was a marked increase in personal names of Scandinavian origin. And finally, many general words entered the language and became part of common English vocabulary.

Scandinavian placenames are particularly common in Yorkshire and Lincolnshire. Some end in *-by*, the Scandinavian word for 'farm' or 'town', e.g. Derby, Grimsby, Rugby, Naseby. Others end in *-thorpe*, 'village' (Althorpe, Astonthorpe, Linthorpe), *-thwaite*, 'clearing' (Braithwaite, Applethwaite, Storthwaite), *-toft*, 'homestead' (Lowestoft, Eastoft, Sandtoft). Although the word *by*, 'dwelling' existed in Old English, the *-by* ending is of Scandinavian origin. This is justified by the fact that placenames with this ending are almost entirely confined to the area of Danelaw (Crystal 1995: 25).

For personal names, a close examination of early Middle English records reveals a strong Scandinavian influence in the north and east, especially in Yorkshire and north Lincolnshire, where over 60 per cent of personal names seem to have been affected. Many of these names end in *-son:* Davidson, Jackson, Henderson.

With two cultures in such close contact for so long, we would expect extensive borrowing. However, most of the Scandinavian words in Old English did not actually occur in written records until the Middle English period. Because of the close relationship between Old Norse and Old English, many Scandinavian words resemble their English cognates so closely that it would be impossible to tell whether a given word was Scandinavian or English. Sometimes, if the meanings of obviously related words differed, semantic contamination took place, as when Old English *dream*, 'joy' acquired the meaning of the related Scandinavian *draum*, 'vision in sleep'.

Some of the commonest words in modern English came into the language at this time. They were made to conform wholly or in part with the English sound and inflectional systems. For example, Pyles and Algeo (1993: 294) note that the very common verbs *get* and *give* come to us not from Old English *gitan* and *gifan*, but from cognate Scandinavian forms. Even the personal pronoun system was affected, with *they, them* and *their* replacing earlier forms of the third person plural. One of the most remarkable borrowings relates to the verb *be*: 'the replacement of *sidon* by *are* is almost certainly the result of Scandinavian influence, as is the spread of the third personal singular *-s* ending in the present tense in other verbs' (Crystal 1995: 25).

A good many words with *sc-/sk-* are of Scandinavian origin: *scathe, scorch, score, scowl, scrape, scrub, skill, skin, skirt, sky*. Scandinavian loans sometimes involve little more than the substitution of one word for another (such as *window* for *vindauga*), but some borrowings expressed new concepts (such as certain Scandinavian legal terms) or new things (for various kinds of Viking warship). A large number of duplicate words also arose from the contact. In such cases, both Old Norse (ON) and Old English (OE) provided ways of denoting the same object or situation. In some cases the ON word was retained, e.g. *egg* vs. OE *ey, sister* vs. OE *sweoster, silver* vs. OE *sealfor*. In others the OE word survived, as in *path* vs. ON *reike, sorrow* vs. ON *site, swell* vs. ON *bolnen*. In a number of cases both words survived, but developed a useful difference in meaning:

ON		OE	
	dike		ditch
	hale		whole
	raise		rise
	sick		ill
	skill		craft
	skirt		shirt

Sometimes, where two forms have survived, one is considered standard and the other dialectal. In the following examples, the first word from OE is standard, while the ON equivalent is dialectal: cast/werpan, yard/garth, church/kirk, leap/laup, no/nay, true/trigg. As a final observation, it should be noted that a number of Scandinavian words continued to be borrowed even during the Modern English period: *muggy, rug, scud, ski*. More recent borrowings include: *geyser, rune, saga, ombudsman*.

2.4.3 Greek loanwords

As a classical language, Greek provided English, as well as other European languages, with a considerable number of technical terms in all branches of human knowledge. This need was perhaps most strongly felt by the English humanists, who wanted their language to be capable of expressing the most refined thoughts, just like Latin and Greek. Some Greek words were borrowed via Latin and French, some were derived from Greek and Latin elements, while others were taken directly from Greek.

