Knowledge of English Language

Bryce Gordon

KNOWLEDGE OF ENGLISH LANGUAGE

KNOWLEDGE OF ENGLISH LANGUAGE

Bryce Gordon

Knowledge of English Language
by Bryce Gordon

Contents

Chapter 1

Introduction

English is a West Germanic language of the Indo-European language family, originally spoken by the inhabitants of early medieval England. It is named after the Angles, one of the ancient Germanic peoples that migrated to the area of Great Britain that later took their name, England. Both names derive from Anglia, a peninsula on the Baltic Sea which is not to be confused with East Anglia, the Eastern part of England which comprises the counties of Norfolk, Suffolk and Essex. English is most closely related to Frisian and Low Saxon, while its vocabulary has been significantly influenced by other Germanic languages, particularly Old Norse (a North Germanic language), as well as Latin and French.

English has developed over the course of more than 1,400 years. The earliest forms of English, a group of West Germanic (Ingvaeonic) dialects brought to Great Britain by Anglo-Saxon settlers in the 5th century, are collectively called Old English. Middle English began in the late 11th century with the Norman conquest of England; this was a period in which English was influenced by Old French, in particular through its Old Norman dialect. Early Modern English began in the late 15th century with the introduction of the printing press to London, the printing of the King James Bible and the start of the Great Vowel Shift.

Modern English has been spreading around the world since the 17th century by the worldwide influence of the British Empire and the United States. Through all types of printed and

electronic media of these countries, English has become the leading language of international discourse and the *lingua franca* in many regions and professional contexts such as science, navigation and law. Modern English grammar is the result of a gradual change from a typical Indo-European dependent-marking pattern, with a rich inflectional morphology and relatively free word order, to a mostly analytic pattern with little inflection, a fairly fixed subject–verb–object word order and a complex syntax. Modern English relies more on auxiliary verbs and word order for the expression of complex tenses, aspect and mood, as well as passive constructions, interrogatives and some negation.

English is the most spoken language in the world and the third-most spoken native language in the world, after Standard Chinese and Spanish. It is the most widely learned second language and is either the official language or one of the official languages in almost 60 sovereign states. There are more people who have learned English as a second language than there are native speakers. As of 2005, it was estimated that there were over 2 billion speakers of English. English is the majority native language in the United States, the United Kingdom, Canada, Australia, New Zealand and Ireland, an official language and the main language of Singapore, and it is widely spoken in some areas of the Caribbean, Africa, South Asia, Southeast Asia and Oceania. It is a co-official language of the United Nations, the European Union and many other world and regional international organisations. It is the most widely spoken Germanic language, accounting for at least 70% of speakers of this Indo-European branch. English speakers are called "Anglophones". There is much variability among the many accents and dialects of English used in different

countries and regions in terms of phonetics and phonology, and sometimes also vocabulary, idioms, grammar, and spelling, but it does not typically prevent understanding by speakers of other dialects and accents, although mutual unintelligibility can occur at extreme ends of the dialect continuum.

Classification

English is an Indo-European language and belongs to the West Germanic group of the Germanic languages. Old Englishoriginated from a Germanic tribal and linguistic continuum along the FrisianNorth Sea coast, whose languages gradually evolved into the Anglic languages in the British Isles, and into the Frisian languages and Low German/Low Saxon on the continent. The Frisian languages, which together with the Anglic languages form the Anglo-Frisian languages, are the closest living relatives of English. Low German/Low Saxon is also closely related, and sometimes English, the Frisian languages, and Low Germanare grouped together as the Ingvaeonic (North Sea Germanic) languages, though this grouping remains debated. Old English evolved into Middle English, which in turn evolved into Modern English. Particular dialects of Old and Middle English also developed into a number of other Anglic languages, including Scots and the extinct Fingallian and Forth and Bargy (Yola) dialects of Ireland.

Like Icelandic and Faroese, the development of English in the British Isles isolated it from the continental Germanic languages and influences, and it has since diverged considerably. English is not mutually intelligible with any continental Germanic language, differing in vocabulary,

syntax, and phonology, although some of these, such as Dutch or Frisian, do show strong affinities with English, especially with its earlier stages.

Unlike Icelandic and Faroese, which were isolated, the development of English was influenced by a long series of invasions of the British Isles by other peoples and languages, particularly Old Norse and Norman French. These left a profound mark of their own on the language, so that English shows some similarities in vocabulary and grammar with many languages outside its linguistic clades—but it is not mutually intelligible with any of those languages either. Some scholars have argued that English can be considered a mixed language or a creole—a theory called the Middle English creole hypothesis. Although the great influence of these languages on the vocabulary and grammar of Modern English is widely acknowledged, most specialists in language contact do not consider English to be a true mixed language.

English is classified as a Germanic language because it shares innovations with other Germanic languages such as Dutch, German, and Swedish. These shared innovations show that the languages have descended from a single common ancestor called Proto-Germanic. Some shared features of Germanic languages include the division of verbs into strong and weak classes, the use of modal verbs, and the sound changes affecting Proto-Indo-European consonants, known as Grimm's and Verner's laws. English is classified as an Anglo-Frisian language because Frisian and English share other features, such as the palatalisation of consonants that were velar consonants in Proto-Germanic (see Phonological history of Old English § Palatalization).

History

Proto-Germanic to Old English

The earliest form of English is calledOld English or Anglo-Saxon (c. year 550–1066). Old English developed from a set of West Germanic dialects, often grouped as Anglo-Frisian or North Sea Germanic, and originally spoken along the coasts of Frisia, Lower Saxony and southern Jutland by Germanic peoples known to the historical record as the Angles, Saxons, and Jutes. From the 5th century, the Anglo-Saxons settled Britain as the Roman economy and administration collapsed. By the 7th century, the Germanic language of the Anglo-Saxons became dominant in Britain, replacing the languages of Roman Britain (43–409): Common Brittonic, a Celtic language, and Latin, brought to Britain by the Roman occupation. *England* and *English* (originally *Ænglaland* and *Ænglisc*) are named after the Angles.

Old English was divided into four dialects: the Anglian dialects (Mercian and Northumbrian) and the Saxon dialects, Kentish and West Saxon. Through the educational reforms of King Alfred in the 9th century and the influence of the kingdom of Wessex, the West Saxon dialect became the standard written variety. The epic poem*Beowulf*is written in West Saxon, and the earliest English poem, *Cædmon's Hymn*, is written in Northumbrian. Modern English developed mainly from Mercian, but the Scots language developed from Northumbrian. A few short inscriptions from the early period of Old English were written using a runic script. By the 6th century, a Latin alphabetwas adopted, written with half-uncialletterforms. It

included the runic letters *wynn*⟨⟩ and *thorn*⟨þ⟩, and the modified Latin letters *eth*⟨ð⟩, and *ash*⟨æ⟩.

Old English is essentially a distinct language from Modern English and is virtually impossible for 21st-century unstudied English-speakers to understand. Its grammar was similar to that of modern German, and its closest relative is Old Frisian. Nouns, adjectives, pronouns, and verbs had many more inflectional endings and forms, and word order was much freer than in Modern English. Modern English has case forms in pronouns (*he*, *him*, *his*) and has a few verb inflections (*speak*, *speaks*, *speaking*, *spoke*, *spoken*), but Old English had case endings in nouns as well, and verbs had more person and number endings.

The translation of Matthew 8:20 from 1000 shows examples of case endings (nominative plural, accusative plural, genitive singular) and a verb ending (present plural):

- *Foxas habbað holu and heofonan fuglas nest*
- Fox-as habb-að hol-u and heofon-an fugl-as nest-∅
- fox-NOM.PL have-PRS.PL hole-ACC.PL and heaven-GEN.SG bird-NOM.PL nest-ACC.PL
- "Foxes have holes and the birds of heaven nests"

Middle English

From the 8th to the 12th century, Old English gradually transformed through language contact into Middle English. Middle English is often arbitrarily defined as beginning with the conquest of England by William the Conqueror in 1066, but it developed further in the period from 1200 to 1450.

First, the waves of Norse colonisation of northern parts of the British Isles in the 8th and 9th centuries put Old English into intense contact with Old Norse, a North Germanic language. Norse influence was strongest in the north-eastern varieties of Old English spoken in the Danelaw area around York, which was the centre of Norse colonisation; today these features are still particularly present in Scots and Northern English. However the centre of norsified English seems to have been in the Midlands around Lindsey, and after 920 CE when Lindsey was reincorporated into the Anglo-Saxon polity, Norse features spread from there into English varieties that had not been in direct contact with Norse speakers. An element of Norse influence that persists in all English varieties today is the group of pronouns beginning with *th-* (*they, them, their*) which replaced the Anglo-Saxon pronouns with *h-* (*hie, him, hera*).

With the Norman conquest of England in 1066, the now norsified Old English language was subject to contact with Old French, in particular with the Old Norman dialect. The Norman language in England eventually developed into Anglo-Norman. Because Norman was spoken primarily by the elites and nobles, while the lower classes continued speaking Anglo-Saxon (English), the main influence of Norman was the introduction of a wide range of loanwords related to politics, legislation and prestigious social domains. Middle English also greatly simplified the inflectional system, probably in order to reconcile Old Norse and Old English, which were inflectionally different but morphologically similar. The distinction between nominative and accusative cases was lost except in personal pronouns, the instrumental case was dropped, and the use of the genitive case was limited to indicating possession. The inflectional system regularised many irregular inflectional

forms, and gradually simplified the system of agreement, making word order less flexible. In the Wycliffe Bible of the 1380s, the verse Matthew 8:20 was written: *Foxis han dennes, and briddis of heuene han nestis* Here the plural suffix *-n* on the verb *have* is still retained, but none of the case endings on the nouns are present. By the 12th century Middle English was fully developed, integrating both Norse and French features; it continued to be spoken until the transition to early Modern English around 1500. Middle English literature includes Geoffrey Chaucer's *The Canterbury Tales*, and Malory's*Le Morte d'Arthur.* In the Middle English period, the use of regional dialects in writing proliferated, and dialect traits were even used for effect by authors such as Chaucer.

Early Modern English

The next period in the history of English was Early Modern English (1500–1700). Early Modern English was characterised by the Great Vowel Shift (1350–1700), inflectional simplification, and linguistic standardisation.

The Great Vowel Shift affected the stressed long vowels of Middle English. It was a chain shift, meaning that each shift triggered a subsequent shift in the vowel system. Mid and open vowels were raised, and close vowels were broken into diphthongs. For example, the word *bite*was originally pronounced as the word *beet* is today, and the second vowel in the word *about* was pronounced as the word *boot* is today. The Great Vowel Shift explains many irregularities in spelling since English retains many spellings from Middle English, and it also explains why English vowel letters have very different pronunciations from the same letters in other languages.

English began to rise in prestige, relative to Norman French, during the reign of Henry V. Around 1430, the Court of Chancery in Westminster began using English in its official documents, and a new standard form of Middle English, known as Chancery Standard, developed from the dialects of London and the East Midlands. In 1476, William Caxton introduced the printing press to England and began publishing the first printed books in London, expanding the influence of this form of English. Literature from the Early Modern period includes the works of William Shakespeare and the translation of the Bible commissioned by King James I. Even after the vowel shift the language still sounded different from Modern English: for example, the consonant clusters/kn ɡn sw/ in *knight*, *gnat*, and *sword* were still pronounced. Many of the grammatical features that a modern reader of Shakespeare might find quaint or archaic represent the distinct characteristics of Early Modern English.

In the 1611 King James Version of the Bible, written in Early Modern English, Matthew 8:20 says, "The Foxes haue holes and the birds of the ayre haue nests." This exemplifies the loss of case and its effects on sentence structure (replacement with subject–verb–object word order, and the use of *of* instead of the non-possessive genitive), and the introduction of loanwords from French (*ayre*) and word replacements (*bird* originally meaning "nestling" had replaced OE *fugol*).

Spread of Modern English

By the late 18th century, the British Empire had spread English through its colonies and geopolitical dominance. Commerce, science and technology, diplomacy, art, and formal

education all contributed to English becoming the first truly global language. English also facilitated worldwide international communication. England continued to form new colonies, and these later developed their own norms for speech and writing. English was adopted in parts of North America, parts of Africa, Australasia, and many other regions. When they obtained political independence, some of the newly independent nations that had multiple indigenous languages opted to continue using English as the official language to avoid the political and other difficulties inherent in promoting any one indigenous language above the others. In the 20th century the growing economic and cultural influence of the United States and its status as a superpower following the Second World War has, along with worldwide broadcasting in English by the BBC and other broadcasters, caused the language to spread across the planet much faster. In the 21st century, English is more widely spoken and written than any language has ever been.

As Modern English developed, explicit norms for standard usage were published, and spread through official media such as public education and state-sponsored publications. In 1755Samuel Johnson published his *A Dictionary of the English Language* which introduced standard spellings of words and usage norms. In 1828, Noah Webster published the *American Dictionary of the English language* to try to establish a norm for speaking and writing American English that was independent of the British standard. Within Britain, non-standard or lower class dialect features were increasingly stigmatised, leading to the quick spread of the prestige varieties among the middle classes.

In modern English, the loss of grammatical case is almost complete (it is now only found in pronouns, such as *he* and *him*, *she* and *her*, *who* and *whom*), and SVO word order is mostly fixed. Some changes, such as the use of do-support, have become universalised. (Earlier English did not use the word "do" as a general auxiliary as Modern English does; at first it was only used in question constructions, and even then was not obligatory. Now, do-support with the verb *have* is becoming increasingly standardised.) The use of progressive forms in *-ing*, appears to be spreading to new constructions, and forms such as *had been being built* are becoming more common. Regularisation of irregular forms also slowly continues (e.g. *dreamed* instead of *dreamt*), and analytical alternatives to inflectional forms are becoming more common (e.g. *more polite* instead of *politer*). British English is also undergoing change under the influence of American English, fuelled by the strong presence of American English in the media and the prestige associated with the US as a world power.

Geographical distribution

As of 2016, 400 million people spoke English as their first language, and 1.1 billion spoke it as a secondary language. English is the largest language by number of speakers. English is spoken by communities on every continent and on islands in all the major oceans.

The countries where English is spoken can be grouped into different categories according to how English is used in each country. The "inner circle" countries with many native speakers of English share an international standard of written

English and jointly influence speech norms for English around the world. English does not belong to just one country, and it does not belong solely to descendants of English settlers. English is an official language of countries populated by few descendants of native speakers of English. It has also become by far the most important language of international communication when people who share no native language meet anywhere in the world.

Three circles of English-speaking countries

The Indian linguist Braj Kachru distinguished countries where English is spoken with a three circles model. In his model,

- the "inner circle" countries have large communities of native speakers of English,
- "outer circle" countries have small communities of native speakers of English but widespread use of English as a second language in education or broadcasting or for local official purposes, and
- "expanding circle" countries are countries where many people learn English as a foreign language.

Kachru based his model on the history of how English spread in different countries, how users acquire English, and the range of uses English has in each country. The three circles change membership over time.

Countries with large communities of native speakers of English (the inner circle) include Britain, the United States, Australia, Canada, Ireland, and New Zealand, where the majority speaks English, and South Africa, where a significant minority speaks English. The countries with the most native English speakers

are, in descending order, the United States (at least 231 million), the United Kingdom (60 million), Canada (19 million), Australia (at least 17 million), South Africa (4.8 million), Ireland (4.2 million), and New Zealand (3.7 million). In these countries, children of native speakers learn English from their parents, and local people who speak other languages and new immigrants learn English to communicate in their neighbourhoods and workplaces. The inner-circle countries provide the base from which English spreads to other countries in the world.

Estimates of the numbers of second language and foreign-language English speakers vary greatly from 470 million to more than 1 billion, depending on how proficiency is defined. Linguist David Crystal estimates that non-native speakers now outnumber native speakers by a ratio of 3 to 1. In Kachru's three-circles model, the "outer circle" countries are countries such as the Philippines, Jamaica, India, Pakistan, Malaysia and Nigeria with a much smaller proportion of native speakers of English but much use of English as a second language for education, government, or domestic business, and its routine use for school instruction and official interactions with the government.

Those countries have millions of native speakers of dialect continua ranging from an English-based creole to a more standard version of English. They have many more speakers of English who acquire English as they grow up through day-to-day use and listening to broadcasting, especially if they attend schools where English is the medium of instruction. Varieties of English learned by non-native speakers born to English-speaking parents may be influenced, especially in their

grammar, by the other languages spoken by those learners. Most of those varieties of English include words little used by native speakers of English in the inner-circle countries, and they may show grammatical and phonological differences from inner-circle varieties as well. The standard English of the inner-circle countries is often taken as a norm for use of English in the outer-circle countries.

In the three-circles model, countries such as Poland, China, Brazil, Germany, Japan, Indonesia, Egypt, and other countries where English is taught as a foreign language, make up the "expanding circle". The distinctions between English as a first language, as a second language, and as a foreign language are often debatable and may change in particular countries over time. For example, in the Netherlands and some other countries of Europe, knowledge of English as a second language is nearly universal, with over 80 percent of the population able to use it, and thus English is routinely used to communicate with foreigners and often in higher education. In these countries, although English is not used for government business, its widespread use puts them at the boundary between the "outer circle" and "expanding circle". English is unusual among world languages in how many of its users are not native speakers but speakers of English as a second or foreign language.

Many users of English in the expanding circle use it to communicate with other people from the expanding circle, so that interaction with native speakers of English plays no part in their decision to use the language. Non-native varieties of English are widely used for international communication, and speakers of one such variety often encounter features of other

varieties. Very often today a conversation in English anywhere in the world may include no native speakers of English at all, even while including speakers from several different countries. This is particularly true of the shared vocabulary of the scientifical and mathematical fields of life.

Pluricentric English

English is a pluricentric language, which means that no one national authority sets the standard for use of the language. Spoken English, for example English used in broadcasting, generally follows national pronunciation standards that are also established by custom rather than by regulation. International broadcasters are usually identifiable as coming from one country rather than another through their accents, but newsreader scripts are also composed largely in international standard written English.

The norms of standard written English are maintainedpurely by the consensus of educated English-speakers around the world, without any oversight by any government or international organisation.

American listeners generally readily understand most British broadcasting, and British listeners readily understand most American broadcasting. Most English speakers around the world can understand radio programmes, television programmes, and films from many parts of the English-speaking world. Both standard and non-standard varieties of English can include both formal or informal styles, distinguished by word choice and syntax and use both technical and non-technical registers.

The settlement history of the English-speaking inner circle countries outside Britain helped level dialect distinctions and produce koineised forms of English in South Africa, Australia, and New Zealand. The majority of immigrants to the United States without British ancestry rapidly adopted English after arrival. Now the majority of the United States population are monolingual English speakers, and English has been given official or co-official status by 30 of the 50 state governments, as well as all five territorial governments of the US, though there has never been an official language at the federal level.

English as a global language

English has ceased to be an "English language" in the sense of belonging only to people who are ethnically English. Use of English is growing country-by-country internally and for international communication. Most people learn English for practical rather than ideological reasons. Many speakers of English in Africa have become part of an "Afro-Saxon" language community that unites Africans from different countries.

As decolonisation proceeded throughout the British Empire in the 1950s and 1960s, former colonies often did not reject English but rather continued to use it as independent countries setting their own language policies. For example, the view of the English language among many Indians has gone from associating it with colonialism to associating it with economic progress, and English continues to be an official language of India. English is also widely used in media and literature, and the number of English language books published annually in India is the third largest in the world after the US and UK. However English is rarely spoken as a

first language, numbering only around a couple hundred-thousand people, and less than 5% of the population speak fluent English in India. David Crystal claimed in 2004 that, combining native and non-native speakers, India now has more people who speak or understand English than any other country in the world, but the number of English speakers in India is very uncertain, with most scholars concluding that the United States still has more speakers of English than India.

Modern English, sometimes described as the first global lingua franca, is also regarded as the first world language. English is the world's most widely used language in newspaper publishing, book publishing, international telecommunications, scientific publishing, international trade, mass entertainment, and diplomacy. English is, by international treaty, the basis for the required controlled natural languages Seaspeak and Airspeak, used as international languages of seafaring and aviation. English used to have parity with French and German in scientific research, but now it dominates that field. It achieved parity with French as a language of diplomacy at the Treaty of Versailles negotiations in 1919. By the time of the foundation of the United Nations at the end of World War II, English had become pre-eminent and is now the main worldwide language of diplomacy and international relations. It is one of six official languages of the United Nations. Many other worldwide international organisations, including the International Olympic Committee, specify English as a working language or official language of the organisation.

Many regional international organisations such as the European Free Trade Association, Association of Southeast

Asian Nations (ASEAN), and Asia-Pacific Economic Cooperation (APEC) set English as their organisation's sole working language even though most members are not countries with a majority of native English speakers. While the European Union (EU) allows member states to designate any of the national languages as an official language of the Union, in practice English is the main working language of EU organisations.

Although in most countries English is not an official language, it is currently the language most often taught as a foreign language. In the countries of the EU, English is the most widely spoken foreign language in nineteen of the twenty-five member states where it is not an official language (that is, the countries other than Ireland and Malta). In a 2012 official Eurobarometer poll (conducted when the UK was still a member of the EU), 38 percent of the EU respondents outside the countries where English is an official language said they could speak English well enough to have a conversation in that language. The next most commonly mentioned foreign language, French (which is the most widely known foreign language in the UK and Ireland), could be used in conversation by 12 percent of respondents.

A working knowledge of English has become a requirement in a number of occupations and professions such as medicine and computing. English has become so important in scientific publishing that more than 80 percent of all scientific journal articles indexed by *Chemical Abstracts* in 1998 were written in English, as were 90 percent of all articles in natural science publications by 1996 and 82 percent of articles in humanities publications by 1995.

International communities such as international business people may use English as an auxiliary language, with an emphasis on vocabulary suitable for their domain of interest. This has led some scholars to develop the study of English as an auxiliary language.

The trademarked Globish uses a relatively small subset of English vocabulary (about 1500 words, designed to represent the highest use in international business English) in combination with the standard English grammar. Other examples include Simple English.

The increased use of the English language globally has had an effect on other languages, leading to some English words being assimilated into the vocabularies of other languages.

This influence of English has led to concerns about language death, and to claims of linguistic imperialism, and has provoked resistance to the spread of English; however the number of speakers continues to increase because many people around the world think that English provides them with opportunities for better employment and improved lives.

Although some scholars mention a possibility of future divergence of English dialects into mutually unintelligible languages, most think a more likely outcome is that English will continue to function as a koineised language in which the standard form unifies speakers from around the world. English is used as the language for wider communication in countries around the world. Thus English has grown in worldwide use much more than any constructed language proposed as an international auxiliary language, including Esperanto.

Phonology

The phonetics and phonology of the English language differ from one dialect to another, usually without interfering with mutual communication. Phonological variation affects the inventory of phonemes (i.e. speech sounds that distinguish meaning), and phonetic variation consists in differences in pronunciation of the phonemes. This overview mainly describes the standard pronunciations of the United Kingdom and the United States: Received Pronunciation (RP) and General American (GA). (See § Dialects, accents, and varieties, below.)

The phonetic symbols used below are from the International Phonetic Alphabet (IPA).

Phonotactics

An English syllable includes a syllable nucleus consisting of a vowel sound. Syllable onset and coda (start and end) are optional. A syllable can start with up to three consonant sounds, as in *sprint* /sprɪnt/, and end with up to four, as in *texts* /teksts/. This gives an English syllable the following structure, (CCC)V(CCCC) where C represents a consonant and V a vowel; the word *strengths* /strɛŋkθs/ is thus an example of the most complex syllable possible in English. The consonants that may appear together in onsets or codas are restricted, as is the order in which they may appear. Onsets can only have four types of consonant clusters: a stop and approximant, as in *play*; a voiceless fricative and approximant, as in *fly* or *sly*; *s* and a voiceless stop, as in *stay*; and *s*, a voiceless stop, and

an approximant, as in *string.* Clusters of nasal and stop are only allowed in codas. Clusters of obstruents always agree in voicing, and clusters of sibilants and of plosives with the same point of articulation are prohibited. Furthermore, several consonants have limited distributions: /h/ can only occur in syllable-initial position, and /ŋ/ only in syllable-final position.

Stress, rhythm and intonation

Stress plays an important role in English. Certain syllablesare stressed, while others are unstressed. Stress is a combination of duration, intensity, vowel quality, and sometimes changes in pitch. Stressed syllables are pronounced longer and louder than unstressed syllables, and vowels in unstressed syllables are frequently reduced while vowels in stressed syllables are not. Some words, primarily short function words but also some modal verbs such as *can*, have weak and strong forms depending on whether they occur in stressed or non-stressed position within a sentence.

Stress in English is phonemic, and some pairs of words are distinguished by stress. For instance, the word *contract* is stressed on the first syllable (/ˈkɒntrækt/*KON-trakt*) when used as a noun, but on the last syllable (/kənˈtrækt/*kən-TRAKT*) for most meanings (for example, "reduce in size") when used as a verb. Here stress is connected to vowel reduction: in the noun "contract" the first syllable is stressed and has the unreduced vowel /ɒ/, but in the verb "contract" the first syllable is unstressed and its vowel is reduced to /ə/. Stress is also used to distinguish between words and phrases, so that a compound word receives a single stress unit, but the corresponding phrase has two: e.g. *a burnout* (/ˈbɜːrnaʊt/) versus *to burn out*

(/ˈbɜːrnˈaʊt/), and *a hotdog* (/ˈhɒtdɒɡ/) versus *a hot dog* (/ˈhɒtˈdɒɡ/).

In terms of rhythm, English is generally described as a stress-timed language, meaning that the amount of time between stressed syllables tends to be equal. Stressed syllables are pronounced longer, but unstressed syllables (syllables between stresses) are shortened. Vowels in unstressed syllables are shortened as well, and vowel shortening causes changes in vowel quality: vowel reduction.

Varieties of English vary the most in pronunciation of vowels. The best known national varieties used as standards for education in non-English-speaking countries are British (BrE) and American (AmE). Countries such as Canada, Australia, Ireland, New Zealand and South Africa have their own standard varieties which are less often used as standards for education internationally. Some differences between the various dialects are shown in the table "Varieties of Standard English and their features".

English has undergone many historical sound changes, some of them affecting all varieties, and others affecting only a few. Most standard varieties are affected by the Great Vowel Shift, which changed the pronunciation of long vowels, but a few dialects have slightly different results. In North America, a number of chain shifts such as the Northern Cities Vowel Shift and Canadian Shift have produced very different vowel landscapes in some regional accents.

Some dialects have fewer or more consonant phonemes and phones than the standard varieties. Some conservative varieties like Scottish English have a voiceless[ʍ] sound in

whine that contrasts with the voiced [w] in *wine*, but most other dialects pronounce both words with voiced [w], a dialect feature called *wine–whine* merger. The unvoiced velar fricative sound /x/is found in Scottish English, which distinguishes *loch*/lɔx/ from *lock*/lɔk/.

Accents like Cockney with "*h*-dropping" lack the glottal fricative /h/, and dialects with *th*-stopping and *th*-fronting like African American Vernacular and Estuary English do not have the dental fricatives /θ, ð/, but replace them with dental or alveolar stops /t, d/ or labiodental fricatives /f, v/. Other changes affecting the phonology of local varieties are processes such as *yod*-dropping, *yod*-coalescence, and reduction of consonant clusters.

General American and Received Pronunciation vary in their pronunciation of historical /r/ after a vowel at the end of a syllable (in the syllable coda).

GA is a rhotic dialect, meaning that it pronounces /r/ at the end of a syllable, but RP is non-rhotic, meaning that it loses /r/ in that position. English dialects are classified as rhotic or non-rhotic depending on whether they elide /r/ like RP or keep it like GA.

There is complex dialectal variation in words with the open front and open back vowels/æ ɑːɒɔː/. These four vowels are only distinguished in RP, Australia, New Zealand and South Africa. In GA, these vowels merge to three /æ ɑɔ/, and in Canadian English, they merge to two /æ ɑ/. In addition, the words that have each vowel vary by dialect. The table "Dialects and open vowels" shows this variation with lexical sets in which these sounds occur.

Grammar

As is typical of an Indo-European language, English follows accusativemorphosyntactic alignment. Unlike other Indo-European languages though, English has largely abandoned the inflectional case system in favor of analytic constructions. Only the personal pronouns retain morphological case more strongly than any other word class.

English distinguishes at least seven major word classes: verbs, nouns, adjectives, adverbs, determiners (including articles), prepositions, and conjunctions.

Some analyses add pronouns as a class separate from nouns, and subdivide conjunctions into subordinators and coordinators, and add the class of interjections.

English also has a rich set of auxiliary verbs, such as *have* and *do*, expressing the categories of mood and aspect. Questions are marked by do-support, wh-movement (fronting of question words beginning with *wh-*) and word order inversion with some verbs.

Some traits typical of Germanic languages persist in English, such as the distinction between irregularly inflected strong stems inflected through ablaut (i.e. changing the vowel of the stem, as in the pairs *speak/spoke* and *foot/feet*) and weak stems inflected through affixation (such as *love/loved*, *hand/hands*). Vestiges of the case and gender system are found in the pronoun system (*he/him, who/whom*) and in the inflection of the copula verb *to be*.

Nouns and noun phrases

English nouns are only inflected for number and possession. New nouns can be formed through derivation or compounding. They are semantically divided into proper nouns (names) and common nouns. Common nouns are in turn divided into concrete and abstract nouns, and grammatically into count nouns and mass nouns.

Most count nouns are inflected for plural number through the use of the plural suffix *-s*, but a few nouns have irregular plural forms. Mass nouns can only be pluralised through the use of a count noun classifier, e.g. *one loaf of bread*, *two loaves of bread.*

Regular plural formation:

- Singular: *cat, dog*
- Plural: *cats, dogs*

Irregular plural formation:

- Singular: *man, woman, foot, fish, ox, knife, mouse*
- Plural: *men, women, feet, fish, oxen, knives, mice*

Possession can be expressed either by the possessive enclitic *-s* (also traditionally called a genitive suffix), or by the preposition *of*. Historically the -s possessive has been used for animate nouns, whereas the *of* possessive has been reserved for inanimate nouns.

Today this distinction is less clear, and many speakers use *-s* also with inanimates. Orthographically the possessive -s is

separated from a singular noun with an apostrophe. If the noun is plural formed with -s the apostrophe follows the -s.

Possessive constructions:

- With -s: *The woman's husband's child*
- With of: *The child of the husband of the woman*

Nouns can form noun phrases (NPs) where they are the syntactic head of the words that depend on them such as determiners, quantifiers, conjunctions or adjectives. Noun phrases can be short, such as *the man*, composed only of a determiner and a noun. They can also include modifiers such as adjectives (e.g. *red*, *tall*, *all*) and specifiers such as determiners (e.g. *the*, *that*). But they can also tie together several nouns into a single long NP, using conjunctions such as *and*, or prepositions such as *with*, e.g. *the tall man with the long red trousers and his skinny wife with the spectacles* (this NP uses conjunctions, prepositions, specifiers, and modifiers). Regardless of length, an NP functions as a syntactic unit. For example, the possessive enclitic can, in cases which do not lead to ambiguity, follow the entire noun phrase, as in *The President of India's wife*, where the enclitic follows *India* and not *President*.

The class of determiners is used to specify the noun they precede in terms of definiteness, where *the* marks a definite noun and *a* or *an* an indefinite one. A definite noun is assumed by the speaker to be already known by the interlocutor, whereas an indefinite noun is not specified as being previously known. Quantifiers, which include *one*, *many*, *some* and *all*, are used to specify the noun in terms of quantity or number. The noun must agree with the number of the determiner, e.g.

one man (sg.) but *all men* (pl.). Determiners are the first constituents in a noun phrase.

Adjectives

Adjectives modify a noun by providing additional information about their referents. In English, adjectives come before the nouns they modify and after determiners. In Modern English, adjectives are not inflected so as toagree in form with the noun they modify, as adjectives in most other Indo-European languages do. For example, in the phrases *the slender boy*, and *many slender girls*, the adjective *slender* does not change form to agree with either the number or gender of the noun.

Some adjectives are inflected for degree of comparison, with the positive degree unmarked, the suffix *-er* marking the comparative, and *-est* marking the superlative: *a small boy*, *the boy is smaller than the girl*, *that boy is the smallest*. Some adjectives have irregular comparative and superlative forms, such as *good*, *better*, and *best*. Other adjectives have comparatives formed by periphrastic constructions, with the adverb *more* marking the comparative, and *most* marking the superlative: *happier* or *more happy*, *the happiest* or *most happy*. There is some variation among speakers regarding which adjectives use inflected or periphrastic comparison, and some studies have shown a tendency for the periphrastic forms to become more common at the expense of the inflected form.

Pronouns, case, and person

English pronouns conserve many traits of case and gender inflection. The personal pronouns retain a difference between

subjective and objective case in most persons (*I/me, he/him, she/her, we/us, they/them*) as well as a gender and animateness distinction in the third person singular (distinguishing *he/she/it*). The subjective case corresponds to the Old English nominative case, and the objective case is used in the sense both of the previous accusative case (for a patient, or direct object of a transitive verb), and of the Old English dative case (for a recipient or indirect object of a transitive verb). The subjective is used when the pronoun is the subject of a finite clause, otherwise the objective is used. While grammarians such as Henry Sweet and Otto Jespersen noted that the English cases did not correspond to the traditional Latin-based system, some contemporary grammars, for example Huddleston & Pullum (2002), retain traditional labels for the cases, calling them nominative and accusative cases respectively.

Possessive pronouns exist in dependent and independent forms; the dependent form functions as a determiner specifying a noun (as in *my chair*), while the independent form can stand alone as if it were a noun (e.g. *the chair is mine*). The English system of grammatical person no longer has a distinction between formal and informal pronouns of address (the old 2nd person singular familiar pronoun *thou* acquired a pejorative or inferior tinge of meaning and was abandoned), and the forms for 2nd person plural and singular are identical except in the reflexive form. Some dialects have introduced innovative 2nd person plural pronouns such as *y'all* found in Southern American English and African American (Vernacular) English or *youse* found in Australian English and *ye* in Hiberno-English.

Pronouns are used to refer to entities deictically or anaphorically. A deictic pronoun points to some person or object by identifying it relative to the speech situation—for example, the pronoun *I* identifies the speaker, and the pronoun *you*, the addressee. Anaphoric pronouns such as *that* refer back to an entity already mentioned or assumed by the speaker to be known by the audience, for example in the sentence *I already told you that.* The reflexive pronouns are used when the oblique argument is identical to the subject of a phrase (e.g. "he sent it to himself" or "she braced herself for impact").

Prepositions

Prepositional phrases (PP) are phrases composed of a preposition and one or more nouns, e.g. *with the dog*, *for my friend*, *to school*, *in England.* Prepositions have a wide range of uses in English. They are used to describe movement, place, and other relations between different entities, but they also have many syntactic uses such as introducing complement clauses and oblique arguments of verbs.

For example, in the phrase *I gave it to him*, the preposition *to* marks the recipient, or Indirect Object of the verb *to give.* Traditionally words were only considered prepositions if they governed the case of the noun they preceded, for example causing the pronouns to use the objective rather than subjective form, "with her", "to me", "for us". But some contemporary grammars such as that of Huddleston & Pullum (2002:598–600) no longer consider government of case to be the defining feature of the class of prepositions, rather defining prepositions as words that can function as the heads of prepositional phrases.

Verbs and verb phrases

English verbs are inflected for tense and aspect and marked for agreement with present-tense third-person singular subject. Only the copula verb *to be* is still inflected for agreement with the plural and first and second person subjects. Auxiliary verbs such as *have* and *be*are paired with verbs in the infinitive, past, or progressive forms. They form complex tenses, aspects, and moods. Auxiliary verbs differ from other verbs in that they can be followed by the negation, and in that they can occur as the first constituent in a question sentence.

Most verbs have six inflectional forms. The primary forms are a plain present, a third-person singular present, and a preterite (past) form. The secondary forms are a plain form used for the infinitive, a gerund-participle and a past participle. The copula verb *to be* is the only verb to retain some of its original conjugation, and takes different inflectional forms depending on the subject. The first-person present-tense form is *am*, the third person singular form is *is*, and the form *are* is used in the second-person singular and all three plurals. The only verb past participle is *been* and its gerund-participle is *being*.

Further aspectual distinctions are shown by auxiliary verbs, primarily *have* and *be*, which show the contrast between a perfect and non-perfect past tense (*I have run* vs. *I was running*), and compound tenses such as preterite perfect (*I had been running*) and present perfect (*I have been running*).

For the expression of mood, English uses a number of modal auxiliaries, such as *can*, *may*, *will*, *shall* and the past tense forms *could*, *might*, *would*, *should*. There are also subjunctive

and imperative moods, both based on the plain form of the verb (i.e. without the third person singular *-s*), for use in subordinate clauses (e.g. subjunctive: *It is important that he run every day*; imperative *Run!*).

An infinitive form, that uses the plain form of the verb and the preposition *to*, is used for verbal clauses that are syntactically subordinate to a finite verbal clause. Finite verbal clauses are those that are formed around a verb in the present or preterite form. In clauses with auxiliary verbs, they are the finite verbs and the main verb is treated as a subordinate clause. For example, *he has to go* where only the auxiliary verb *have* is inflected for time and the main verb *to go* is in the infinitive, or in a complement clause such as *I saw him leave*, where the main verb is *to see* which is in a preterite form, and *leave* is in the infinitive.

Phrasal verbs

English also makes frequent use of constructions traditionally called phrasal verbs, verb phrases that are made up of a verb root and a preposition or particle which follows the verb. The phrase then functions as a single predicate. In terms of intonation the preposition is fused to the verb, but in writing it is written as a separate word. Examples of phrasal verbs are *to get up*, *to ask out*, *to back up*, *to give up*, *to get together*, *to hang out*, *to put up with*, etc. The phrasal verb frequently has a highly idiomatic meaning that is more specialised and restricted than what can be simply extrapolated from the combination of verb and preposition complement (e.g. *lay off* meaning *terminate someone's employment*). In spite of the idiomatic meaning, some grammarians, including Huddleston &

Pullum (2002:274), do not consider this type of construction to form a syntactic constituent and hence refrain from using the term "phrasal verb". Instead, they consider the construction simply to be a verb with a prepositional phrase as its syntactic complement, i.e. *he woke up in the morning* and *he ran up in the mountains* are syntactically equivalent.

Adverbs

The function of adverbs is to modify the action or event described by the verb by providing additional information about the manner in which it occurs. Many adverbs are derived from adjectives by appending the suffix *-ly*. For example, in the phrase *the woman walked quickly*, the adverb *quickly* is derived in this way from the adjective *quick*. Some commonly used adjectives have irregular adverbial forms, such as *good* which has the adverbial form *well*.

Syntax

Modern English syntax language is moderately analytic. It has developed features such as modal verbs and word order as resources for conveying meaning. Auxiliary verbs mark constructions such as questions, negative polarity, the passive voice and progressive aspect.