Even before the Norman conquest in 1066, a number of Greek words had entered English by way of Latin, in addition to some very early loans such as *church*, which may have come into Germanic directly from Greek. From the Middle English period onwards, Latin and French were the immediate sources of most loanwords ultimately from Greek. But it was only in the Early Modern English period, from around 1500, that Greek influence became most noticeable. Greek words that came into English via Latin include: allegory, anaesthesia, chaos, dilemma, drama, enthusiasm, history, metaphor, paradox, phenomenon, rhythm, theory, zone. Those coming via French include: centre, character, chronicle, democracy, ecstasy, harmony, machine, pause, tyrant. The following came directly from Greek, though some are combinations unknown in classical times: acronym, autocracy, idiosyncracy, pathos, telegram, xylophone. Incidentally, the terms 'lexis', 'lexeme', 'lexical', 'lexicographer', 'dictionary' and 'vocabulary' are all derived from Greek and Latin elements.

Many of the Greek loanwords were considered learned, and some still are; but others have passed into the stock of more or less everyday vocabulary. Although Greek had considerable prestige as a classical language, it did not have the same influence as Latin, which was the language of literature, science, and religion, as well as the medium of instruction in European universities well into the seventeenth century and in some cases beyond. However, Greek studies received an enormous boost when Greek scholars fled to Europe following the conquest of Constantinople in 1453 by the Turks.

2.4.4 French loanwords

Borrowing from French has occurred ever since the Middle Ages, although not always on the same scale. Before 1066, there were close contacts between the English and French cultures following the exile to Normandy of Edward the Confessor, the son of Aethelred II and Emma, daughter of the Duke of Normandy. In fact, Edward lived there

for 25 years, returning in 1041 with many French courtiers. When he acceded to the throne, several of the French nobles were given high positions. Furthermore, the monastic revival started in France, and many English monks would have studied there. The linguistic consequences of these contacts was the borrowing of some French words into Old English, e.g. *servian*, 'serve', *bacun*, 'bacon', *arblast*, 'weapon', *prisun*, 'prison', *castel*, 'castle', *cancelere*, 'chancellor'.

Following William, Duke of Normandy's accession to the English throne in 1066, Norman French became the language of government, the courts, and the new upper social classes. Within twenty years of the invasion, almost all the religious houses were under French-speaking superiors, and several new foundations were solely French. Such a context was favourable for the development of French-English bilingualism: many English people learned French to gain advantage from the aristocracy, while many Norman French learned English in their daily contacts with local communities.

As the period progressed, the influence of French vocabulary on English became increasingly noticeable. It is estimated that by the end of the thirteenth century, some 10,000 French words had come into English. These words were in part to do with law and administration, but they also included words from fields such as medicine, arts, fashion, and everyday life. According to Crystal (1995: 46) over 70 per cent were nouns, a great proportion of which were abstract terms constructed using French affixes, such as *con-*, *trans-*, *pre-*, *-ance*, *-tion*, *-ment*. Crystal concludes that three-quarters of these French loans are still in use in English today.

The following words, which have to do with lay and spiritual administration, are of French origin: government, administration, attorney, chancellor, court, crime (replacing OE *sin*, which is then confined to religious vocabulary), judge, jury. Words in the religious sphere include: abbot, clergy, preach, sacrament. Words designating titles of nobility include: prince, duke, marquess, viscount, baron – and their female equivalents. In military usage English borrowed: army, captain, corporal, lieutenant, sergeant, soldier. All the names of the best-known precious stones are French: amethyst, diamond, emerald, garnet, pearl, ruby, sapphire, topaz, turquoise. French names were given to various animals when served up as food: beef, mutton, pork, veal – and their OE equivalents are restricted to the live animal (cow, sheep, etc.). Words were also borrowed for culinary processes: boil, fry, roast, stew.