Basic constituent order

English word order has moved from the Germanic verb-second (V2) word order to being almost exclusively subject–verb–object (SVO). The combination of SVO order and use of auxiliary verbs

often creates clusters of two or more verbs at the centre of the sentence, such as *he had hoped to try to open it.*

In most sentences, English only marks grammatical relations through word order. The subject constituent precedes the verb and the object constituent follows it. The example below demonstrates how the grammatical roles of each constituent are marked only by the position relative to the verb:

• *The dog*	• *bites*	• *the man*
• S	• V	• O
• *The man*	• *bites*	• *the dog*
• S	• V	• O

An exception is found in sentences where one of the constituents is a pronoun, in which case it is doubly marked, both by word order and by case inflection, where the subject pronoun precedes the verb and takes the subjective case form, and the object pronoun follows the verb and takes the objective case form. The example below demonstrates this double marking in a sentence where both object and subject are represented with a third person singular masculine pronoun:

• *He*	• *hit*	• *him*
• S	• V	• O

Indirect objects (IO) of ditransitive verbs can be placed either as the first object in a double object construction (S V IO O), such as *I gave Jane the book* or in a prepositional phrase, such as *I gave the book to Jane.*

Clause syntax

In English a sentence may be composed of one or more clauses, that may, in turn, be composed of one or more phrases (e.g. Noun Phrases, Verb Phrases, and Prepositional Phrases). A clause is built around a verb and includes its constituents, such as any NPs and PPs. Within a sentence, there is always at least one main clause (or matrix clause) whereas other clauses are subordinate to a main clause. Subordinate clauses may function as arguments of the verb in the main clause. For example, in the phrase *I think (that) you are lying*, the main clause is headed by the verb *think*, the subject is *I*, but the object of the phrase is the subordinate clause *(that) you are lying*. The subordinating conjunction *that* shows that the clause that follows is a subordinate clause, but it is often omitted. Relative clauses are clauses that function as a modifier or specifier to some constituent in the main clause: For example, in the sentence *I saw the letter that you received today*, the relative clause *that you received today* specifies the meaning of the word *letter*, the object of the main clause. Relative clauses can be introduced by the pronouns *who*, *whose*, *whom* and *which* as well as by *that* (which can also be omitted.) In contrast to many other Germanic languages there is no major differences between word order in main and subordinate clauses.

Auxiliary verb constructions

English syntax relies on auxiliary verbs for many functions including the expression of tense, aspect, and mood. Auxiliary verbs form main clauses, and the main verbs function as heads of a subordinate clause of the auxiliary verb. For example, in

the sentence *the dog did not find its bone*, the clause *find its bone* is the complement of the negated verb *did not*. Subject–auxiliary inversionis used in many constructions, including focus, negation, and interrogative constructions.

The verb *do* can be used as an auxiliary even in simple declarative sentences, where it usually serves to add emphasis, as in "I *did* shut the fridge." However, in the negated and inverted clauses referred to above, it is used because the rules of English syntax permit these constructions only when an auxiliary is present. Modern English does not allow the addition of the negating adverb *not* to an ordinary finite lexical verb, as in **I know not*—it can only be added to an auxiliary (or copular) verb, hence if there is no other auxiliary present when negation is required, the auxiliary *do* is used, to produce a form like *I do not (don't) know.* The same applies in clauses requiring inversion, including most questions—inversion must involve the subject and an auxiliary verb, so it is not possible to say **Know you him?*; grammatical rules require *Do you know him?*

Negation is done with the adverb *not*, which precedes the main verb and follows an auxiliary verb. A contracted form of not - *n't*can be used as an enclitic attaching to auxiliary verbs and to the copula verb *to be*. Just as with questions, many negative constructions require the negation to occur with do-support, thus in Modern English *I don't know him* is the correct answer to the question *Do you know him?*, but not **I know him not*, although this construction may be found in older English.

Passive constructions also use auxiliary verbs. A passive construction rephrases an active construction in such a way

that the object of the active phrase becomes the subject of the passive phrase, and the subject of the active phrase is either omitted or demoted to a role as an oblique argument introduced in a prepositional phrase.

They are formed by using the past participle either with the auxiliary verb *to be* or *to get*, although not all varieties of English allow the use of passives with *get*. For example, putting the sentence *she sees him* into the passive becomes *he is seen (by her)*, or *he gets seen (by her)*.

Questions

Both yes–no questions and *wh*-questions in English are mostly formed using subject–auxiliary inversion (*Am I going tomorrow?*,*Where can we eat?*), which may require *do*-support (*Do you like her?*, *Where did he go?*).

In most cases, interrogative words (*wh*-words; e.g. *what*, *who*, *where*, *when*, *why*, *how*) appear in a fronted position. For example, in the question *What did you see?*, the word *what* appears as the first constituent despite being the grammatical object of the sentence. (When the *wh*-word is the subject or forms part of the subject, no inversion occurs: *Who saw the cat?*.) Prepositional phrases can also be fronted when they are the question's theme, e.g. *To whose house did you go last night?*.

The personal interrogative pronoun *who* is the only interrogative pronoun to still show inflection for case, with the variant *whom* serving as the objective case form, although this form may be going out of use in many contexts.

Discourse level syntax

While English is a subject-prominent language, at the discourse level it tends to use a topic-comment structure, where the known information (topic) precedes the new information (comment). Because of the strict SVO syntax, the topic of a sentence generally has to be the grammatical subject of the sentence. In cases where the topic is not the grammatical subject of the sentence, it is often promoted to subject position through syntactic means. One way of doing this is through a passive construction, *the girl was stung by the bee.* Another way is through a cleft sentence where the main clause is demoted to be a complement clause of a copula sentence with a dummy subject such as *it* or *there*, e.g. *it was the girl that the bee stung*, *there was a girl who was stung by a bee.* Dummy subjects are also used in constructions where there is no grammatical subject such as with impersonal verbs (e.g., *it is raining*) or in existential clauses (*there are many cars on the street*). Through the use of these complex sentence constructions with informationally vacuous subjects, English is able to maintain both a topic-comment sentence structure and a SVO syntax.

Focus constructions emphasise a particular piece of new or salient information within a sentence, generally through allocating the main sentence level stress on the focal constituent. For example, *the girl was stung by a bee* (emphasising it was a bee and not, for example, a wasp that stung her), or *The girl was stung by a bee* (contrasting with another possibility, for example that it was the boy). Topic and focus can also be established through syntactic dislocation, either preposing or postposing the item to be focused on

relative to the main clause. For example, *That girl over there, she was stung by a bee*, emphasises the girl by preposition, but a similar effect could be achieved by postposition, *she was stung by a bee, that girl over there*, where reference to the girl is established as an "afterthought".

Cohesion between sentences is achieved through the use of deictic pronouns as anaphora (e.g. *that is exactly what I mean* where *that* refers to some fact known to both interlocutors, or *then* used to locate the time of a narrated event relative to the time of a previously narrated event). Discourse markers such as *oh*, *so* or *well*, also signal the progression of ideas between sentences and help to create cohesion. Discourse markers are often the first constituents in sentences. Discourse markers are also used for stance taking in which speakers position themselves in a specific attitude towards what is being said, for example, *no way is that true!* (the idiomatic marker *no way!* expressing disbelief), or *boy! I'm hungry* (the marker *boy* expressing emphasis). While discourse markers are particularly characteristic of informal and spoken registers of English, they are also used in written and formal registers.

Vocabulary

It is generally stated that English has around 170,000 words, or 220,000 if obsolete words are counted; this estimate is based on the last full edition of the *Oxford English Dictionary* from 1989. Over half of these words are nouns, a quarter adjectives, and a seventh verbs. There is one count that puts the English vocabulary at about 1 million words—but that count presumably includes words such as Latin species names, scientific terminology, botanical terms, prefixed and suffixed

words, jargon, foreign words of extremely limited English use, and technical acronyms.

Due to its status as an international language, English adopts foreign words quickly, and borrows vocabulary from many other sources. Early studies of English vocabulary by lexicographers, the scholars who formally study vocabulary, compile dictionaries, or both, were impeded by a lack of comprehensive data on actual vocabulary in use from good-quality linguistic corpora, collections of actual written texts and spoken passages.

Many statements published before the end of the 20th century about the growth of English vocabulary over time, the dates of first use of various words in English, and the sources of English vocabulary will have to be corrected as new computerised analysis of linguistic corpus data becomes available.

Word formation processes

English forms new words from existing words or roots in its vocabulary through a variety of processes. One of the most productive processes in English is conversion, using a word with a different grammatical role, for example using a noun as a verb or a verb as a noun. Another productive word-formation process is nominal compounding, producing compound words such as *babysitter* or *ice cream* or *homesick*. A process more common in Old English than in Modern English, but still productive in Modern English, is the use of derivational suffixes (-*hood*, -*ness*, -*ing*, -*ility*) to derive new words from existing words (especially those of Germanic origin) or stems

(especially for words of Latin or Greek origin). Formation of new words, called neologisms, based on Greek and/or Latin roots (for example *television* or *optometry*) is a highly productive process in English and in most modern European languages, so much so that it is often difficult to determine in which language a neologism originated. For this reason, lexicographer Philip Gove attributed many such words to the "international scientific vocabulary" (ISV) when compiling *Webster's Third New International Dictionary* (1961). Another active word-formation process in English are acronyms, words formed by pronouncing as a single word abbreviations of longer phrases, e.g. *NATO*, *laser*).

Word origins

English, besides forming new words from existing words and their roots, also borrows words from other languages. This adoption of words from other languages is commonplace in many world languages, but English has been especially open to borrowing of foreign words throughout the last 1,000 years. The most commonly used words in English are West Germanic. The words in English learned first by children as they learn to speak, particularly the grammatical words that dominate the word count of both spoken and written texts, are mainly the Germanic words inherited from the earliest periods of the development of Old English.

But one of the consequences of long language contact between French and English in all stages of their development is that the vocabulary of English has a very high percentage of "Latinate" words (derived from French, especially, and also from other Romance languages and Latin). French words from

various periods of the development of French now make up one-third of the vocabulary of English. Linguist Anthony Lacoudre estimated that over 40,000 English words are of French origin and may be understood without orthographical change by French speakers. Words of Old Norse origin have entered the English language primarily from the contact between Old Norse and Old English during colonisation of eastern and northern England. Many of these words are part of English core vocabulary, such as *egg* and *knife.*

English has also borrowed many words directly from Latin, the ancestor of the Romance languages, during all stages of its development. Many of these words had earlier been borrowed into Latin from Greek. Latin or Greek are still highly productive sources of stems used to form vocabulary of subjects learned in higher education such as the sciences, philosophy, and mathematics. English continues to gain new loanwords and calques ("loan translations") from languages all over the world, and words from languages other than the ancestral Anglo-Saxon language make up about 60% of the vocabulary of English.

English has formal and informal speech registers; informal registers, including child-directed speech, tend to be made up predominantly of words of Anglo-Saxon origin, while the percentage of vocabulary that is of Latinate origin is higher in legal, scientific, and academic texts.

English loanwords and calques in other languages

English has had a strong influence on the vocabulary of other languages. The influence of English comes from such factors as

opinion leaders in other countries knowing the English language, the role of English as a world *lingua franca*, and the large number of books and films that are translated from English into other languages. That pervasive use of English leads to a conclusion in many places that English is an especially suitable language for expressing new ideas or describing new technologies. Among varieties of English, it is especially American English that influences other languages. Some languages, such as Chinese, write words borrowed from English mostly as calques, while others, such as Japanese, readily take in English loanwords written in sound-indicating script. Dubbed films and television programmes are an especially fruitful source of English influence on languages in Europe.

Writing system

Since the ninth century, English has been written in a Latin alphabet (also called Roman alphabet). Earlier Old English texts in Anglo-Saxon runes are only short inscriptions. The great majority of literary works in Old English that survive to today are written in the Roman alphabet. The modern English alphabet contains 26 letters of the Latin script: a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z (which also have capital forms: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z).

The spelling system, or orthography, of English is multi-layered and complex, with elements of French, Latin, and Greek spelling on top of the native Germanic system. Further complications have arisen through sound changes with which the orthography has not kept pace. Compared to European

languages for which official organisations have promoted spelling reforms, English has spelling that is a less consistent indicator of pronunciation, and standard spellings of words that are more difficult to guess from knowing how a word is pronounced. There are also systematic spelling differences between British and American English. These situations have prompted proposals for spelling reform in English.

Although letters and speech sounds do not have a one-to-one correspondence in standard English spelling, spelling rules that take into account syllable structure, phonetic changes in derived words, and word accent are reliable for most English words. Moreover, standard English spelling shows etymological relationships between related words that would be obscured by a closer correspondence between pronunciation and spelling, for example the words *photograph*, *photography*, and *photographic*, or the words *electricity* and *electrical*. While few scholars agree with Chomsky and Halle (1968) that conventional English orthography is "near-optimal", there is a rationale for current English spelling patterns. The standard orthography of English is the most widely used writing system in the world. Standard English spelling is based on a graphomorphemic segmentation of words into written clues of what meaningful units make up each word.

Readers of English can generally rely on the correspondence between spelling and pronunciation to be fairly regular for letters or digraphs used to spell consonant sounds. The letters *b*, *d*, *f*, *h*, *j*, *k*, *l*, *m*, *n*, *p*, *r*, *s*, *t*, *v*, *w*, *y*, *z* represent, respectively, the phonemes /b, d, f, h, dʒ, k, l, m, n, p, r, s, t, v, w, j, z/. The letters *c* and *g* normally represent /k/ and /ɡ/, but there is also a soft *c* pronounced /s/, and a soft *g*

pronounced /dʒ/. The differences in the pronunciations of the letters *c* and *g* are often signalled by the following letters in standard English spelling. Digraphs used to represent phonemes and phoneme sequences include *ch* for /tʃ/, *sh* for /ʃ/, *th* for /θ/ or /ð/, *ng* for /ŋ/, *qu* for /kw/, and *ph* for /f/ in Greek-derived words. The single letter *x* is generally pronounced as /z/ in word-initial position and as /ks/ otherwise. There are exceptions to these generalisations, often the result of loanwords being spelled according to the spelling patterns of their languages of origin or residues of proposals by scholars in the early period of Modern English to follow the spelling patterns of Latin for English words of Germanic origin.

For the vowel sounds of the English language, however, correspondences between spelling and pronunciation are more irregular. There are many more vowel phonemes in English than there are single vowel letters (*a*, *e*, *i*, *o*, *u*, *w*, *y*). As a result, some "long vowels" are often indicated by combinations of letters (like the *oa* in *boat*, the *ow* in *how*, and the *ay* in *stay*), or the historically based silent *e* (as in *note* and *cake*).

The consequence of this complex orthographic history is that learning to read and write can be challenging in English. It can take longer for school pupils to become independently fluent readers of English than of many other languages, including Italian, Spanish, and German. Nonetheless, there is an advantage for learners of English reading in learning the specific sound-symbol regularities that occur in the standard English spellings of commonly used words. Such instruction greatly reduces the risk of children experiencing reading difficulties in English. Making primary school teachers more aware of the primacy of morpheme representation in English

may help learners learn more efficiently to read and write English. English writing also includes a system of punctuation marks that is similar to those used in most alphabetic languages around the world. The purpose of punctuation is to mark meaningful grammatical relationships in sentences to aid readers in understanding a text and to indicate features important for reading a text aloud.

Dialects, accents, and varieties

Dialectologists identify many English dialects, which usually refer to regional varieties that differ from each other in terms of patterns of grammar, vocabulary, and pronunciation. The pronunciation of particular areas distinguishes dialects as separate regional accents. The major native dialects of English are often divided by linguists into the two extremely general categories of British English (BrE) and North American English (NAE). There also exists a third common major grouping of English varieties: Southern Hemisphere English, the most prominent being Australian and New Zealand English.

Britain and Ireland

Since the English language first evolved in Britain and Ireland, the archipelago is home to the most diverse dialects, particularly in England. Within the United Kingdom, the Received Pronunciation (RP), an educated dialect of South East England, is traditionally used as the broadcast standard and is considered the most prestigious of the British dialects. The spread of RP (also known as BBC English) through the media has caused many traditional dialects of rural England to

recede, as youths adopt the traits of the prestige variety instead of traits from local dialects. At the time of the Survey of English Dialects, grammar and vocabulary differed across the country, but a process of lexical attrition has led most of this variation to disappear.Nonetheless, this attrition has mostly affected dialectal variation in grammar and vocabulary, and in fact, only 3 percent of the English population actually speak RP, the remainder speaking in regional accents and dialects with varying degrees of RP influence. There is also variability within RP, particularly along class lines between Upper and Middle-class RP speakers and between native RP speakers and speakers who adopt RP later in life. Within Britain, there is also considerable variation along lines of social class, and some traits though exceedingly common are considered "non-standard" and are associated with lower class speakers and identities. An example of this is H-dropping, which was historically a feature of lower-class London English, particularly Cockney, and can now be heard in the local accents of most parts of England—yet it remains largely absent in broadcasting and among the upper crust of British society.

English in Englandcan be divided into four major dialect regions, Southwest English, South East English, Midlands English, and Northern English. Within each of these regions several local subdialects exist: Within the Northern region, there is a division between the Yorkshire dialects and the Geordie dialect spoken in Northumbria around Newcastle, and the Lancashire dialects with local urban dialects in Liverpool (Scouse) and Manchester (Mancunian). Having been the centre of Danish occupation during the Viking Invasions, Northern English dialects, particularly the Yorkshire dialect, retain Norse features not found in other English varieties.

Since the 15th century, southeastern England varieties have centred on London, which has been the centre from which dialectal innovations have spread to other dialects. In London, the Cockney dialect was traditionally used by the lower classes, and it was long a socially stigmatised variety. The spread of Cockney features across the south-east led the media to talk of Estuary English as a new dialect, but the notion was criticised by many linguists on the grounds that London had been influencing neighbouring regions throughout history. Traits that have spread from London in recent decades include the use of intrusive R (*drawing* is pronounced *drawring*/ˈdrɔːrɪŋ/), *t*-glottalisation (*Potter* is pronounced with a glottal stop as *Po'er*/poʔʌ/), and the pronunciation of *th-* as /f/ (*thanks* pronounced *fanks*) or /v/ (*bother* pronounced *bover*).

Scots is today considered a separate language from English, but it has its origins in early Northern Middle English and developed and changed during its history with influence from other sources, particularly Scots Gaelic and Old Norse. Scots itself has a number of regional dialects. And in addition to Scots, Scottish English comprises the varieties of Standard English spoken in Scotland; most varieties are Northern English accents, with some influence from Scots.

In Ireland, various forms of English have been spoken since the Norman invasions of the 11th century. In County Wexford, in the area surrounding Dublin, two extinct dialects known as Forth and Bargy and Fingallian developed as offshoots from Early Middle English, and were spoken until the 19th century. Modern Irish English, however, has its roots in English colonisation in the 17th century. Today Irish English is divided

into Ulster English, the Northern Ireland dialect with strong influence from Scots, and various dialects of the Republic of Ireland. Like Scottish and most North American accents, almost all Irish accents preserve the rhoticity which has been lost in the dialects influenced by RP.

North America

North American English is fairly homogeneous compared to British English. Today, American accent variation is often increasing at the regional level and decreasing at the very local level, though most Americans still speak within a phonological continuum of similar accents, known collectively as General American (GA), with differences hardly noticed even among Americans themselves (such as Midland and Western American English). In most American and Canadian English dialects, rhoticity (or *r*-fulness) is dominant, with non-rhoticity (*r*-dropping) becoming associated with lower prestige and social class especially after World War II; this contrasts with the situation in England, where non-rhoticity has become the standard.

Separate from GA are American dialects with clearly distinct sound systems, historically including Southern American English, English of the coastal Northeast (famously including Eastern New England English and New York City English), and African American Vernacular English, all of which are historically non-rhotic. Canadian English, except for the Atlantic provinces and perhaps Quebec, may be classified under GA as well, but it often shows the raising of the vowels/aɪ/ and /aʊ/ before voiceless consonants, as well as distinct norms for written and pronunciation standards.

In Southern American English, the most populous American "accent group" outside of GA, rhoticity now strongly prevails, replacing the region's historical non-rhotic prestige.

Southern accents are colloquially described as a "drawl" or "twang," being recognised most readily by the Southern Vowel Shift initiated by glide-deleting in the /aɪ/ vowel (e.g. pronouncing *spy* almost like *spa*), the "Southern breaking" of several front pure vowels into a gliding vowel or even two syllables (e.g. pronouncing the word "press" almost like "pray-us"), the pin–pen merger, and other distinctive phonological, grammatical, and lexical features, many of which are actually recent developments of the 19th century or later.

Today spoken primarily by working- and middle-class African Americans, African-American Vernacular English (AAVE) is also largely non-rhotic and likely originated among enslaved Africans and African Americans influenced primarily by the non-rhotic, non-standard older Southern dialects. A minority of linguists, contrarily, propose that AAVE mostly traces back to African languages spoken by the slaves who had to develop a pidgin or Creole English to communicate with slaves of other ethnic and linguistic origins.

AAVE's important commonalities with Southern accents suggests it developed into a highly coherent and homogeneous variety in the 19th or early 20th century. AAVE is commonly stigmatised in North America as a form of "broken" or "uneducated" English, as are white Southern accents, but linguists today recognise both as fully developed varieties of English with their own norms shared by a large speech community.

Australia and New Zealand

Since 1788, English has been spoken in Oceania, and Australian English has developed as a first language of the vast majority of the inhabitants of the Australian continent, its standard accent being General Australian. The English of neighbouring New Zealandhas to a lesser degree become an influential standard variety of the language. Australian and New Zealand English are each other's closest relatives with few differentiating characteristics, followed by South African English and the English of southeastern England, all of which have similarly non-rhotic accents, aside from some accents in the South Island of New Zealand. Australian and New Zealand English stand out for their innovative vowels: many short vowels are fronted or raised, whereas many long vowels have diphthongised. Australian English also has a contrast between long and short vowels, not found in most other varieties. Australian English grammar aligns closely to British and American English; like American English, collective plural subjects take on a singular verb (as in *the government is* rather than *are*). New Zealand English uses front vowels that are often even higher than in Australian English.

Singapore

The development of Singapore English started from at least 1819 when British statesmanStamford Raffles arrived in the lands that now make up Singapore to establish a trading port. It generally resembles British English and is often used in more formal settings such as the workplace or when communicating with people of authority such as employers, teachers, and government officials. Singapore English acts as

the "bridge" among different ethnic groups in Singapore, and in addition to being one of the four official languages in the country, it is considered de facto as the main language of communication. Standard Singapore English retains British spelling and grammar.

The standard Singaporean accent used to be officially Received Pronunciation (RP), prevalent during news broadcasts and in radio. However, a standard Singaporean accent, quite independent of any external standard, including RP, started to emerge. A 2003 study by the National Institute of Education in Singapore suggests that a standard Singaporean pronunciation is emerging and is on the cusp of being standardised. Singaporean accents can also be said to be largely non-rhotic.

In addition to Singapore English, Singlish is an English-based creole language spoken in Singapore. Unlike SSE, Singlish includes many discourse particles and loan words from various Asian languages such as Malay, Japanese, Mandarin and Hokkien. Although it is controversially regarded as "low prestige" especially by the government, most Singaporeans view Singlish as a unique Singaporean identity and continues to be used in informal communication among Singaporeans, and for new citizens, immigrants or tourists to learn more about Singaporean culture.

Philippines

The first significant exposure of the Philippines to the English language occurred in 1762 when the British occupied Manila during the Seven Years' War, but this was a brief episode that had no lasting influence. English later became more important

and widespread during American rule between 1898 and 1946, and remains an official language of the Philippines. Today, the use of English is ubiquitous in the Philippines, from street signs and marquees, government documents and forms, courtrooms, the media and entertainment industries, the business sector, and other aspects of daily life. One such usage that is also prominent in the country is in speech, where most Filipinos from Manila would use or have been exposed to Taglish, a form of code-switching between Tagalog and English. A similar code-switching method is used by urban native speakers of Visayan languages called Bislish.

Africa, the Caribbean, and South Asia

English is spoken widely in southern Africa and is an official or co-official language in several countries. In South Africa, English has been spoken since 1820, co-existing with Afrikaans and various African languages such as the Khoe and Bantu languages. Today, about 9 percent of the South African population speaks South African English (SAE) as a first language. SAE is a non-rhotic variety, which tends to follow RP as a norm. It is alone among non-rhotic varieties in lacking intrusive r. There are different L2 varieties that differ based on the native language of the speakers. Most phonological differences from RP are in the vowels. Consonant differences include the tendency to pronounce /p, t, tʃ, k/ without aspiration (e.g. *pin* pronounced [pɪn] rather than as [pʰɪn] as in most other varieties), while r is often pronounced as a flap [ɾ] instead of as the more common fricative.

Nigerian English is a dialect of English spoken in Nigeria. It is based on British English, but in recent years, because of

influence from the United States, some words of American English origin have made it into Nigerian English. Additionally, some new words and collocations have emerged from the language, which come from the need to express concepts specific to the culture of the nation (e.g. *senior wife*). Over 150 million Nigerians speak English.

Several varieties of English are also spoken in the Caribbean islands that were colonial possessions of Britain, including Jamaica, and the Leeward and Windward Islands and Trinidad and Tobago, Barbados, the Cayman Islands, and Belize. Each of these areas is home both to a local variety of English and a local English-based creole, combining English and African languages. The most prominent varieties are Jamaican English and Jamaican Creole. In Central America, English-based creoles are spoken in on the Caribbean coasts of Nicaragua and Panama. Locals are often fluent both in the local English variety and the local creole languages and code-switching between them is frequent, indeed another way to conceptualise the relationship between Creole and Standard varieties is to see a spectrum of social registers with the Creole forms serving as "basilect" and the more RP-like forms serving as the "acrolect", the most formal register.

Most Caribbean varieties are based on British English and consequently, most are non-rhotic, except for formal styles of Jamaican English which are often rhotic. Jamaican English differs from RP in its vowel inventory, which has a distinction between long and short vowels rather than tense and lax vowels as in Standard English. The diphthongs /ei/ and /ou/ are monophthongs [eː] and [oː] or even the reverse diphthongs [ie] and [uo] (e.g. *bay* and *boat* pronounced [bʲeː] and [bʷoːt]).

Often word-final consonant clusters are simplified so that "child" is pronounced [tʃail] and "wind" [win].

As a historical legacy, Indian English tends to take RP as its ideal, and how well this ideal is realised in an individual's speech reflects class distinctions among Indian English speakers. Indian English accents are marked by the pronunciation of phonemes such as /t/ and /d/ (often pronounced with retroflex articulation as [ʈ] and [ɖ]) and the replacement of /θ/ and /ð/ with dentals [t̪] and [d̪]. Sometimes Indian English speakers may also use spelling based pronunciations where the silent ⟨h⟩ found in words such as *ghost*is pronounced as an Indian voiced aspirated stop [ɡʱ].

Chapter 2
English Grammar

English grammar is the way in which meanings are encoded into wordings in the English language. This includes the structure of words, phrases, clauses, sentences, and whole texts.

This article describes a generalized, present-day Standard English – a form of speech and writing used in public discourse, including broadcasting, education, entertainment, government, and news, over a range of registers, from formal to informal. Divergences from the grammar described here occur in some historical, social, cultural and regional varieties of English, although these are more minor than differences in pronunciation and vocabulary.

Modern English has largely abandoned the inflectionalcase system of Indo-European in favor of analytic constructions. The personal pronouns retain morphological case more strongly than any other word class (a remnant of the more extensive Germanic case system of Old English). For other pronouns, and all nouns, adjectives, and articles, grammatical function is indicated only by word order, by prepositions, and by the "Saxon genitive or English possessive" (-*'s*).

Eight "word classes" or "parts of speech"are commonly distinguished in English: nouns, determiners, pronouns, verbs, adjectives, adverbs, prepositions, and conjunctions. Nouns form the largest word class, and verbs the second-largest.

Unlike nouns in almost all other Indo-European languages, English nouns do not have grammatical gender.

Word classes and phrases

Nouns, verbs, adjectives, and adverbs form open classes – word classes that readily accept new members, such as the noun *celebutante* (a celebrity who frequents the fashion circles), and other similar relatively new words. The others are considered to be closed classes. For example, it is rare for a new pronoun to enter the language. Determiners, traditionally classified along with adjectives, have not always been regarded as a separate part of speech. Interjections are another word class, but these are not described here as they do not form part of the clause and sentence structure of the language.

Linguists generally accept nine English word classes: nouns, verbs, adjectives, adverbs, pronouns, prepositions, conjunctions, determiners, and exclamations. English words are not generally marked for word class. It is not usually possible to tell from the form of a word which class it belongs to except, to some extent, in the case of words with inflectional endings or derivational suffixes. On the other hand, most words belong to more than one-word class. For example, *run* can serve as either a verb or a noun (these are regarded as two different lexemes). Lexemes may be inflected to express different grammatical categories. The lexeme *run* has the forms *runs*, *ran*, *runny*, *runner*, and *running*. Words in one class can sometimes be derived from those in another. This has the potential to give rise to new words. The noun *aerobics*has recently given rise to the adjective *aerobicized*.

Words combine to form phrases. A phrase typically serves the same function as a word from some particular word class. For example, *my very good friend Peter* is a phrase that can be used in a sentence as if it were a noun, and is therefore called a noun phrase. Similarly, adjectival phrases and adverbial phrases function as if they were adjectives or adverbs, but with other types of phrases, the terminology has different implications. For example, a verb phrase consists of a verb together with any objects and other dependents; a prepositional phrase consists of a preposition and its complement (and is therefore usually a type of adverbial phrase); and a determiner phrase is a type of noun phrase containing a determiner.

Nouns

Many common suffixes form nouns from other nouns or from other types of words, such as *-age* (as in *shrinkage*), *-hood* (as in *sisterhood*), and so on, although many nouns are base forms not containing any such suffix (such as *cat*, *grass*, *France*). Nouns are also often created by conversion of verbs or adjectives, as with the words *talk* and *reading* (*a boring talk*, *the assigned reading*).

Nouns are sometimes classified semantically (by their meanings) as proper nouns and common nouns (*Cyrus*, *China* vs. *frog*, *milk*) or as concrete nouns and abstract nouns (*book*, *laptop* vs. *embarrassment*, *prejudice*). A grammatical distinction is often made between count (countable) nouns such as *clock* and *city*, and non-count (uncountable) nouns such as *milk* and *decor*. Some nouns can function both as

countable and as uncountable such as the word "wine" (*This is a good wine*, *I prefer red wine*).

Countable nouns generally have singular and plural forms. In most cases the plural is formed from the singular by adding -*[e]s* (as in *dogs*, *bushes*), although there are also irregular forms (*woman/women*, *foot/feet*, etc.), including cases where the two forms are identical (*sheep*, *series*). For more details, see English plural. Certain nouns can be used with plural verbs even though they are singular in form, as in *The government were …* (where *the government* is considered to refer to the people constituting the government). This is a form of synesis; it is more common in British than American English. See English plural § Singulars with collective meaning treated as plural.

English nouns are not marked for case as they are in some languages, but they have possessive forms, through the addition of -*'s* (as in *John's*, *children's*) or just an apostrophe (with no change in pronunciation) in the case of -*[e]s* plurals and sometimes other words ending with -*s* (*the dogs' owners*, *Jesus' love*). More generally, the ending can be applied to noun phrases (as in *the man you saw yesterday's sister*); see below. The possessive form can be used either as a determiner (*John's cat*) or as a noun phrase (*John's is the one next to Jane's*).

The status of the possessive as an affix or a clitic is the subject of debate. It differs from the noun inflection of languages such as German, in that the genitive ending may attach to the last word of the phrase. To account for this, the possessive can be analysed, for instance as a clitic

construction (an "encliticpostposition") or as an inflection of the last word of a phrase ("edge inflection").

Phrases

Noun phrases are phrases that function grammatically as nouns within sentences, for example as the subject or object of a verb. Most noun phrases have a noun as their head.

In this structure:

- the*determiner* may be an article (*the*, *a[n]*) or other equivalent word, as described in the following section. In many contexts, it is required for a noun phrase to include some determiner.
- *pre-modifiers* include adjectives and some adjective phrases (such as *red*, *really lovely*), and noun adjuncts (such as *college* in the phrase *the college student*). Adjectival modifiers usually come before noun adjuncts.
- a *complement* or *postmodifier* may be a prepositional phrase (*... of London*), a relative clause (like *...which we saw yesterday*), certain adjective or participial phrases (*... sitting on the beach*), or a dependent clause or infinitive phrase appropriate to the noun (like *... that the world is round* after a noun such as *fact* or *statement*, or *... to travel widely* after a noun such as *desire*).

An example of a noun phrase that includes all of the above-mentioned elements is *that rather attractive young college student to whom you were talking*. Here *that* is the determiner, *rather attractive* and *young* are adjectival pre-modifiers, *college*

is a noun adjunct, *student* is the noun serving as the head of the phrase, and *to whom you were talking* is a post-modifier (a relative clause in this case). Notice the order of the pre-modifiers; the determiner *that* must come first and the noun adjunct *college* must come after the adjectival modifiers.

Coordinating conjunctions such as *and*, *or*, and *but* can be used at various levels in noun phrases, as in *John, Paul, and Mary*; *the matching green coat and hat*; *a dangerous but exciting ride*; *a person sitting down or standing up*. See § Conjunctions below for more explanation.

Noun phrases can also be placed in *apposition* (where two consecutive phrases refer to the same thing), as in *that president, Abraham Lincoln, ...* (where *that president* and *Abraham Lincoln* are in apposition). In some contexts, the same can be expressed by a prepositional phrase, as in *the twin curses of famine and pestilence* (meaning "the twin curses" that are "famine and pestilence").

Particular forms of noun phrases include:

- phrases formed by the determiner *the* with an adjective, as in *the homeless*, *the English* (these are plural phrases referring to homeless people or English people in general);
- phrases with a pronoun rather than a noun as the head (see below);
- phrases consisting just of a possessive;
- infinitive and gerund phrases, in certain positions;
- certain clauses, such as *that* clauses and relative clauses like *what he said*, in certain positions.

Gender

A system of grammatical gender, whereby every nounwas treated as either masculine, feminine or neuter, existed in Old English, but fell out of use during the Middle English period. Modern English retains features relating to natural gender, namely the use of certain nouns and pronouns (such as *he* and *she*) to refer specifically to persons or animals of one or other genders and certain others (such as *it*) for sexless objects – although feminine pronouns are sometimes used when referring to ships (and more uncommonly some airplanes and analogous machinery) and nation-states.

Some aspects of gender usage in English have been influenced by the movement towards a preference for gender-neutral language. Animals are triple-gender nouns, being able to take masculine, feminine and neuter pronouns. Generally there is no difference between male and female in English nouns. However, gender is occasionally exposed by different shapes or dissimilar words when referring to people or animals.

Many nouns that mention people's roles and jobs can refer to either a masculine or a feminine subject, for instance "cousin", "teenager", "teacher", "doctor", "student", "friend", and "colleague".

- Jane is my friend. She is a dentist.
- Paul is my cousin. He is a dentist.

Often the gender distinction for these neutral nouns is established by inserting the words "male" or "female".

- Sam is a female doctor.

- No, he is not my boyfriend; he is just a male friend.
- I have three female cousins and two male cousins.

Rarely, nouns illustrating things with no gender are referred to with a gendered pronoun to convey familiarity. It is also standard to use the gender-neutral pronoun (it).

- I love my car. She (the car) is my greatest passion.
- France is popular with her (France's) neighbors at the moment.
- I traveled from England to New York on the *Queen Elizabeth*; she (Queen Elizabeth) is a great ship.

Determiners

English determiners constitute a relatively small class of words. They include the articles *the* and *a[n]*; certain demonstrative and interrogative words such as *this*, *that*, and *which*; possessives such as *my* and *whose* (the role of determiner can also be played by noun possessive forms such as *John's* and *the girl's*); various quantifying words like *all*, *some*, *many*, *various*; and numerals (*one*, *two*, etc.).

There are also many phrases (such as *a couple of*) that can play the role of determiners.

Determiners are used in the formation of noun phrases (see above). Many words that serve as determiners can also be used as pronouns (*this*, *that*, *many*, etc.).

Determiners can be used in certain combinations, such as **all the** water and **the many** problems.

In many contexts, it is required for a noun phrase to be completed with an article or some other determiner. It is not grammatical to say just *cat sat on table*; one must say *my cat sat on the table.*

The most common situations in which a complete noun phrase can be formed without a determiner are when it refers generally to a whole class or concept (as in *dogs are dangerous* and *beauty is subjective*) and when it is a name (*Jane*, *Spain*, etc.). This is discussed in more detail at English articles and Zero article in English.

Pronouns

Personal

The personal pronouns of modern standard English are presented in the table above. They are *I, you, she, he, it, we*, and *they*. The personal pronouns are so-called not because they apply to persons (which other pronouns also do), but because they participate in the system of grammatical person (1st, 2nd, 3rd).

The second-person forms such as *you* are used with both singular and plural reference. In the Southern United States, y'all (you all) is used as a plural form, and various other phrases such as *you guys* are used in other places. An archaic set of second-person pronouns used for singular reference is *thou, thee, thyself, thy, thine,* which are still used in religious services and can be seen in older works, such as Shakespeare's—in such texts, the *you* set of pronouns are used for plural reference, or with singular reference as a formal V-form. *You* can also be used as an indefinite pronoun, referring

to a person in general (see generic *you*), compared to the more formal alternative, *one* (reflexive *oneself*, possessive *one's*).

The third-person singular forms are differentiated according to the sex of the referent. For example, *she* is used to refer to a female person, sometimes a female animal, and sometimes an object to which female characteristics are attributed, such as a ship or a country. A male person, and sometimes a male animal, is referred to using *he*. In other cases *it* can be used. (See Gender in English.) The word *it* can also be used as a dummy subject, in sentences like *It is going to be sunny this afternoon.*

The third-person form *they* is used with both plural and singular referents. Historically, singular *they* was restricted to quantificational constructions such as *Each employee should clean their desk* and referential cases where the referent's gender was unknown. However, it is increasingly used when the referent's gender is irrelevant or when the referent is neither male nor female.

The possessive determiners such as *my* are used as determiners together with nouns, as in *my old man*, *some of his friends*. The second possessive forms like *mine* are used when they do not qualify a noun: as pronouns, as in *mine is bigger than yours*, and as predicates, as in *this one is mine*. Note also the construction *a friend of mine* (meaning "someone who is my friend"). See English possessive for more details.

Demonstrative

The demonstrative pronouns of English are *this* (plural *these*), and *that* (plural *those*), as in *these are good, I like that*. Note

that all four words can also be used as determiners (followed by a noun), as in *those cars*. They can also form the alternative pronominal expressions *this/that one*, *these/those ones*.

Interrogative

The interrogative pronouns are *who*, *what*, and *which* (all of them can take the suffix *-ever* for emphasis). The pronoun *who* refers to a person or people; it has an oblique form *whom* (though in informal contexts this is usually replaced by *who*), and a possessive form (pronoun or determiner) *whose*. The pronoun *what* refers to things or abstracts.

The word *which* is used to ask about alternatives from what is seen as a closed set: *which (of the books) do you like best?* (It can also be an interrogative determiner: *which book?*; this can form the alternative pronominal expressions *which one* and *which ones*.) *Which*, *who*, and *what* can be either singular or plural, although *who* and *what* often take a singular verb regardless of any supposed number. For more information see *who*.

In Old and Middle English, the roles of the three words were different from their roles today. "The interrogative pronoun *hwā* 'who, what' had only singular forms and also only distinguished between non-neuter and neuter, the neuter nominative form being *hwæt*." Note that neuter and non-neuter refers to the grammatical gender system of the time, rather than the so-called natural gender system of today. A small holdover of this is the ability of relative (but not interrogative) *whose* to refer to non-persons (e.g., *the car whose door won't open*).

All the interrogative pronouns can also be used as relative pronouns, though *what* is quite limited in its use; see below for more details.

Relative

The main relative pronouns in English are *who* (with its derived forms *whom* and *whose*), *which*, and *that.*

The relative pronoun *which* refers to things rather than persons, as in *the shirt, which used to be red, is faded.* For persons, *who* is used (*the man who saw me was tall*). The oblique case form of *who* is *whom*, as in *the man whom I saw was tall*, although in informal registers*who* is commonly used in place of *whom.*

The possessive form of *who* is *whose* (for example, *the man whose car is missing*); however the use of *whose* is not restricted to persons (one can say *an idea whose time has come*).

The word *that* as a relative pronoun is normally found only in restrictive relative clauses (unlike *which* and *who*, which can be used in both restrictive and unrestrictive clauses). It can refer to either persons or things, and cannot follow a preposition. For example, one can say *the song that* [or *which*] *I listened to yesterday*, but *the song to which* [not *to that*] *I listened yesterday.* The relative pronoun *that* is usually pronounced with a reduced vowel (schwa), and hence differently from the demonstrative *that* (see Weak and strong forms in English). If *that* is not the subject of the relative clause, it can be omitted (*the song I listened to yesterday*).