During the Middle English period, some words were borrowed from Norman French spoken in England (Anglo-Norman), others from Central French, which later became the standard in France. It is

possible to tell by the form of a word whether it is of Norman or Central French origin. For instance, Old French *w* was retained in Norman French, but changed to *gw* and then *g* in standard French. Hence, the words *wage* and *warrant* were borrowed from Norman French, while *gage* and *guarantee* came from Central French. With similar sound changes, we may show that *chapter* and *chattel* came from Central French, while *cattle* is from Norman French.

It is reckoned that during this period the rate of adoption of French words by English was greater than it had ever been before or has ever been since. A statistical survey by Jespersen (1954: 86–7) of a thousand French loanwords from the *Oxford English Dictionary* shows that nearly half of the loanwords were adopted during the period. Many of the words borrowed duplicated words that already existed in English. In some cases, one of the words supplanted the other, in others both were retained but developed slightly different meanings or connotations. For example, *leod* gave way to *people*, *wlitig* to *beautiful*, and *stow* to *place*. Hundreds of Old English words were lost in this way. Cases where both OE and French words survive include: doom/ judgment, hearty/cordial, house/mansion. Conscious of the difficulties caused by such doublets, specialists compiled bilingual word lists to aid intelligibility between English and French.

Borrowing from French continued during the Modern English period, though not on the same scale. There are cases where the same French word was borrowed at different times in the history of English. For example, *gentle*, *genteel* and *jaunty* are all from French *gentil*; the last two were borrowed during the seventeenth century. Similarly, *chief* first occurred in English in the fourteenth century, and *chef* in the nineteenth. Generally speaking, doublets show by their pronunciation the approximate time of their adoption. So, the pronunciation of the 'ch' shows that *chamber*, *champion*, *chance* were borrowed in Middle English times, while *chauffeur*, *chevron*, *chiffon* were borrowed in the Modern English period. Similarly, the position of the main stress is frequently evidence of the period of borrowing. Thus, *carriage*, *courage*, *village* came into English in the Middle English period and have acquired initial stress, following the normal English pattern, while more recent borrowings, such as *prestige*, *balloon*, *mademoiselle*, do not follow this pattern.

In general, French words borrowed since the seventeenth century are less completely naturalized than older loans, e.g. amateur, boulevard, crochet, detour, ensemble, liaison, massage, nuance, rapport, vignette. Later borrowings from French also reflect French dominance in the spheres of fashion, lifestyle, arts and sciences: gown, luxury, romance, tragedy, engineer, physician. And we must not forget gastronomic terms: casserole, crème brûlée, hors d'oeuvre, restaurant.

Let us also note, finally, that English has some loan translations from French, e.g. *marriage of convenience* (*mariage de convenance*), *that goes without saying* (*cela va sans dire*), *trial balloon* (*ballon d'essai*), *reason of state* (*raison d'état*).

2.4.5 German and Dutch loans

From the Middle Ages on, commerical relationships have existed between Flemish/Dutch and English-speaking peoples. In these contacts, English borrowed from Dutch and other forms of Low German. Because of the eminence of the Dutch in seafaring activities, their language has contributed many nautical terms to English, e.g. bowline, bowsprit, buoy, commodore, cruise, deck, skipper, smuggle, yacht. The Dutch and Flemish were also famous for their cloth making and associated commercial activities. England imported the goods together with the words denoting them, e.g. cambric, duck (cloth), jacket, nap, spool. Other commercial terms include: dollar, groat, guilder, mart.

Loanwords from other Low German dialects include: boor (lover), broke, isinglass, luck, skate, snap, wagon.

A number of loanwords came into English through contact between Americans and Dutch settlers, especially in the New York area, e.g. Dutch American food items such as *coleslaw* (*koolsla*, 'cabbage salad'), *cookie*, *cranberry*, *waffle*. The diversity of the contacts accounts for the wide range of loans: boodle, boss, caboose, dope, lowery, noop, Santa Claus (Sante Klaas, 'Saint Nicholas'), spook. From South African Dutch (Afrikaans), English has borrowed: apartheid, commandeer, commando, kraal, outspan, spoor, trek, veld.