The word *what*can be used to form a free relative clause – one that has no antecedent and that serves as a complete noun phrase in itself, as in *I like what he likes*. The words *whatever* and *whichever*can be used similarly, in the role of either pronouns (*whatever he likes*) or determiners (*whatever book he likes*). When referring to persons, *who(ever)* (and *whom(ever)*) can be used in a similar way (but not as determiners).

"There"

The word *there*is used as a pronoun in some sentences, playing the role of a dummy subject, normally of an intransitive verb. The "logical subject" of the verb then appears as a complement after the verb.

This use of *there* occurs most commonly with forms of the verb *be* in existential clauses, to refer to the presence or existence of something. For example: *There is a heaven*; *There are two cups on the table*; *There have been a lot of problems lately*. It can also be used with other verbs: *There exist two major variants*; *There occurred a very strange incident.*

The dummy subject takes the number (singular or plural) of the logical subject (complement), hence it takes a plural verb if the complement is plural. In informal English, however, the contraction*there's* is often used for both singular and plural.

The dummy subject can undergo inversion, *Is there a test today?*and*Never has there been a man such as this.* It can also appear without a corresponding logical subject, in short sentences and question tags: *There wasn't a discussion, was there? There was.*

The word *there* in such sentences has sometimes been analyzed as an adverb, or as a dummy predicate, rather than as a pronoun. However, its identification as a pronoun is most consistent with its behavior in inverted sentences and question tags as described above.

Because the word *there* can also be a deictic adverb (meaning "at/to that place"), a sentence like *There is a river* could have either of two meanings: "a river exists" (with *there* as a pronoun), and "a river is in that place" (with *there* as an adverb). In speech, the adverbial *there* would be given stress, while the pronoun would not – in fact, the pronoun is often pronounced as a weak form, /ðə(r)/.

Reciprocal

The English reciprocal pronouns are *each other* and *one another*. Although they are written with a space, they're best thought of as single words. No consistent distinction in meaning or use can be found between them. Like the reflexive pronouns, their use is limited to contexts where an antecedent precedes it. In the case case of the reciprocals, they need to appear in the same clause as the antecedent.

Other

Other pronouns in English are often identical in form to determiners (especially quantifiers), such as *many*, *a little*, etc. Sometimes, the pronoun form is different, as with *none* (corresponding to the determiner *no*), *nothing*, *everyone*, *somebody*, etc. Many examples are listed as indefinite pronouns. Another indefinite (or impersonal) pronoun is *one*

(with its reflexive form *oneself* and possessive *one's*), which is a more formal alternative to generic *you.*

Verbs

The basic form of an English verb is not generally marked by any ending, although there are certain suffixes that are frequently used to form verbs, such as *-ate* (*formulate*), *-fy* (*electrify*), and *-ise/ize* (*realise/realize*). Many verbs also contain prefixes, such as *un-* (*unmask*), *out-* (*outlast*), *over-* (*overtake*), and *under-* (*undervalue*). Verbs can also be formed from nouns and adjectives by zero derivation, as with the verbs *snare*, *nose*, *dry*, and *calm.*

Most verbs have three or four inflected forms in addition to the base form: a third-person singular present tense form in *-(e)s* (*writes*, *botches*), a present participle and gerund form in *-ing* (*writing*), a past tense (*wrote*), and – though often identical to the past tense form – a past participle (*written*). Regular verbs have identical past tense and past participle forms in *-ed*, but there are 100 or so irregular English verbs with different forms (see list). The verbs *have*, *do* and *say* also have irregular third-person present tense forms (*has*, *does*/dʌz/, *says*/sɛz/). The verb *be* has the largest number of irregular forms (*am, is, are* in the present tense, *was, were* in the past tense, *been* for the past participle).

Most of what are often referred to as verb tenses (or sometimes aspects) in English are formed using auxiliary verbs. Apart from what are called the simple present (*write*, *writes*) and simple past (*wrote*), there are also continuous (progressive) forms (*am/is/are/was/were writing*), perfect forms

(*have/has/had written*, and the perfect continuous *have/has/had been writing*), future forms (*will write*, *will be writing*, *will have written*, *will have been writing*), and conditionals (also called "future in the past") with *would* in place of *will*. The auxiliaries *shall* and *should* sometimes replace *will* and *would* in the first person. For the uses of these various verb forms, see English verbs and English clause syntax.

The basic form of the verb (*be, write, play*) is used as the infinitive, although there is also a "to-infinitive" (*to be*, *to write*, *to play*) used in many syntactical constructions.

There are also infinitives corresponding to other aspects: *(to) have written*, *(to) be writing*, *(to) have been writing*. The second-person imperative is identical to the (basic) infinitive; other imperative forms may be made with *let* (*let us go*, or *let's go*; *let them eat cake*).

A form identical to the infinitive can be used as a present subjunctive in certain contexts: *It is important that he **follow** them* or *... that he **be** committed to the cause*. There is also a past subjunctive (distinct from the simple past only in the possible use of *were* instead of *was*), used in some conditional sentences and similar: *if I were* (or *was*) *rich ...*; *were he to arrive now ...*; *I wish she were* (or *was*) *here*. For details see English subjunctive.

The passive voiceis formed using the verb *be* (in the appropriate tense or form) with the past participle of the verb in question: *cars are driven, he was killed, I am being tickled, it is nice to be pampered*, etc. The performer of the action may be

introduced in a prepositional phrase with *by* (as in *they were killed by the invaders*).

The English modal verbs consist of the core modals *can*, *could*, *may*, *might*, *must*, *shall*, *should*, *will*, *would*, as well as *ought (to)*, *had better*, and in some uses *dare* and *need*. These do not inflect for person or number, do not occur alone, and do not have infinitive or participle forms (except synonyms, as with *be/being/been able (to)* for the modals *can/could*). The modals are used with the basic infinitive form of a verb (*I can swim, he may be killed*, *we dare not move*, *need they go?*), except for *ought*, which takes *to* (*you ought to go*). Modals can indicate the condition, probability, possibility, necessity, obligation and ability exposed by the speaker's or writer's attitude or expression.

The copula*be*, along with the modal verbs and the other auxiliaries, form a distinct class, sometimes called "special verbs" or simply "auxiliaries". These have different syntax from ordinary lexical verbs, especially in that they make their interrogative forms by plain inversion with the subject, and their negative forms by adding *not* after the verb (*could I ...? I could not ...*). Apart from those already mentioned, this class may also include *used to* (although the forms *did he use to?* and *he didn't use to* are also found), and sometimes *have* even when not an auxiliary (forms like *have you a sister?* and *he hadn't a clue* are possible, though becoming less common). It also includes the auxiliary *do* (*does*, *did*); this is used with the basic infinitive of other verbs (those not belonging to the "special verbs" class) to make their question and negation forms, as well as emphatic forms (*do I like you?*;*he doesn't*

speak English; *we did close the fridge*). For more details of this, see *do*-support.

Some forms of the copula and auxiliaries often appear as contractions, as in *I'm* for *I am*, *you'd* for *you would* or *you had*, and *John's* for *John is*. Their negated forms with following *not* are also often contracted (see § Negation below). For detail see English auxiliaries and contractions.

Phrases

A verb together with its dependents, excluding its subject, may be identified as a verb phrase (although this concept is not acknowledged in all theories of grammar). A verb phrase headed by a finite verbmay also be called a predicate. The dependents may be objects, complements, and modifiers (adverbs or adverbial phrases). In English, objects and complements nearly always come after the verb; a direct object precedes other complements such as prepositional phrases, but if there is an indirect object as well, expressed without a preposition, then that precedes the direct object: *give me the book*, but *give the book to me.* Adverbial modifiers generally follow objects, although other positions are possible (see under § Adverbs below). Certain verb–modifier combinations, particularly when they have independent meaning (such as *take on* and *get up*), are known as "phrasal verbs".

For details of possible patterns, see English clause syntax. See the Non-finite clauses section of that article for verb phrases headed by non-finite verb forms, such as infinitives and participles.

Adjectives

English adjectives, as with other word classes, cannot in general be identified as such by their form, although many of them are formed from nouns or other words by the addition of a suffix, such as *-al* (*habitual*), *-ful* (*blissful*), *-ic* (*atomic*), *-ish* (*impish*, *youngish*), *-ous* (*hazardous*), etc.; or from other adjectives using a prefix: *disloyal*, *irredeemable*, *unforeseen*, *overtired.*

Adjectives may be used attributively, as part of a noun phrase (nearly always preceding the noun they modify; for exceptions see postpositive adjective), as in *the big house*, or predicatively, as in *the house is big.* Certain adjectives are restricted to one or other use; for example, *drunken* is attributive (*a drunken sailor*), while *drunk* is usually predicative (*the sailor was drunk*).

Comparison

Many adjectives have comparative and superlative forms in *-er* and *-est*, such as *faster* and *fastest* (from the positive form *fast*). Spelling rules which maintain pronunciation apply to suffixing adjectives just as they do for similar treatment of regular past tense formation; these cover consonant doubling (as in *bigger* and *biggest*, from *big*) and the change of *y* to *i* after consonants (as in *happier* and *happiest*, from *happy*).

The adjectives *good* and *bad* have the irregular forms *better, best* and *worse, worst*; also *far* becomes *farther, farthest* or *further, furthest.* The adjective *old* (for which the regular *older* and *oldest* are usual) also has the irregular forms *elder* and

eldest, these generally being restricted to use in comparing siblings and in certain independent uses. For the comparison of adverbs, see Adverbs below.

Many adjectives, however, particularly those that are longer and less common, do not have inflected comparative and superlative forms. Instead, they can be qualified with *more* and *most*, as in *beautiful, more beautiful, most beautiful* (this construction is also sometimes used even for adjectives for which inflected forms do exist).

Certain adjectives are classed as ungradable. These represent properties that cannot be compared on a scale; they simply apply or do not, as with *pregnant*, *dead*, *unique*. Consequently, comparative and superlative forms of such adjectives are not normally used, except in a figurative, humorous or imprecise context. Similarly, such adjectives are not normally qualified with modifiers of degree such as *very* and *fairly*, although with some of them it is idiomatic to use adverbs such as *completely*. Another type of adjective sometimes considered ungradable is those that represent an extreme degree of some property, such as *delicious* and *terrified*.

Phrases

An adjective phrase is a group of words that plays the role of an adjective in a sentence. It usually has a single adjective as its head, to which modifiers and complementsmay be added.

Adjectives can be modified by a preceding adverb or adverb phrase, as in *very warm*, *truly imposing*, *more than a little excited.* Some can also be preceded by a noun or quantitative phrase, as in *fat-free*, *two-meter-long.*

Complements following the adjective may include:

- prepositional phrases: *proud of him*, *angry at the screen*, *keen on breeding toads*;
- infinitive phrases: *anxious to solve the problem*, *easy to pick up*;
- content clauses, i.e. *that* clauses and certain others: *certain that he was right*, *unsure where they are*;
- after comparatives, phrases or clauses with *than*: *better than you*, *smaller than I had imagined*.

An adjective phrase may include both modifiers before the adjective and a complement after it, as in *very difficult to put away*.

Adjective phrases containing complements after the adjective cannot normally be used as attributive adjectives *before* a noun.

Sometimes they are used attributively after the noun, as in *a woman proud of being a midwife* (where they may be converted into relative clauses: *a woman who is proud of being a midwife*), but it is wrong to say **a proud of being a midwife woman*. Exceptions include very brief and often established phrases such as *easy-to-use*. (Certain complements can be moved to after the noun, leaving the adjective before the noun, as in *a better man than you*, *a hard nut to crack*.)

Certain attributive adjective phrases are formed from other parts of speech, without any adjective as their head, as in *a two-bedroom house*, *a no-jeans policy*.

Adverbs

Adverbs perform a wide range of functions. They typically modify verbs (or verb phrases), adjectives (or adjectival phrases), or other adverbs (or adverbial phrases). However, adverbs also sometimes qualify noun phrases (**only** *the boss*; **quite** *a lovely place*), pronouns and determiners (**almost** *all*), prepositional phrases (**halfway** *through the movie*), or whole sentences, to provide contextual comment or indicate an attitude (**Frankly**, *I don't believe you*). They can also indicate a relationship between clauses or sentences (*He died, and* **consequently** *I inherited the estate*).

Many English adverbs are formed from adjectives by adding the ending *-ly*, as in *hopefully*, *widely*, *theoretically* (for details of spelling and etymology, see *-ly*). Certain words can be used as both adjectives and adverbs, such as *fast*, *straight*, and *hard*; these are flat adverbs. In earlier usage more flat adverbs were accepted in formal usage; many of these survive in idioms and colloquially. (*That's just* **plain** *ugly*.) Some adjectives can also be used as flat adverbs when they actually describe the subject. (*The streaker ran* **naked**, *not* ***The streaker ran nakedly*.) The adverb corresponding to the adjective *good* is *well* (note that *bad* forms the regular *badly*, although *ill* is occasionally used in some phrases).

There are also many adverbs that are not derived from adjectives, including adverbs of time, of frequency, of place, of degree and with other meanings. Some suffixes that are commonly used to form adverbs from nouns are *-ward[s]* (as in *homeward[s]*) and *-wise* (as in *lengthwise*).

Most adverbs form comparatives and superlatives by modification with *more* and *most*: *often*, *more often*, *most often*;*smoothly*, *more smoothly*, *most smoothly* (see also comparison of adjectives, above). However, a few adverbs retain irregular inflection for comparative and superlative forms: *much*, *more*, *most*; *a little*, *less*, *least*; *well*, *better*, *best*; *badly*, *worse*, *worst*; *far*, *further* (*farther*), *furthest* (*farthest*); or follow the regular adjectival inflection: *fast*, *faster*, *fastest*; *soon*, *sooner*, *soonest*; etc.

Adverbs indicating the manner of an action are generally placed after the verb and its objects (*We considered the proposal **carefully***), although other positions are often possible (*We **carefully** considered the proposal*). Many adverbs of frequency, degree, certainty, etc. (such as *often*, *always*, *almost*, *probably*, and various others such as *just*) tend to be placed before the verb (*they **usually** have chips*), although if there is an auxiliary or other "special verb" (see § Verbs above), then the normal position for such adverbs is after that special verb (or after the first of them, if there is more than one): *I have **just** finished the crossword*; *She can **usually** manage a pint*; *We are **never** late*; *You might **possibly** have been unconscious*. Adverbs that provide a connection with previous information (such as *next*, *then*, *however*), and those that provide the context (such as time or place) for a sentence, are typically placed at the start of the sentence: ***Yesterday** we went on a shopping expedition.* If the verb has an object, the adverb comes after the object (*He finished the test **quickly***). When there is more than one types of adverb, they usually appear in the order: manner, place, time (*His arm was hurt **severely at home yesterday***).

A special type of adverb is the adverbial particle used to form phrasal verbs (such as *up* in *pick up*, *on* in *get on*, etc.) If such a verb also has an object, then the particle may precede or follow the object, although it will normally follow the object if the object is a pronoun (*pick the pen up* or *pick up the pen*, but *pick it up*).

Phrases

An adverb phrase is a phrase that acts as an adverb within a sentence. An adverb phrase may have an adverb as its head, together with any modifiers (other adverbs or adverb phrases) and complements, analogously to the adjective phrases described above. For example:*very sleepily*; *all too suddenly*; *oddly enough*; *perhaps shockingly for us*.

Another very common type of adverb phrase is the prepositional phrase, which consists of a preposition and its object: *in the pool*; *after two years*;*for the sake of harmony*.

Prepositions

Prepositions form a closed word class, although there are also certain phrases that serve as prepositions, such as *in front of*. A single preposition may have a variety of meanings, often including temporal, spatial and abstract. Many words that are prepositions can also serve as adverbs. Examples of common English prepositions (including phrasal instances) are *of*, *in*, *on*, *over*, *under*, *to*, *from*, *with*, *in front of*, *behind*, *opposite*, *by*, *before*, *after*, *during*, *through*, *in spite of* or *despite*, *between*, *among*, etc.

A preposition is usually used with a noun phrase as its complement. A preposition together with its complement is called a prepositional phrase. Examples are *in England*, *under the table*, *after six pleasant weeks*, *between the land and the sea*. A prepositional phrase can be used as a complement or post-modifier of a noun in a noun phrase, as in *the man in the car*, *the start of the fight*; as a complement of a verb or adjective, as in *deal with the problem*, *proud of oneself*; or generally as an adverb phrase (see above).

English allows the use of "stranded" prepositions. This can occur in interrogative and relative clauses, where the interrogative or relative pronoun that is the preposition's complement is moved to the start (fronted), leaving the preposition in place. This kind of structure is avoided in some kinds of formal English. For example:

- *What are you talking about?* (Possible alternative version: *About what are you talking?*)
- *The song that you were listening to …* (more formal: *The song to which you were listening …*)

Notice that in the second example the relative pronoun *that* could be omitted.

Stranded prepositions can also arise in passive voice constructions and other uses of passive past participial phrases, where the complement in a prepositional phrase can become zero in the same way that a verb's direct object would: *it was looked at*; *I will be operated on*; *get your teeth seen to.* The same can happen in certain uses of infinitive phrases: *he is nice to talk to*; *this is the page to make copies of.*

Conjunctions

Conjunctions express a variety of logical relations between items, phrases, clauses and sentences. The principal coordinating conjunctions in English are:*and*, *or*, *but*, *nor*, *so*, *yet*, and *for*. These can be used in many grammatical contexts to link two or more items of equal grammatical status, for example:

- Noun phrases combined into a longer noun phrase, such as *John, Eric, and Jill*, *the red coat or the blue one*. When *and* is used, the resulting noun phrase is plural. A determiner does not need to be repeated with the individual elements: *the cat, the dog, and the mouse* and *the cat, dog, and mouse* are both correct. The same applies to other modifiers. (The word *but*can be used here in the sense of "except": *nobody but you*.)
- Adjective or adverb phrases combined into a longer adjective or adverb phrase: *tired but happy*, *over the fields and far away*.
- Verbs or verb phrases combined as in *he washed, peeled, and diced the turnips* (verbs conjoined, object shared); *he washed the turnips, peeled them, and diced them* (full verb phrases, including objects, conjoined).
- Other equivalent items linked, such as prefixes linked in *pre- and post-test counselling*, numerals as in *two or three buildings*, etc.
- Clauses or sentences linked, as in *We came, **but** they wouldn't let us in. They wouldn't let us in, **nor** would they explain what we had done wrong.*

There are also correlative conjunctions, where as well as the basic conjunction, an additional element appears before the first of the items being linked. The common correlatives in English are:

- *either ... or* (**either** *a man* **or** *a woman*);
- *neither ... nor* (**neither** *clever* **nor** *funny*);
- *both ... and* (*they* **both** *punished* **and** *rewarded them*);
- *not ... but*, particularly in *not only ... but also* (**not** *exhausted* **but** *exhilarated*, **not only** *football* **but also** *many other sports*).

Subordinating conjunctions make relations between clauses, making the clause in which they appear into a subordinate clause. Some common subordinating conjunctions in English are:

- conjunctions of time, including *after, before, since, until, when, while*;
- conjunctions of cause and effect, including *because, since, now that, as, in order that, so*;
- conjunctions of opposition or concession, such as *although, though, even though, whereas, while*;
- conjunctions of condition: such as *if, unless, only if, whether or not, even if, in case (that)*;
- the conjunction *that*, which produces content clauses, as well as words that produce interrogative content clauses: *whether, where, when, how*, etc.

Subordinating conjunction generally comes at the very start of its clause, although many of them can be preceded by

qualifying adverbs, as in *probably because …, especially if ….* The conjunction *that* can be omitted after certain verbs, as in *she told us (that) she was ready.* (For the use of *that* in relative clauses, see § Relative pronouns above.)