When compared to Low German, High German has had comparatively little impact on English. Words have been borrowed in specialist fields such as geology and mineralogy, e.g. cobalt, feldspar, gneiss, nickel, quartz, seltzer, zinc. Some food and drink terms have accompanied their items: delicatessen, frankfurter, noodle, schnapps. Otherwise, it is a small miscellany of borrowings: angst, ersatz, Gestalt, hinterland, leitmotiv, rucksack, umlaut, waltz, Weltanschauung.

2.4.6 Romance loans other than from French

In addition to Latin and French, English has borrowed from other Romance languages such as Spanish, Portuguese and Italian.

Most of the borrowings from Spanish and Portuguese have occurred from the sixteenth century onwards. English has borrowed directly from Spanish, and to a lesser extent from Portuguese, but in addition many non-European words from the colonies found their way

into English via these two languages. Many of these loanwords came from the New World: alligator (*el lagarto*, 'the lizard'), avocado, barracuda, canoe, chocolate, cigar, cockroach, domino, embargo, mosquito (little fly), peccadillo, potato, sombrero, tobacco, tomato, tornado, tortilla, vanilla.

In the nineteenth century it became fashionable for Americans to adopt words from Spanish, accounting for the following: bonanza, canyon, lasso, mustang, patio, ranch, sierra, siesta, stampede. The twentieth century is characterized by loan translations such as *moment of truth* (*momento de la verdad* – referring to the 'moment of kill' in a bullfight). Virtually all loanwords from Portuguese have entered English during the modern period, including: albino, copra, flamingo, madeira, mango, marmalade, molasses, palaver, teak.

Italian has had a particular significance for musical vocabulary, and other arts. Borrowing has occurred over a long period, starting in the sixteenth century with: duo, fugue, madrigal, violin. These were followed in the seventeenth century with: allegro, largo, opera, piano, presto, solo, sonata. But it was during the eighteenth century that interest in Italian music reached its peak in England, and there were numerous loanwords, including: adagio, andante, aria, cantata, concerto, crescendo, duet, finale, forte, obligato, oratorio, soprano, trio, trombone, viola. The process continued in the nineteenth century with the adoption of: alto, cadenza, legato, piccolo, prima-donna.

Other loanwords from Italian include: balcony, balloon, carnival, dilettante, fresco, ghetto, grotto, incognito, inferno, lagoon, malaria, miniature, portico, regatta, stiletto, studio, torso, umbrella, vendetta, volcano. Some Italian words came via French: cartoon, citron, corridor, gazette, porcelain. Italian food has also left its mark on English vocabulary, often by way of American English: lasagne, macaroni, pizza, scampi, etc.

2.4.7 Loans from the East

A number of words of Arabic origin were borrowed during the Middle English period, mostly to do with science and commerce. Some came via French or Latin. Examples include: admiral, amber, camphor, cipher, cotton, lute, mattress, orange, saffron, syrup, zenith. The Arabic definite article *al* is retained in one form or another in: alchemy, alcohol, algebra, algorithm, alkali, almanac, azimuth, elixir, hazard. Borrowing from Arabic has continued up to modern times, sometimes via Italian or French, including the following items: assassin, calibre, carat, garble, giraffe, hashish, lemon, magazine, sherbet.

Other Semitic languages have contributed little directly, though a number of Hebrew words have come into English, either directly (*kibbutz*) or by way of French (*amen*, *hallelujah*, *rabbi*, *sabbath*) or Yiddish (*chutzpah*).

A few words have been borrowed from Persian: caravan, bazaar, shah, shawl. Other Persian loanwords have come via Latin or French: azure, musk, paradise, scarlet, tiger.

From the Indian subcontinent, English has borrowed a few words from Sanskrit (avatar, karma, mahatma, yoga), some from Tamil (curry, pariah), and a number from Hindi/Urdu: bangle, dungaree, jungle, pajamas, shampoo.

Words borrowed from the Far East and Australia are comparatively few in number, except in the local varieties (see 6.2). But there are some very common words from that area. *Silk* may ultimately be from Chinese, though there is no known etymon in Chinese languages. Also ultimately from Chinese are *judo*, *tycoon* and *kamikaze*, but they came into English via Japanese. Directly from Japanese are: kimono, samurai, soy(a). Australia has given English *kangaroo* and *boomerang*, among a few others.

2.4.8 Loanwords from other sources

English has borrowed a few words from West African languages, mostly via Portuguese and Spanish, e.g. *banana* and *yam*, which first appeared towards the end of the sixteenth century. Likewise, *voodoo*, with its variant *hoodoo*, which came in through American English, is of African origin. *Gorilla* is apparently African in origin. More recent borrowings from this source include: okra, chimpanzee, gnu, safari, zebra.

Most of the other borrowings have been made in modern times. *Sable*, however, came into English in Middle English times, via French from Slavic languages; *polka* came via French in the nineteenth century from Czech. *Mammoth* was borrowed in the eighteenth century directly from Russian. Other more recent borrowings from Russian have not become completely naturalized: bolshevik, czar, glasnost, intelligentsia (ultimately from Latin), perestroika, tundra, vodka. From Hungarian, English has borrowed directly *goulash* and *paprika*; while *coach* came via French from Hungarian *kosci*. Turkish words in English include: fez, shish kebab.

From Native American languages have come: moccassin, toboggan, tomahawk, skunk. And many American place names have their origin in these languages, e.g. Chicago, Michigan, Saratoga, Tallahassie.

2.4.9 Sources of most recent loans

English still borrows, and is likely to continue borrowing from other languages of the world. However, borrowing in recent times is characterized by two main factors: the frequency of borrowing is considerably reduced; and English seems to be spreading its tentacles to reach and borrow from less and less known languages. A study by Garland Cannon (1987) of more than a thousand recent loanwords from 84 languages shows that 'about 25% are from French, 8% each from Japanese and Spanish, 7% each from Italian and Latin, 6% each from African languages, German and Greek, 4% each from Russian and Yiddish, 3% from Chinese, and progressively smaller percentages from Arabic, Portuguese, Hindi, Sanskrit, Hebrew, Afrikaans, Malayo-Polynesian, Vietnamese, Amerindian languages, Swedish, Bengali, Danish, Indonesian, Korean, Persian, Amharic, Eskimo-Aleut, Irish, Norwegian, and 30 other languages' (quoted in Pyles and Algeo 1993: 310).

French is still the largest supplier of words to English, which may be because of the geographical proximity of France and Britain. We would, therefore, expect more French words to enter via British English than American. Similarly, Spanish loanwords are often borrowed from American Spanish into American English. The increase in the importance of Japanese as a source of loans is probably a consequence of the increased commercial importance of Japan in the world generally. The decline of Latin as a source of loanwords may have a dual explanation: on the one hand, it may be argued that English has already borrowed so much of Latin vocabulary that there is relatively little left to be borrowed; on the other hand, rather than borrowing directly from Latin, English now often makes new Latinate words from English morphemes originally from Latin.

EXERCISE 2/7

The previous sections have illustrated just how many words in the modern English vocabulary are borrowed from other languages. Without going back over these sections, peruse the etymologies in your dictionary and list five words that have been borrowed into English at some time or other from each of the following languages:

French, Italian, Spanish, Latin, Greek.

2.5 Creating new English words

This section briefly reviews the processes used in creating new words in English, other than by borrowing. They are discussed in more detail in Chapter 4.