Case

Although English has largely lost its case system, personal pronouns still have three morphological cases that are simplified forms of the nominative, objective and genitive cases:

- The *nominative case* (*subjective pronouns* such as *I, he, she, we, they, who, whoever*), used for the subject of a finite verb and sometimes for the complement of a copula.
- The *oblique case* (*object pronouns* such as *me, him, her, us, it, us, them, whom, whomever*), used for the direct or indirect object of a verb, for the object of a preposition, for an absolute disjunct, and sometimes for the complement of a copula.
- The *genitive case* (*possessive pronouns* such as *my/mine, his, her(s), our(s), its, our(s), their, theirs, whose*), used for a grammatical possessor. This is not always considered to be a case; see English possessive § Status of the possessive as a grammatical case.

Most English personal pronouns have five forms: the nominative and oblique case forms, the possessive case, which has both a *determiner* form (such as *my, our*) and a distinct *independent* form (such as *mine, ours*) (with two exceptions:

the third person singular masculine and the third person singular neuter *it*, which use the same form for both determiner and independent [*his car*, *it is his*]), and a distinct *reflexive* or *intensive* form (such as *myself, ourselves*). The interrogative personal pronoun *who* exhibits the greatest diversity of forms within the modern English pronoun system, having definite nominative, oblique, and genitive forms (*who, whom, whose*) and equivalently coordinating indefinite forms (*whoever, whomever*, and *whosever*).

Forms such as *I*, *he*, and *we* are used for the subject ("**I** kicked the ball"), whereas forms such as *me*, *him* and *us* are used for the object ("John kicked **me**").

Declension

Nouns have distinct singular and plural forms; that is, they *decline* to reflect their grammatical number; consider the difference between *book* and *books*. In addition, a few English pronouns have distinct nominative (also called subjective) and oblique (or objective) forms; that is, they decline to reflect their relationship to a verb or preposition, or case. Consider the difference between *he* (subjective) and *him* (objective), as in "He saw it" and "It saw him"; similarly, consider *who*, which is subjective, and the objective *whom.*

Further, these pronouns and a few others have distinct possessive forms, such as *his* and *whose.* By contrast, nouns have no distinct nominative and objective forms, the two being merged into a single *plain case.* For example, *chair* does not change form between "the chair is here" (subject) and "I saw the chair" (direct object). Possession is shown by the clitic-*'s*

attached to a possessive noun phrase, rather than by declension of the noun itself.

Negation

As noted above under § Verbs, a finite indicative verb (or its clause) is negated by placing the word *not* after an auxiliary, modal or other "special" verb such as *do*, *can* or *be*. For example, the clause *I go*is negated with the appearance of the auxiliary *do*, as *I do not go* (see *do*-support). When the **affirmative** already uses auxiliary verbs (*I am going*), no other auxiliary *verbs*are added to negate the clause (*I am not going*). (Until the period of early Modern English, negation was effected without additional auxiliary verbs: *I go not*.)

Most combinations of auxiliary verbs etc. with *not* have contracted forms: *don't*, *can't*, *isn't*, etc. (Also the uncontracted negated form of *can* is written as a single word *cannot*.) On the inversion of subject and verb (such as in questions; see below), the subject may be placed after a contracted negated form: *Should he not pay?*or*Shouldn't he pay?*

Other elements, such as noun phrases, adjectives, adverbs, infinitive and participial phrases, etc., can be negated by placing the word *not* before them: *not the right answer*, *not interesting*, *not to enter*, *not noticing the train*, etc.

When other negating words such as *never*, *nobody*, etc. appear in a sentence, the negating *not* is omitted (unlike its equivalents in many languages): *I saw nothing* or *I didn't see anything*, but not (except in non-standard speech) **I didn't see nothing* (see Double negative). Such negating words generally

have corresponding negative polarity items (*ever* for *never*, *anybody* for *nobody*, etc.) which can appear in a negative context but are not negative themselves (and can thus be used after a negation without giving rise to double negatives).

Clause and sentence structure

A typical sentence contains one independent clause and possibly one or more dependent clauses, although it is also possible to link together sentences of this form into longer sentences, using coordinating conjunctions (see above). A clause typically contains a subject (a noun phrase) and a predicate (a verb phrase in the terminology used above; that is, a verb together with its objects and complements). A dependent clause also normally contains a subordinating conjunction (or in the case of relative clauses, a relative pronoun, or phrase containing one).

Word order

English word order has moved from the Germanic verb-second (V2) word order to being almost exclusively subject–verb–object (SVO). The combination of SVO order and use of auxiliary verbs often creates clusters of two or more verbs at the center of the sentence, such as *he had hoped to try to open it.* In most sentences, English marks grammatical relations only through word order. The subject constituent precedes the verb and the object constituent follows it. The Object–subject–verb (OSV) may on occasion be seen in English, usually in the future tense or used as a contrast with the conjunction "but", such as

in the following examples: "Rome I shall see!", "I hate oranges, but apples I'll eat!".

Questions

Like many other Western European languages, English historically allowed questions to be formed by inverting the positions of the verb and subject. Modern English permits this only in the case of a small class of verbs ("special verbs"), consisting of auxiliaries as well as forms of the copula*be* (see subject–auxiliary inversion). To form a question from a sentence which does not have such an auxiliary or copula present, the auxiliary verb *do* (*does*, *did*) needs to be inserted, along with inversion of the word order, to form a question (see *do*-support). For example:

- She can dance. → Can she dance? (inversion of subject *she* and auxiliary *can*)
- I am sitting here. → Am I sitting here? (inversion of subject *I* and copula *am*)
- The milk goes in the fridge. → Does the milk go in the fridge? (no special verb present; *do*-support required)

The above concerns yes-no questions, but inversion also takes place in the same way after other questions, formed with interrogative words such as *where*, *what*, *how*, etc. An exception applies when the interrogative word is the subject or part of the subject, in which case there is no inversion. For example:

- I go. →Where do I go? (*wh*-question formed using inversion, with *do*-support required in this case)

- He goes. →Who goes? (no inversion, because the question word *who* is the subject)

Note that inversion does not apply in indirect questions: *I wonder where he is* (not **... where is he*). Indirect yes-no questions can be expressed using *if* or *whether* as the interrogative word: *Ask them whether/if they saw him.*

Negative questions are formed similarly; however, if the verb undergoing inversion has a contraction with *not*, then it is possible to invert the subject with this contraction as a whole. For example:

- John is going. (affirmative)
- John is not going. / John isn't going. (negative, with and without contraction)
- Isn't John going? / Is John not going? (negative question, with and without contraction respectively)

See also English auxiliaries and contractions § Contractions and inversion.

Dependent clauses

The syntax of a dependent clause is generally the same as that of an independent clause, except that the dependent clause usually begins with a subordinating conjunction or relative pronoun (or phrase containing such). In some situations (as already described) the conjunction or relative pronoun *that* can be omitted. Another type of dependent clause with no subordinating conjunction is the conditional clause formed by inversion (see below).

Other uses of inversion

The clause structure with an inverted subject and verb, used to form questions as described above, is also used in certain types of declarative sentences. This occurs mainly when the sentence begins with adverbial or other phrases that are essentially negative or contain words such as *only*, *hardly*, etc.: *Never have I known someone so stupid; Only in France can such food be tasted.*

In elliptical sentences (see below), inversion takes place after *so* (meaning "also") as well as after the negative *neither*:*so do I, neither does she.*

Inversion can also be used to form conditional clauses, beginning with *should*, *were* (subjunctive), or *had*, in the following ways:

- *should I win the race* (equivalent to *if I win the race*);
- *were he a soldier* (equivalent to *if he were a soldier*);
- *were he to win the race* (equivalent to *if he were to win the race*, i.e. *if he won the race*);
- *had he won the race* (equivalent to *if he had won the race*).

Other similar forms sometimes appear but are less common. There is also a construction with subjunctive *be*, as in *be he alive or dead* (meaning "no matter whether he is alive or dead").

Use of inversion to express a third-person imperative is now mostly confined to the expression *long live X,* meaning "let X live long".

Imperatives

In an imperative sentence (one giving an order), there is usually no subject in the independent clause: *Go away until I call you.* It is possible, however, to include *you* as the subject for emphasis: **You** *stay away from me.*

Elliptical constructions

Many types of elliptical construction are possible in English, resulting in sentences that omit certain redundant elements. Various examples are given in the article on Ellipsis.

Some notable elliptical forms found in English include:

- Short statements of the form *I can*, *he isn't*, *we mustn't*. Here the verb phrase (understood from the context) is reduced to a single auxiliary or other "special" verb, negated if appropriate. If there is no special verb in the original verb phrase, it is replaced by *do/does/did*: *he does*, *they didn't*.
- Clauses that omit the verb, in particular those like *me too*, *nor me*, *me neither*. The latter forms are used after negative statements. (Equivalents including the verb: *I do too* or *so do I*; *I don't either* or *neither do I*.)
- Tag questions, formed with a special verb and pronoun subject: *isn't it?*;*were there?*; *am I not?*

History of English grammars

The **history of English grammars** begins late in the sixteenth century with the *Pamphlet for Grammar* by William Bullokar. In

the early works, the structure and rules of English grammarwere based on those of Latin. A more modern approach, incorporating phonology, was introduced in the nineteenth century.

Sixteenth to eighteenth centuries

The first English grammar, *Pamphlet for Grammar* by William Bullokar, written with the seeming goal of demonstrating that English was quite as rule-bound as Latin, was published in 1586. Bullokar's grammar was faithfully modeled on William Lily's Latin grammar, *Rudimenta Grammatices* (1534). Lily's grammar was being used in schools in England at that time, having been "prescribed" for them in 1542 by Henry VIII. Although Bullokar wrote his grammar in English and used a "reformed spelling system" of his own invention, many English grammars, for much of the century after Bullokar's effort, were to be written in Latin; this was especially so for books whose authors were aiming to be scholarly. Christopher Cooper's *Grammatica Linguæ Anglicanæ* (1685) was the last English grammar written in Latin.

The yoke of Latin grammar writing bore down oppressively on much of the early history of English grammars. The goal of grammarians was to assimilate a reading and writing system that taught English speakers of all different social classes the same equitable pattern, relying on a set of new guidelines taken from their Latin language rules. Any attempt by one author to assert an independent grammatical rule for English was quickly followed by equal declarations by others of truth of the corresponding Latin-based equivalent. Even as late as the early nineteenth century, Lindley Murray, the author of one of

the most widely used grammars of the day, was having to cite "grammatical authorities" to bolster the claim that grammatical cases in English are different from those in Ancient Greek or Latin.

The focus on tradition, however, belied the role that other social forces had already begun to play in the early seventeenth century. In particular, increasing commerce, and the social changes it wrought, created new impetus for grammar writing. On the one hand, greater British role in international trade created demand for English grammars for speakers of other languages. Consequently, grammars were published in various European languages in the second half of the seventeenth century. On the other hand, English grammars were being written for "non-learned, native-speaker audiences" in Britain, such as women, merchants, tradesmen, and children. With education becoming more widespread by the early eighteenth century, many grammars, such as John Brightland's *A Grammar of the English tongue* (1759) and James Greenwood's *Essay towards a practical English grammar*, were intended for those without a Latin background, including the "fair sex" and children.

If by the end of the seventeenth century English grammar writing had made a modest start, totaling 16 new grammars since Bullokar's *Pamphlet* of 115 years before, by the end of the eighteenth, the pace was positively brisk; 270 new titles were added during that century. 83 percent of these titles were published in the late eighteenth century. Both publishing and demand, moreover, would continue to mushroom. The first half of the nineteenth century would see the appearance of almost 900 new books on English grammar. Showing little originality,

most new books took the tack of claiming—as justification for their appearance—that the needs of their particular target audience were still unmet or that a particular "grammatical point" had not been treated adequately in the preexisting texts, or oftentimes both. Texts that were both utilitarian and egalitarian were proliferating everywhere. Edward Shelley's *The people's grammar; or English grammar without difficulties for 'the million'* (1848), for example, was written for "the mechanic and hard-working youth, in their solitary struggles for the acquirement of knowledge." Similarly, William Cobbett's popular mid-century book was titled, *A Grammar of the English Language, In a Series of Letters: Intended for the Use of Schools and of Young Persons in General, but more especially for the use of Soldiers, Sailors, Apprentices, and Plough-Boys.*

Eighteenth-century prescriptive grammars

In 1745, Ann Fisher published her English Grammar which has been argued to have had influence on grammarians in the 18th century to follow and printed in more than 30 numbered editions, making it one of the most popular in addition to being the earliest English grammar. Later, Robert Lowth, Bishop of Oxford and thereafter of London, scholar of Hebrew poetry, and for a short time Oxford Professor of Poetry, was one of the best known of the widely emulated grammarians of the 18th century. A self-effacing clergyman, he published *A Short Introduction to English Grammar, with critical notes* (1762), his only work on the subject, without the author's name on the title page. His influence extended, through the works of his students Lindley Murray and William Cobbett, well into the

late 19th century. He would also become, among prescriptive grammarians, the target of choice for the criticism meted out by later descriptivist linguists. Lowth wrote against preposition stranding, using "whose" as the possessive case of "which", and using "who" instead of "whom" in certain cases.

In America in 1765, the American Rev. Dr. Samuel Johnson, founder and first president of King's College in New York City (now Columbia University) published in New York *An English Grammar; the First Easy Rudiments of Grammar Applied to the English Tongue.* It "appears to have been the first English grammar prepared by an American and published in America." In 1767, Johnson combined it with a Hebrew grammar, and published it as *An English and Hebrew grammar, being the first short rudiments of those two languages*, suggesting the languages be taught together to children, which went to four more imprints by 1776. Johnson developed his grammars independently of Lowth, but later corresponded and exchanged grammars with him.

In 2003, scholar Karen Cajka described nine English women who published grammars in the late eighteenth century: Ellin Devis, Dorothea Du Bois, Mrs. M. C. Edwards, Mrs. Eves, Ellenor Fenn (aka Mrs. Teachwell and Mrs. Lovechild), Ann Fisher, Jane Gardiner née Arden, Blanche Mercy, and Mrs. Taylor. They "together published a total of twelve discrete grammars, with over one hundred documented editions appearing well into the nineteenth century". The study of English grammar was seen as important in learning how to write English well, and in learning other languages later. It held a strong significance to many people in the United States with little to no income, and sparse educational backgrounds,

ranging from former slaves, to rail splitters or weavers. Learning it permitted individuals like these to speak and write the language with passionate fluency, helping them expand on their careers.

Nineteenth century to present

It was during the nineteenth century that modern-language studies became systematized. In the case of English, this happened first in continental Europe, where it was studied by historical and comparative linguists. In 1832, Danish philologist Rasmus Rask published an English grammar, *Engelsk Formlære*, part of his extensive comparative studies in the grammars of Indo-European languages. German philologist Jacob Grimm, the elder of the Brothers Grimm, included English grammar in his monumental grammar of Germanic languages, *Deutsche Grammatik* (1822–1837). German historical linguist Eduard Adolf Maetzner published his 1,700 page *Englische Grammatik* between 1860 and 1865; an English translation, *An English grammar: methodical, analytical and historical* appeared in 1874. Contributing little new to the intrinsic scientific study of English grammar, these works nonetheless showed that English was being studied seriously by the first professional linguists.

As phonology became a full-fledged field, *spoken English* began to be studied scientifically as well, generating by the end of the nineteenth century an international enterprise investigating the structure of the language. This enterprise comprised scholars at various universities, their students who were training to be teachers of English, and journals publishing new research. All the pieces were in place for new "large-scale

English grammars" which combined the disparate approaches of the previous decades. The first work to lay claim to the new scholarship was British linguist Henry Sweet's *A new English grammar: logical and historical*, published in two parts, Phonology and Accidence (1892) and Syntax (1896), its title suggesting not only continuity and contrast with Maetzner's earlier work, but also kinship with the contemporary *A New English Dictionary on Historical Principles* (begun 1884), later the *Oxford English Dictionary* (1895). Two other contemporary English grammars were also influential. *English Grammar: Past and Present*, by John Collinson Nesfield, was originally written for the market in colonial India. It was later expanded to appeal to students in Britain as well, from young men preparing for various professional examinations to students in "Ladies' Colleges." Other books by Nesfield include *A Junior Course In English Composition, A Senior Course In English Composition*, but it was his *A Manual Of English Grammar and Composition* that proved really successful both in Britain and her colonies—so much so that it formed the basis for many other grammar and composition primers including but not limited to *Warriner's English Grammar and Composition*, and *High School English Grammar and Composition*, casually called Wren & Martin, by P. C. Wren and H. Martin.*Grammar of spoken English* (1924), by H. E. Palmer, written for the teaching and study of English as a foreign language, included a full description of the intonation patterns of English.

The next set of wide-ranging English grammars were written by Danish and Dutch linguists. Danish linguist Otto Jespersen, who had coauthored a few books with Henry Sweet, began work on his seven-volume *Modern English grammar on historical principles* in the first decade of the twentieth century. The first

volume, *Sounds and Spellings*, was published in 1909; it then took forty years for the remaining volumes on syntax (volumes 2 through 5), morphology (volume 6), and syntax again (volume 7), to be completed. Jespersen's original contribution was in analyzing the various parts of a sentence in terms of categories that he named, *rank*, *junction*, and *nexus*, forgoing the usual word classes. His ideas would inspire the later work of Noam Chomsky and Randolph Quirk.

The Dutch tradition of writing English grammars, which began with Thomas Basson's *The Conjugations in Englische and Netherdutche* in the same year—1586—as William Bullokar's first English grammar (written in English), gained renewed strength in the early 20th century in the work of three grammarians: Hendrik Poutsma, Etsko Kruisinga, and Reinard Zandvoort. Poutsma's *Grammar of late modern English*, published between 1904 and 1929 and written for "continental, especially Dutch students," selected all its examples from English literature.

Timeline of English grammars

- 1551. John Hart*The Opening of the Unreasonable Writing of our Inglish Toung*
- 1586. William Bullokar: *Brief Grammar of English.*
- 1594. Paul Greaves: *Grammatica Anglicana.*
- 1612. Thomas Tomkis *De Analogia Anglicani Sermoni liber Grammaticus*, Royal Manuscript Collection of the British Library (12.F.xviii).
- 1617. Alexander Hume: *Orthographie and Congruitie of the Britan Tongue.*
- 1619/1621. Alexander Gill: *Logonomia Anglica.*

- 1634. Charles Butler: *English Grammar.*
- 1640. Ben Jonson: *The English Grammar.*
- 1646. Joshua Poole: *The English Accidence.*
- 1653. John Wallis: *Grammatica Linguæ Anglicanæ.*
- 1654. Jeremiah Wharton: *The English Grammar.*
- 1662. James Howell: *A New English Grammar.*
- 1669. John Newton: *School Pastime for Young Children: or the Rudiments of Grammar.*
- 1671. Thomas Lye: *The Child's Delight.*
- 1685. Christopher Cooper: *Grammatica Linguæ Anglicanæ.*
- 1688. Guy Miège: *The English Grammar.*
- 1693. Joseph Aickin: *The English grammar.*
- 1700. A. Lane: *A Key to the Art of Letters.*
- 1745. Ann Fisher*A New Grammar.*
- 1761. Joseph Priestley: The Rudiments of English Grammar:Adapted to the Use of Schools.
- 1762. Robert Lowth: *A short introduction to English grammar: with critical notes.*
- 1763. John Ash: *Grammatical institutes: or, An easy introduction to Dr. Lowth's English grammar.*
- 1765. William Ward: *An Essay on English Grammar.*
- 1766. Samuel Johnson: *A dictionary of the English Language...: to which is prefixed, a Grammar of the English Language.*
- 1772. Joseph Priestley: *The Rudiments of English Grammar: Adapted to the Use of Schools.*
- 1775. Ellin Devis: *The Accidence*
- 1772. Dorothea Du Bois: *Short English Grammar*
- 1795. Lindley Murray: *English grammar: adapted to the different classes of learners.*
- 1799. Jane Gardiner: *Young Ladies' Grammar*

- 1804. Noah Webster: *A Grammatical Institute of the English Language.*
- 1809. William Hazlitt: *A New and Improved Grammar of the English Tongue*
- 1818. William Cobbett: *A Grammar of the English Language, In a Series of Letters.*
- 1850. William Chauncey Fowler: *English grammar: The English language in its elements and forms.*
- 1874. Eduard Adolf Maetzner, *An English grammar: methodical, analytical, and historical. With a treatise on the orthography, prosody, inflections and syntax of the English tongue, and numerous authorities cited in order of historical development.* (English translation of *Englische Grammatik* (1860–65)).
- 1892/98. Henry Sweet: *A New English Grammar, Logical and Historical* (Part 1: Introduction, Phonology, and Accidence; Part 2: Syntax).
- 1904–1929. Hendrik Poutsma: *A Grammar of Modern English* (5 volumes).
- 1909–1932. Etsko Kruisinga: *A Handbook of Present-day English*
- 1909–1949. Otto Jespersen: *A Modern English Grammar on Historical Principles.*
- 1931/1935. George O. Curme: *A Grammar of the English Language.*
- 1945. R. W. Zandvoort: *A Handbook of English Grammar.*
- 1952. Charles C. Fries: *The Structure of English: An Introduction to the Construction of English Sentences.*
- 1984. M. A. K. Halliday: *An Introduction to Functional Grammar.*

- 1985. Randolph Quirk, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik: *A Comprehensive Grammar of the English Language.*
- 1999. Douglas Biber, Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan: *Longman Grammar of Spoken and Written English.*
- 2002. Rodney Huddleston and Geoffrey Pullum: *The Cambridge Grammar of the English Language.*
- 2006. Ronald Carter and Michael McCarthy: *The Cambridge Grammar of English.*
- 2011. Bas Aarts: *Oxford Modern English Grammar.*

Chapter 3

English Phonology

Like many other languages, English has wide variation in pronunciation, both historically and from dialect to dialect. In general, however, the regional dialects of English share a largely similar (but not identical) phonological system. Among other things, most dialects have vowel reduction in unstressedsyllables and a complex set of phonological features that distinguish fortis and lenisconsonants (stops, affricates, and fricatives).

Phonological analysis of English often concentrates on or uses, as a reference point, one or more of the prestige or standard accents, such as Received Pronunciation for England, General American for the United States, and General Australian for Australia. Nevertheless, many other dialects of English are spoken, which have developed independently from these standardized accents, particularly regional dialects. Information about these standardized accents functions only as a *limited* guide to all of English phonology, which one can later expand upon once one becomes more familiar with some of the many other dialects of English that are spoken.

Phonemes

A phoneme of a language or dialect is an abstraction of a speech sound or of a group of different sounds which are all perceived to have the same function by speakers of that particular language or dialect. For example, the English word

through consists of three phonemes: the initial "th" sound, the "r" sound, and a vowel sound. The phonemes in this and many other English words do not always correspond directly to the letters used to spell them (English orthography is not as strongly phonemic as that of many other languages).

The number and distribution of phonemes in English vary from dialect to dialect, and also depend on the interpretation of the individual researcher. The number of consonant phonemes is generally put at 24 (or slightly more). The number of vowels is subject to greater variation; in the system presented on this page there are 20–25 vowel phonemes in Received Pronunciation, 14–16 in General American and 19–20 in Australian English. The pronunciation keys used in dictionaries generally contain a slightly greater number of symbols than this, to take account of certain sounds used in foreign words and certain noticeable distinctions that may not be—strictly speaking—phonemic.

Consonants

The following table shows the 24 consonant phonemes found in most dialects of English, plus /x/, whose distribution is more limited. Fortis consonants are always voiceless, aspirated in syllable onset (except in clusters beginning with /s/), and sometimes alsoglottalized to an extent in syllable coda (most likely to occur with /t/, see T-glottalization), while lenis consonants are always unaspirated and un-glottalized, and generally partially or fully voiced. The alveolars are usually apical, i.e. pronounced with the tip of the tongue touching or approaching the roof of the mouth, though some speakers produce them laminally, i.e. with the blade of the tongue.

Sonorants

- The pronunciation of /l/ varies by dialect:
- Received Pronunciation has two main allophones of /l/: the clear or plain [l] (the "light L"), and the dark or velarized[ɫ] (the "dark L"). The clear variant is used before vowels when they are in the same syllable, and the dark variant when the /l/ precedes a consonant or is in syllable-final position before silence.
- In South Wales, Ireland, and the Caribbean, /l/ is usually clear, and in North Wales, Scotland, Australia, and New Zealand it is usually dark.
- In General American and Canada, /l/ is generally dark, but to varying degrees: before stressed vowels it is neutral or only slightly velarized. In southern U.S. accents it is noticeably clear between vowels, and in some other positions.
- In urban accents of Southern England, as well as New Zealand and some parts of the United States, /l/ can be pronounced as an approximant or semivowel ([w], [o], [ʊ]) at the end of a syllable (*l*-vocalization).
- Depending on dialect, /r/ has at least the following allophones in varieties of English around the world (see Pronunciation of English /r/):
- postalveolar approximant[ɹ] (the most common realization of the /r/ phoneme, occurring in most dialects, RP and General American included)
- retroflex approximant[ɻ] (occurs in most Irish dialects and some American dialects)

- labiodental approximant[ʋ] (occurs in south-east England and some London accents; known as *r-labialization*)

- alveolar flap[ɾ] (occurs in most Scottish, Welsh, Indian and some South African dialects, some conservative dialects in England and Ireland; not to be confused with flapping of /t/ and /d/)

- alveolar trill[r] (occurs in some very conservative Scottish dialects and some Indian, South African and Welsh accents)

- voiced uvular fricative[ʁ] (occurs in northern Northumbria, largely disappeared; known as the Northumbrian burr)

- In most dialects/r/ is labialized[ɹʷ] in many positions, as in *reed*[ɹʷiːd] and *tree*[tɹ̥ʷiː]; in the latter case, the /t/ may be slightly labialized as well.

- In some rhotic accents, such as General American, /r/ when not followed by a vowel is realized as an r-coloring of the preceding vowel or its coda: *nurse*[ˈnɚs], *butter*[ˈbʌɾɚ].

- The distinctions between the nasals are neutralized in some environments. For example, before a final /p/, /t/ or /k/ there is nearly always only one nasal sound that can appear in each case: [m], [n] or [ŋ] respectively (as in the words *limp*, *lint*, *link* – note that the *n* of *link* is pronounced [ŋ]). This effect can even occur across syllable or word boundaries, particularly in stressed syllables: *synchrony* is pronounced [ˈsɪŋkɹəni] whereas *synchronic* may be pronounced either as [sɪŋŋkɹɒnɪk] or as [sɪnˈkɹɒnɪk]. For other possible syllable-final combinations, see § Coda in the Phonotactics section below.

Obstruents

In most dialects, the fortis stops and affricate /p, t, tʃ, k/ have various different allophones, and are distinguished from the lenis stops and affricate /b, d, dʒ, g/ by several phonetic features.

- The allophones of the fortes /p, t, tʃ, k/ include:
- aspirated[pʰ, tʰ, kʰ] when they occur in the onset of a stressed syllable, as in *potato*. In clusters involving a following liquid, the aspiration typically manifests as the devoicing of this liquid. These sounds are unaspirated [p, t, k] after /s/ within the same syllable, as in *stan, span, scan*, and at the ends of syllables, as in *mat, map, mac*. The voiceless fricatives are always unaspirated, but a notable exception to this are English-speaking areas of Wales, where they are often aspirated.
- In many accents of English, fortis stops /p, t, k, tʃ/ are glottalized in some positions. This may be heard either as a glottal stop preceding the oral closure ("pre-glottalization" or "glottal reinforcement") or as a substitution of the glottal stop [ʔ] for the oral stop (glottal replacement). /tʃ/ can only be pre-glottalized. Pre-glottalization normally occurs in British and American English when the fortis consonant phoneme is followed by another consonant or when the consonant is in final position. Thus*football* and *catching* are often pronounced [ˈfʊʔtbɔːl] and [ˈkæʔtʃɪŋ], respectively. Glottal replacement often happens in cases such as those just given, so that *football*is frequently

pronounced[ˈfʊʔbɔːl]. In addition, however, glottal replacement is increasingly common in British English when /t/ occurs between vowels if the preceding vowel is stressed; thus *better* is often pronounced by younger speakers as [ˈbeʔə]. Such *t*-glottalization also occurs in many British regional accents, including Cockney, where it can also occur at the end of words, and where /p/ and /k/are sometimes treated the same way.

- Among stops, both fortes and lenes:
- May have no audible release[p̚, b̚, t̚, d̚, k̚, g̚] in the word-final position. These allophones are more common in North America than Great Britain.
- Always have a 'masked release' before another plosive or affricate (as in *rubbed*[ˈrʌˑb̚d̥]), i.e. the release of the first stop is made after the closure of the second stop. This also applies when the following stop is homorganic (articulated in the same place), as in *to**p** **p**layer*. A notable exception to this is Welsh English, where stops are usually released in this environment.
- The affricates /tʃ, dʒ/ have a mandatory fricative release in all environments.
- Very often in the United States and Canada, and less frequently in Australia and New Zealand, both /t/ and /d/ can be pronounced as a voiced flap[ɾ] in certain positions: when they come between a preceding stressed vowel (possibly with intervening /r/) and precede an unstressed vowel or syllabic/l/. Examples include *wa**t**er*, *bo**tt**le*, *pe**t**al*, *pe**dd**le* (the last two words sound alike when flapped). The flap may even appear at word boundaries, as in *pu**t it** on*.

When the combination /nt/ appears in such positions, some American speakers pronounce it as a nasalized flap that may become indistinguishable from /n/, so *winter*[ˈwɪɾ̃ɚ]may be pronounced similarly or identically to *winner*[ˈwɪnɚ].

- Yod-coalescence is a process that palatalizes the clusters/dj/, /tj/, /sj/ and /zj/ into [dʒ], [tʃ], [ʃ] and [ʒ] respectively, frequently occurring with clusters that would be considered to span a syllable boundary.

- Yod-coalescence in stressed syllables, such as in *tune* and *dune*, occurs in Australian, Cockney, Estuary English, Hiberno-English (some speakers), Newfoundland English, South African English, and to a certain extent in New Zealand English and Scottish English (many speakers). This can lead to additional homophony; for instance, *dew* and *due* come to be pronounced the same as *Jew*.

- In certain varieties—such as Australian English, South African English, and New Zealand English—/sj/ and /zj/ in stressed syllables can coalesce into [ʃ] and [ʒ], respectively. In Australian English for example, *assume* is pronounced [əˈʃʉːm] by some speakers. Furthermore, some British, Canadian, American, New Zealand and Australian speakers may change the /s/ sound to /ʃ/ before /tr/, so that a word having a cluster of ⟨str⟩ like in **str**ewnwould be pronounced[ʃtruːn].

- The postalveolar consonants/tʃ, dʒ, ʃ, ʒ/are strongly labialized: [tʃʷ dʒʷʃʷʒʷ].

- In addition to /tʃ, dʒ/, clusters /ts, dz, tr, dr, tθ, dð, pf, bv/ also have affricate-like realizations in certain

positions (as in *cats, roads, tram, dram, eighth, width, cupful, obvious*; see also § Onset), but usually only /tʃ, dʒ/ are considered to constitute the monophonemic affricates of English because (among other reasons) only they are found in all of morpheme-initial, -internal, and -final positions, and native speakers typically perceive them as single units.

Vowels

- English, much like other Germanic languages, has a particularly large number of vowel phonemes, and in addition the vowels of English differ considerably between dialects. Consequently, corresponding vowels may be transcribed with various symbols depending on the dialect under consideration. When considering English as a whole, lexical sets are often used, each named by a word containing the vowel or vowels in question. For example, the LOT set consists of words which, like *lot*, have /ɒ/ in Received Pronunciation and /ɑ/ in General American. The "LOT vowel" then refers to the vowel that appears in those words in whichever dialect is being considered, or (at a greater level of abstraction) to a diaphoneme, which represents this interdialectal correspondence. A commonly used system of lexical sets, devised by John C. Wells, is presented below; for each set, the corresponding phonemes are given for RP and General American, using the notation that will be used on this page.

The differences between these tables can be explained as follows:

- General American lacks a phoneme corresponding to RP /ɒ/ (LOT, CLOTH), instead using /ɑ/ in the LOT words and generally /ɔ/ in the CLOTH words. In a few North American accents, namely in Eastern New England (Boston), Western Pennsylvania (Pittsburgh), and to some degree in Pacific Northwest (Seattle, Portland) and Eastern Canadian English, LOT words do not have the vowel of PALM (the *father–bother* merger has not occurred) but instead merge with CLOTH/THOUGHT.

- Although the notation /ʌ/is used for the vowel of STRUT in RP and General American, the actual pronunciation is closer to a near-open central vowel[ɐ]. The symbol ⟨ʌ⟩ continues to be used for reasons of tradition (it was historically a back vowel) and because it is still back in other varieties.

- RP transcriptions use ⟨e⟩ rather than ⟨ɛ⟩ largely for convenience and historical tradition; it does not necessarily represent a different sound from the General American phoneme, although the RP vowel may be described as somewhat less open than the American one.

- The different notations used for the vowel of GOAT in RP and General American (/əʊ/ and /oʊ/) reflect a difference in the most common phonetic realizations of that vowel.

- The triphthongs given in the RP table are usually regarded as sequences of two phonemes (a diphthong

plus /ə/); however, in RP, these sequences frequently undergo smoothing into single diphthongs or even monophthongs.

- The different notations used here for some of the Australian vowels reflect the phonetic realization of those vowels in Australian: a central [ʉː] rather than [uː] in GOOSE, a more closed [e] rather than [ɛ] in DRESS, a close-mid [oː] rather than traditional RP's [ɔː] in THOUGHT, an open-mid [ɔ] rather than traditional RP's [ɒ] in LOT, an opener [a] rather than somewhat closer [ʌ] in STRUT, a central [aː] rather a back [ɑː] in CALM and START, and somewhat different pronunciations of most of the diphthongs. Note that central [ʉː] in GOOSE, close-mid [oː] in THOUGHT, open-mid [ɔ] in LOT and near-open [ɐ] in STRUT are standard realizations in modern RP and the difference between modern RP and Australian English in these vowels lies almost only in transcription rather than pronunciation.

- Both Australian /eː/ and RP /eə/ are long monophthongs, the difference between them lies in tongue height: Australian /eː/ is close-mid[eː], whereas the corresponding RP vowel is open-mid [ɛː].

- Australian has the *bad–lad* split, with distinctive short and long variants in various words of the TRAP set: a long phoneme /æː/in words like *bad* contrasts with a short /æ/ in words like *lad.* (A similar split is found in the accents of some speakers in southern England.)

Other points to be noted are these:

- The vowel /æ/ is coming to be pronounced more open (approaching [a]) by many modern RP speakers. In American speech, however, there is a tendency for it to become more closed, tenser and even diphthongized (to something like [eə]), particularly in certain environments, such as before a nasal consonant. Some American accents, for example those of New York City, Philadelphia and Baltimore, make a marginal phonemic distinction between /æ/ and /eə/, although the two occur largely in mutually exclusive environments. See /æ/ raising.

- A significant number of words (the BATH group) have /æ/ in General American, but /ɑː/ in RP. The pronunciation varies between /æ/ and /aː/ in Australia, with speakers from South Australia using /aː/ more extensively than speakers from other regions.

- In General American and Canadian (which are rhotic accents, where /r/ is pronounced in positions where it does not precede a vowel), many of the vowels can be r-colored by way of realization of a following /r/. This is often transcribed phonetically using a vowel symbol with an added retroflexion diacritic[˞]; thus the symbol [ɚ] has been created for an r-colored schwa (sometimes called schwar) as in LETT**ER**, and the vowel of START can be modified to make [ɑ˞] so that the word *start* may be transcribed [stɑ˞t]. Alternatively, the START sequence might be written[stɑɚt] to indicate an r-colored offglide. The vowel of NURSE is generally always r-colored in

these dialects, and this can be written [ɚ] (or as a syllabic [ɻ]).

- In modern RP and other dialects, many words from the CURE group are coming to be pronounced by an increasing number of speakers with the NORTH vowel (so *sure* is often pronounced like *shore*). Also the RP vowel /eə/ may be monophthongized to [ɛː].
- The vowels of FLEECE and GOOSE are commonly pronounced as narrow diphthongs, approaching [ɪi] and [ʊu], in RP. Near-RP speakers may have particularly marked diphthongization of the type [əi] and [əu ~ əʉ], respectively. In General American, the pronunciation varies between a monophthong and a diphthong.

Allophones of vowels

Listed here are some of the significant cases of allophony of vowels found within standard English dialects.

- Vowels are shortened when followed in a syllable by a voiceless (fortis) consonant. This is known as *pre-fortis clipping*. Thus in the following word pairs the first item has a shortened vowel while the second has a normal length vowel: 'right' /raɪt/ – 'ride' /raɪd/; 'face' /feɪs/ – 'phase' /feɪz/; 'advice' /ədvaɪs/ – 'advise' /ədvaɪz/.
- In many accents of English, tense vowels undergo breaking before /l/, resulting in pronunciations like [pʰiəɫ] for *peel*, [pʰuəɫ] for *pool*, [pʰeəɫ] for *pail*, and [pʰoəɫ] for *pole*.

- In RP, the vowel /əʊ/ may be pronounced more back, as [ɒʊ], before syllable-final /l/, as in *goal*. In Australian English the vowel /əʉ/ is similarly backed to [ɔʊ] before /l/. A similar phenomenon may occur in Southern American English.

- The vowel /ə/ is often pronounced [ɐ] in open syllables.

- The PRICE and MOUTH diphthongs may be pronounced with a less open starting point when followed by a voiceless consonant; this is chiefly a feature of Canadian speech (Canadian raising), but is also found in parts of the United States. Thus *writer* may be distinguished from *rider* even when flapping causes the /t/ and /d/ to be pronounced identically.

Unstressed syllables

Unstressed syllables in English may contain almost any vowel, but in practice vowels in stressed and unstressed syllables tend to use different inventories of phonemes. In particular, long vowels are used less often in unstressed syllables than stressed syllables. Additionally there are certain sounds—characterized by central position and weakness—that are particularly often found as the nuclei of unstressed syllables. These include:

- schwa, [ə], as in COMM**A** and (in non-rhotic dialects) LETT**ER** (COMMA–LETTER merger); also in many other positions such as *a*bout, *phot*o*graph*, *padd*o*ck*, etc. This sound is essentially restricted to unstressed syllables exclusively. In the approach

presented here it is identified as a phoneme /ə/, although other analyses do not have a separate phoneme for schwa and regard it as a reduction or neutralization of other vowels in syllables with the lowest degree of stress.

- r-colored schwa, [ɚ], as in LETT**ER** in General American and some other rhotic dialects, which can be identified with the underlying sequence /ər/.

- syllabic consonants: [l̩] as in *bott**le***, [n̩] as in *butt**on***, [m̩] as in *rhyth**m***. These may be phonemized either as a plain consonant or as a schwa followed by a consonant; for example *button*may be represented as /ˈbʌtn̩/ or /ˈbʌtən/ (see above under Consonants).

- [ɨ], as in *ros**e**s* and *mak**i**ng*. This can be identified with the phoneme /ɪ/, although in unstressed syllables it may be pronounced more centrally, and for some speakers (particularly in Australian and New Zealand and some American English) it is merged with /ə/ in these syllables (weak vowel merger). Among speakers who retain the distinction there are many cases where free variation between /ɪ/ and /ə/is found, as in the second syllable of *typ**i**cal*. (The OED has recently adopted the symbol ⟨ɪ̈⟩ to indicate such cases.)

- [ʉ], as in *arg**u**ment*, *t**o**day*, for which similar considerations apply as in the case of [ɨ]. (The symbol ⟨ʊ̈⟩is sometimes used in these cases, similarly to ⟨ɪ̈⟩.) Some speakers may also have a rounded schwa, [ɵ], used in words like *omission*[ɵˈmɪʃən].