2.5.1 Root creation

'Root creation' refers to the building of a word that has no relationship whatsoever with any previously existing word. An often quoted example is *Kodak*, which was first used in print in the *USA Patent Office Gazette* of 1888. According to George Eastman, who invented the word and the device it referred to, *Kodak* is a 'purely arbitrary combination of letters, not derived in whole or part from any existing word' (quoted in Pyles and Algeo 1993: 258). Most tradenames, however, which may look like root creation, are usually suggested by already existing words. For example, *Vaseline* was created from German *Wasser*, 'water' and Greek *elaion*, 'oil'; *Kleenex* derives from *clean* and *curtex*. In the course of history, very few words have been coined by root creation.

2.5.2 Echoic words

'Echoic' or 'onomatopoeic' words find their origin in the specific sound that they are meant to represent, e.g. bang, burp, cuckoo, ping, splash, tinkle. Bloomfield (1933: 156) distinguishes two types of echoic words: 'imitative' and 'symbolic'. The first type intends to imitate the sound that it represents: meow, moo, bow-wow, vroom. The second has a less direct association with the sound: bump, flick. Symbolic words often come in sets, which either rhyme (bump, lump, clump, hump) or alliterate (flick, flash, flip, flop). Both imitative and symbolic words may be subject to the process of doubling, sometimes with a little variation, e.g. bow-wow, choo-choo, peewee.

2.5.3 Ejaculations

'Ejaculations' are words that attempt to imitate instinctive vocal responses to emotional situations (Pyles and Algeo 1993). They are also termed 'natural utterances', but they have become conventionalized and so become lexical items, e.g. *ha-ha* and its variant *ho-ho* for laughter. The words *pish* and *pshaw* express disdain, contempt, impatience or irritation. *Phew* imitates the reaction to a bad

smell or to the avoidance of a disaster. To signify agreement, *uh-huh* is used.

2.5.4 Word formation

Word formation uses existing language material – words and morphemes – to create new lexical items. The processes were already well established in Old English. The major processes are compounding and affixation, which will be introduced here and elaborated on in Chapter 4.

The following nouns were produced by compounding, the combination of two words (free morphemes): birthday, craftsman, grandfather, highway. Compounds are also found among adjectives (newborn, red-hot, banana-flavoured, freeze-dried), verbs (download, safeguard, stagemanage), prepositions/adverbs (inside, inspite of, afterwards).

Affixation involves the use of prefixes and suffixes to form new words. As the language evolved, so the range of affixes increased, especially by borrowing from French and Latin. Here are some examples: celebrat-ion, duck-ling, forget-ful, green-ish, nap-kin, pre-figure, un-know-able.

2.6 Characteristics of modern English vocabulary

Among the characteristics of present-day English vocabulary, we shall discuss the size of the vocabulary and the frequency of occurrence of Anglo-Saxon words, before concluding with a discussion of the 'Englishness' of the vocabulary.

2.6.1 The size of the vocabulary

One way of answering the question, 'How large is the English lexicon?' would be to compare two of the biggest dictionaries of the English language. David Crystal compared the unabridged *Websters Third New International* (1961), which claims over 450,000 entries, and the integrated second edition of the *Oxford English Dictionary* (1989), which claims over 600,000 entries. Crystal's sample comparison revealed a remarkable lack of identity between the headword lists: the *Websters* and *Oxford* dictionaries had only 21 headwords in common out of a possible 57, less than two-fifths. If this disparity were to be repeated across all the entries, then the combined lexicon of both dictionaries would exceed 750,000. The discrepancies observed may be accounted for by differences in editorial emphasis. For example, the *Oxford* has far more historical references and British dialect terms than *Websters*, which in turn has many more local American items. Furthermore, neither work would claim a comprehensive coverage of the 'New Englishes', used in areas such as India, West Africa, Singapore, where thousands of new lexemes are coming into the language. Finally, the tradition of lexicography, which gives preference to the written language as the test for inclusion, will exclude thousands of words and expressions that have never been recorded even if they are common in current spoken use.

Crystal argues that, even if we restrict the issue to standard vocabulary, there are many items that could be included as part of the lexicon, but which are not usually found in a dictionary. Many of the abbreviated forms and acronyms have a clear lexical status (e.g. BA, FBI, NATO, UNICEF), as well as fauna and flora local to diverse English-speaking areas. Crystal concludes his analysis by noting that it is difficult to see how even a conservative estimate of English vocabulary could go much below one million lexemes. More inclusive counts, to encompass all of scientific nomenclature, might easily double this figure. Only a small fraction of these totals is, of course, learned and used by any one of us (Crystal 1995: 119).

2.6.2 Frequency of occurrence and use of Anglo-Saxon words

In addition to being short and concrete, words from Anglo-Saxon have two other characteristics: they are the most frequent in the language; and they are considered 'warmer' than words of foreign origin.

The most frequent two hundred words in both British (BrE) and American English (AmE) consist overwhelmingly of one syllable. There are a few two-syllable words (40 in AmE, 24 in BrE) and a handful of trisyllabic forms (3 in AmE, 2 in BrE). Only AmE has a single four-syllable item, the word *American* itself. It has been estimated that for the 10,000 most frequent words in English, nearly 32 per cent have their origin in Old English (Crystal 1995: 18). In the one-million-word 'Brown University Corpus of Written American English' from the 1960s, the hundred most frequently used items are almost all from Anglo-Saxon. The exceptions are a few Scandinavian loans (e.g. *they*, *are*). The first words borrowed from Romance sources are *just* (at 105th) and *people* (at 107th) (Crystal 1995: 124). These statistics show the importance of the inherited Germanic vocabulary in the central core of English vocabulary. However, it should be noted that the relative frequency of words varies not only according to the type of text but also according to the stylistic level. Generally speaking, formal

style and specialized language use a greater proportion of foreign loans than does everyday conversation.

As regards the second characteristic, it is argued that native English words tend to be preferred in everyday speech because they are vague and convey many shades of meaning, as opposed to loanwords, which are more precise and restricted, and so are more difficult to handle. Furthermore, native English words are considered more human and emotional, whereas many polysyllabic loans from Greek, Latin or the Romance language are considered cold and formal. For example, in an informal everyday situation, when faced with the choice between *initiate*, *commence* and *start*, or between *nourishment*, *nutrition* and *food*, most people would opt for the short, Anglo-Saxon word. In formal situations, however, it may seem more appropriate to *allude* to a *nauseating odour* or even an *obnoxious effluvium* rather than a *nasty smell*.

2.6.3 'English' vocabulary

This chapter has shown the very cosmopolitan nature of English vocabulary, a prime example of a lexically mixed language. How then can we speak of the 'English' vocabulary in the face of this avalanche of words from other languages?

It should be noted, first, that the predominance of foreign words is felt only with reference to the total word stock, for example if we consider the 616,500 word forms in the *Oxford English Dictionary* (1989). But when we look at the items actually used in writing and speech, we find that the front-runners, i.e. the most common in the language, are all native English (Anglo-Saxon) words (see examples in 2.3.1 and Exercise 2/6). Moreover, a good many foreign words have been assimilated to the pronunciation and spelling of English, so they are no longer recognized as 'foreign' by native speakers. And the grammatical structures used have remained essentially the same throughout the various stages in the development of the language. English has, from Middle English times, been unable to afford to be purist in its attitude to the lexicon.

EXERCISE 2/8

Here are some English words that still betray which language they have been borrowed from. Without looking them up, make an intelligent guess at their language of origin. Then look in the Key to Exercises.

addendum (plural, addenda), baguette, cannelloni, con brio, criterion (plural, criteria), id est (i.e.), in loco parentis, mañana, sang-froid, vis-à-vis, zucchini

2.7 Summary

In this chapter we have examined the origins of English words. In the main, we have focused our attention first on the native English vocabulary, and second on words borrowed from other languages. Before discussing these main categories of words, it was necessary not only to place English in the context of world languages, but also to show how English vocabulary evolved from the Old to the Modern English periods. Finally, we touched on the creation of new words in English, before highlighting the most important characteristics of modern English vocabulary.