- [i], as in *happ**y***, *coff**ee***, in many dialects (others have [ɪ] in this position). The phonemic status of this [i] is

not easy to establish. Some authors consider it to correspond phonemically with a close front vowel that is neither the vowel of KIT nor that of FLEECE; it occurs chiefly in contexts where the contrast between these vowels is neutralized, implying that it represents an archiphoneme, which may be written /i/. Many speakers, however, do have a contrast in pairs of words like *studied* and *studded* or *taxis* and *taxes*; the contrast may be [i] vs. [ɪ], [ɪ] vs. [ə] or [i] vs. [ə], hence some authors consider that the *happY*-vowel should be identified phonemically either with the vowel of KIT or that of FLEECE, depending on speaker. See also *happy*-tensing.

- [u], as in *infl**u**ence*, *t**o** each*. This is the back rounded counterpart to [i] described above; its phonemic status is treated in the same works as cited there.

Vowel reduction in unstressed syllables is a significant feature of English. Syllables of the types listed above often correspond to a syllable containing a different vowel ("full vowel") used in other forms of the same morpheme where that syllable is stressed. For example, the first *o* in *photograph*, being stressed, is pronounced with the GOAT vowel, but in *photography*, where it is unstressed, it is reduced to schwa. Also, certain common words (*a*, *an*, *of*, *for*, etc.) are pronounced with a schwa when they are unstressed, although they have different vowels when they are in a stressed position (see Weak and strong forms in English).

Some unstressed syllables, however, retain full (unreduced) vowels, i.e. vowels other than those listed above. Examples are

the /æ/ in **a**mbition and the /aɪ/ in *fin**i**te*. Some phonologists regard such syllables as not being fully unstressed (they may describe them as having *tertiary stress*); some dictionaries have marked such syllables as having secondary stress. However linguists such as Ladefoged and Bolinger (1986) regard this as a difference purely of vowel quality and not of stress, and thus argue that vowel reduction itself is phonemic in English. Examples of words where vowel reduction seems to be distinctive for some speakers include *chickar**ee*** vs. *chicor**y*** (the latter has the reduced vowel of HAPP**Y**, whereas the former has the FLEECE vowel without reduction), and *Phar**aoh*** vs. *farr**ow*** (both have the GOAT vowel, but in the latter word it may reduce to [ɵ]).

Lexical stress

Lexical stress is phonemic in English. For example, the noun ***in**crease* and the verb *in**crease**are distinguished by the positioning of the stress on the first syllable in the former, and on the second syllable in the latter. (See initial-stress-derived noun.) Stressed syllables in English are louder than non-stressed syllables, as well as being longer and having a higher pitch.

In traditional approaches, in any English word consisting of more than one syllable, each syllable is ascribed one of three degrees of stress: *primary*, *secondary* or *unstressed*. Ordinarily, in each such word there will be exactly one syllable with primary stress, possibly one syllable having secondary stress, and the remainder are unstressed. For example, the word *am**az**ing* has primary stress on the second syllable, while

the first and third syllables are unstressed, whereas the word ***organiza**tion* has primary stress on the fourth syllable, secondary stress on the first, and the second, third, and fifth unstressed. This is often shown in pronunciation keys using the IPA symbols for primary and secondary stress (which are ˈ and ˌ respectively), placed before the syllables to which they apply. The two words just given may therefore be represented (in RP) as /əˈmeɪzɪŋ/ and /ˌɔːgənaɪˈzeɪʃən/.

Some analysts identify an additional level of stress (*tertiary* stress). This is generally ascribed to syllables that are pronounced with less force than those with secondary stress, but nonetheless contain a "full" or "unreduced" vowel (vowels that are considered to be reduced are listed under English phonology § Unstressed syllables above). Hence the third syllable of *organization*, if pronounced with /aɪ/ as shown above (rather than being reduced to /ɪ/ or /ə/), might be said to have tertiary stress. (The precise identification of secondary and tertiary stress differs between analyses; dictionaries do not generally show tertiary stress, although some have taken the approach of marking all syllables with unreduced vowels as having at least secondary stress.)

In some analyses, then, the concept of lexical stress may become conflated with that of vowel reduction. An approach which attempts to separate these two is provided by Peter Ladefoged, who states that it is possible to describe English with only one degree of stress, as long as unstressed syllables are phonemically distinguished for vowel reduction. In this approach, the distinction between primary and secondary stress is regarded as a phonetic or prosodic detail rather than a phonemic feature – primary stress is seen as an example of

the predictable "tonic" stress that falls on the *final* stressed syllable of a prosodic unit. For more details of this analysis, see Stress and vowel reduction in English.

For stress as a prosodic feature (emphasis of particular words within utterances), see § Prosodic stress below.

Phonotactics

Phonotactics is the study of the sequences of phonemes that occur in languages and the sound structures that they form. In this study it is usual to represent consonants in general with the letter C and vowels with the letter V, so that a syllable such as 'be' is described as having CV structure. The IPA symbol used to show a division between syllables is the dot [.]. Syllabification is the process of dividing continuous speech into discrete syllables, a process in which the position of a syllable division is not always easy to decide upon.

Most languages of the world syllabify CVCV and CVCCV sequences as /CV.CV/ and /CVC.CV/ or /CV.CCV/, with consonants preferentially acting as the onset of a syllable containing the following vowel. According to one view, English is unusual in this regard, in that stressed syllables attract following consonants, so that ˈCVCV and ˈCVCCV syllabify as /ˈCVC.V/ and /ˈCVCC.V/, as long as the consonant cluster CC is a possible syllable coda; in addition, /r/ preferentially syllabifies with the preceding vowel even when both syllables are unstressed, so that CVrV occurs as /CVr.V/. This is the analysis used in the *Longman Pronunciation Dictionary*. However, this view is not widely accepted, as explained in the following section.

Syllable structure

The syllable structure in English is (C)V(C), with a near maximal example being *strengths* (/strɛŋkθs/, although it can be pronounced /strɛŋθs/). From the phonetic point of view, the analysis of syllable structures is a complex task: because of widespread occurrences of articulatory overlap, English speakers rarely produce an audible release of individual consonants in consonant clusters. This coarticulation can lead to articulatory gestures that seem very much like deletions or complete assimilations. For example, *hundred pounds* may sound like [hʌndɹɪ**b** paʊndz] and *jumped back* (in slow speech, [dʒʌmptbæk]) may sound like [dʒʌmpbæk], but X-ray and electropalatographic studies demonstrate that inaudible and possibly weakened contacts or lingual gestures may still be made. Thus the second /d/ in *hundred pounds* does not entirely assimilate to a labial place of articulation, rather the labial gesture co-occurs with the alveolar one; the "missing" [t] in *jumped back* may still be articulated, though not heard.

Division into syllables is a difficult area, and different theories have been proposed. A widely accepted approach is the maximal onset principle: this states that, subject to certain constraints, any consonants in between vowels should be assigned to the following syllable. Thus the word *leaving* should be divided /ˈliː.vɪŋ/ rather than */ˈliːv.ɪŋ/, and *hasty* is /ˈheɪ.sti/ rather than */ˈheɪs.ti/ or */ˈheɪst.i/. However, when such a division results in an onset cluster which is not allowed in English, the division must respect this. Thus if the word *extra* were divided */ˈɛ.kstrə/ the resulting onset of the second syllable would be /kstr/, a cluster which does not occur initially in English. The division /ˈɛk.strə/ is therefore

preferred. If assigning a consonant or consonants to the following syllable would result in the preceding syllable ending in an unreduced short vowel, this is avoided. Thus the word *comma* (in RP) should be divided /ˈkɒm.ə/ and not */ˈkɒ.mə/, even though the latter division gives the maximal onset to the following syllable.

In some cases, no solution is completely satisfactory: for example, in British English (RP) the word *hurry* could be divided /ˈhʌ.ri/ or /ˈhʌr.i/, but the former would result in an analysis with a syllable-final /ʌ/ (which is held to be non-occurring) while the latter would result in a syllable final /r/ (which is said not to occur in this accent). Some phonologists have suggested a compromise analysis where the consonant in the middle belongs to both syllables, and is described as ambisyllabic. In this way, it is possible to suggest an analysis of *hurry* which comprises the syllables /hʌr/ and /ri/, the medial /r/ being ambisyllabic. Where the division coincides with a word boundary, or the boundary between elements of a compound word, it is not usual in the case of dictionaries to insist on the maximal onset principle in a way that divides words in a counter-intuitive way; thus the word *hardware* would be divided /ˈhɑː.dweə/ by the M.O.P., but dictionaries prefer the division /ˈhɑːd.weə/.

In the approach used by the *Longman Pronunciation Dictionary*, Wells claims that consonants syllabify with the preceding rather than following vowel when the preceding vowel is the nucleus of a more salient syllable, with stressed syllables being the most salient, reduced syllables the least, and full unstressed vowels ("secondary stress") intermediate. But there are lexical differences as well, frequently but not exclusively

with compound words. For example, in *dolphin* and *selfish,* Wells argues that the stressed syllable ends in /lf/, but in *shellfish,* the /f/ belongs with the following syllable: /ˈdɒlf.ɪn, ˈself.ɪʃ/→[ˈdɒlfïn, ˈselfïʃ], but /ˈʃel.fɪʃ/→[ˈʃel·fɪʃ], where the /l/ is a little longer and the /ɪ/ is not reduced. Similarly, in *toe-strap* Wells argues that the second /t/ is a full plosive, as usual in syllable onset, whereas in *toast-rack* the second /t/ is in many dialects reduced to the unreleased allophone it takes in syllable codas, or even elided: /ˈtoʊ.stræp/, /ˈtoʊst.ræk/→[ˈto·ʊstɹæp, ˈtoʊs(t̚)ɹæk]; likewise *nitrate*/ˈnaɪ.treɪt/→[ˈnaɪt̥ɹeɪt] with a voiceless /r/ (and for some people an affricated *tr* as in *tree*), vs *night-rate*/ˈnaɪt.reɪt/→[ˈnaɪt̚ɹeɪt] with a voiced /r/.Cues of syllable boundaries include aspiration of syllable onsets and (in the US) flapping of coda /t, d/*(a tease*/əˈtiːz/→[əˈtʰiːz] vs. *at ease*/ətˈiːz/→[əɾˈiːz]), epenthetic stops like [t] in syllable codas (*fence*/ˈfens/→[ˈfents] but *inside*/ɪnˈsaɪd/→[ɪnˈsaɪd]), and r-colored vowels when the /r/ is in the coda vs. labialization when it is in the onset *(key-ring*/ˈkiː.rɪŋ/→[ˈkiːɹʷɪŋ] but *fearing*/ˈfiːr.ɪŋ/→[ˈfɪəɹɪŋ]).

Other onsets

Certain English onsets appear only in contractions: e.g. /zbl/ (*'sblood*), and /zw/ or /dzw/ (*'swounds* or *'dswounds*). Some, such as /pʃ/ (*pshaw*), /fw/ (*fwoosh*), or /vr/ (*vroom*), can occur in interjections. An archaic voiceless fricative plus nasal exists, /fn/ (*fnese*), as does an archaic /snj/ (*snew*).

Several additional onsets occur in loan words (with varying degrees of anglicization) such as /bw/ (*bwana*), /mw/ (*moiré*), /nw/ (*noire*), /tsw/ (*zwitterion*), /zw/ (*zwieback*), /dv/

(*Dvorak*), /kv/ (*kvetch*), /ʃv/ (*schvartze*), /tv/ (*Tver*), /tsv/ (*Zwickau*), /kdʒ/ (*Kjell*), /kʃ/ (*Kshatriya*), /tl/ (*Tlaloc*), /vl/ (*Vladimir*), /zl/ (*zloty*), /tsk/ (*Tskhinvali*), /hm/ (*Hmong*), /km/ (*Khmer*), and /ŋ/ (*Nganasan*).

Some clusters of this type can be converted to regular English phonotactics by simplifying the cluster: e.g. /(d)z/ (*dziggetai*), /(h)r/ (*Hrolf*), /kr(w)/ (*croissant*), /(ŋ)w/ (*Nguyen*), /(p)f/ (*pfennig*), /(f)θ/ (*phthalic*), /(t)s/ (*tsunami*), /(ŋ)k/ (*!kung*), and /k(ŋ)/ (*Xhosa*).

Others can be replaced by native clusters differing only in voice: /zb ~ sp/ (*sbirro*), and /zɡr ~ skr/ (*sgraffito*).

Nucleus

The following can occur as the nucleus:

- All vowel sounds
- /m/, /n/ and /l/ in certain situations (see below under word-level patterns)
- /r/ in rhotic varieties of English (e.g. General American) in certain situations (see below under word-level patterns)

Coda

Most (in theory, all) of the following except those that end with /s/, /z/, /ʃ/, /ʒ/, /tʃ/ or /dʒ/can be extended with /s/ or /z/ representing the morpheme -s/-z. Similarly, most (in theory, all) of the following except those that end with /t/ or /d/can be extended with /t/ or /d/ representing the morpheme -t/-d.

Wells (1990) argues that a variety of syllable codas are possible in English, even /ntr, ndr/ in words like *entry*/ˈɛntr.i/ and *sundry*/ˈsʌndr.i/, with /tr, dr/ being treated as affricates along the lines of /tʃ, dʒ/. He argues that the traditional assumption that pre-vocalic consonants form a syllable with the following vowel is due to the influence of languages like French and Latin, where syllable structure is CVC.CVC regardless of stress placement.

Syllable-level patterns

- Syllables may consist of a single vowel, meaning that onset and coda are not mandatory.
- The consonant /ŋ/ does not occur in syllable-initial position.
- The consonant /h/ does not occur in syllable-final position.
- Onset clusters ending in /j/are followed by /uː/ or its variants (see note 5 above).
- Long vowels and diphthongs are not found before /ŋ/, except for the mimetic words *boing* and *oink*, unassimilated foreign words such as Burmese *aung* and proper names such as *Taung*, and American-type pronunciations of words like *strong* (which have /ɔŋ/ or /ɑŋ/). The short vowels /ɛ, ʊ/ occur before /ŋ/ only in assimilated non-native words such as *ginseng* and *Song* (name of a Chinese dynasty) or non-finally in some dialects in words like *strength* and *length* as well as in varieties without the *foot-strut* split.
- /ʊ/ is rare in syllable-initial position (although in the northern half of England, [ʊ] is used for /ʌ/ and is common at the start of syllables).

- Stop + /w/ before /uː, ʊ, ʌ, aʊ/ (all presently or historically /u(ː)/) are excluded.
- Sequences of /s/ + C₁ + V̆ + C₁, where C₁ is a consonant other than /t/ and V̆ is a short vowel, are virtually nonexistent.

Word-level patterns

- /ə/ does not occur in stressed syllables.
- /ʒ/ does not occur in word-initial position in native English words, although it can occur syllable-initially as in *luxurious*/lʌɡˈʒʊəriəs/ in American English, and at the start of borrowed words such as *genre.*
- /m/, /n/, /l/ and, in rhotic varieties, /r/ can be the syllable nucleus (i.e. a syllabic consonant) in an unstressed syllable following another consonant, especially /t/, /d/, /s/ or /z/. Such syllables are often analyzed phonemically as having an underlying /ə/ as the nucleus. See above under Consonants.
- The short vowels are checked vowels, in that they cannot occur without a coda in a word-final stressed syllable. (This does not apply to /ə/, which does not occur in stressed syllables at all.)

Prosody

The prosodic features of English – stress, rhythm, and intonation – can be described as follows.

Prosodic stress

Prosodic stress is extra stress given to words or syllables when they appear in certain positions in an utterance, or when they receive special emphasis.

According to Ladefoged's analysis (as referred to under Lexical stress § Notes above), English normally has prosodic stress on the final stressed syllable in an intonation unit. This is said to be the origin of the distinction traditionally made at the lexical level between primary and secondary stress: when a word like *admiration* (traditionally transcribed as something like /ˌædmɪˈreɪʃən/) is spoken in isolation, or at the end of a sentence, the syllable *ra* (the final stressed syllable) is pronounced with greater force than the syllable *ad*, although when the word is not pronounced with this final intonation there may be no difference between the levels of stress of these two syllables.

Prosodic stress can shift for various pragmatic functions, such as focus or contrast. For instance, in the dialogue *Is it brunch tomorrow? No, it's**dinner** tomorrow*, the extra stress shifts from the last stressed syllable of the sentence, *to**mor**row*, to the last stressed syllable of the emphasized word, ***din**ner.*

Grammatical function words are usually prosodically unstressed, although they can acquire stress when emphasized (as in *Did you find the cat? Well, I found **a** cat*). Many English function words have distinct strong and weak pronunciations; for example, the word *a* in the last example is pronounced /eɪ/, while the more common unstressed *a* is pronounced /ə/. See Weak and strong forms in English.

Rhythm

English is claimed to be a *stress-timed* language. That is, stressed syllables tend to appear with a more or less regular rhythm, while non-stressed syllables are shortened to accommodate this. For example, in the sentence *One make of car is better than another*, the syllables *one*, *make*, *car*, *bett-* and *-noth-* will be stressed and relatively long, while the other syllables will be considerably shorter. The theory of stress-timing predicts that each of the three unstressed syllables in between *bett-* and *-noth-* will be shorter than the syllable *of* between *make* and *car*, because three syllables must fit into the same amount of time as that available for *of*. However, it should not be assumed that all varieties of English are stress-timed in this way. The English spoken in the West Indies, in Africa and in India are probably better characterized as syllable-timed, though the lack of an agreed scientific test for categorizing an accent or language as stress-timed or syllable-timed may lead one to doubt the value of such a characterization.

Intonation

Phonological contrasts in intonation can be said to be found in three different and independent domains. In the work of Halliday the following names are proposed:

- *Tonality* for the distribution of continuous speech into tone groups.
- *Tonicity* for the placing of the principal accent on a particular syllable of a word, making it the *tonic*

syllable. This is the domain also referred to as prosodic stress or sentence stress.

- *Tone* for the choice of pitch movement on the tonic syllable. (The use of the term *tone* in this sense should not be confused with the tone of tone languages, such as Chinese.)

These terms ("the Three Ts") have been used in more recent work, though they have been criticized for being difficult to remember. American systems such as ToBI also identify contrasts involving boundaries between intonation phrases (Halliday's *tonality*), placement of pitch accent (*tonicity*), and choice of tone or tones associated with the pitch accent (*tone*).

Example of phonological contrast involving placement of intonation unit boundaries (boundary marked by comma):

- Those who ran quickly, escaped. (the only people who escaped were those who ran quickly)
- Those who ran, quickly escaped. (the people who ran escaped quickly)

Example of phonological contrast involving placement of tonic syllable (marked by capital letters):

- I have plans to LEAVE. (= I am planning to leave)
- I have PLANS to leave. (= I have some drawings to leave)

Example of phonological contrast (British English) involving choice of tone (\ = falling tone, \/ = fall-rise tone)

- She didn't break the record because of the \ WIND. (= she did not break the record, because the wind held her up)
- She didn't break the record because of the \/ WIND. (= she did break the record, but not because of the wind)

There is typically a contrast involving tone between wh-questions and yes/no questions, the former having a falling tone (e.g. "Where did you \PUT it?") and the latter a rising tone (e.g. "Are you going /OUT?"), though studies of spontaneous speech have shown frequent exceptions to this rule. Tag questions asking for information are said to carry rising tones (e.g. "They are coming on Tuesday, /AREN'T they?") while those asking for confirmation have falling tone (e.g. "Your name's John, \ISN'T it.").

History of English pronunciation

The pronunciation system of English has undergone many changes throughout the history of the language, from the phonological system of Old English, to that of Middle English, through to that of the present day. Variation between dialects has always been significant. Former pronunciations of many words are reflected in their spellings, as English orthography has generally not kept pace with phonological changes since the Middle English period.

The English consonant system has been relatively stable over time, although a number of significant changes have occurred. Examples include the loss (in most dialects) of the [ç] and [x] sounds still reflected by the ⟨gh⟩ in words like *night* and

taught, and the splitting of voiced and voiceless allophones of fricatives into separate phonemes (such as the two different phonemes represented by ⟨th⟩). There have also been many changes in consonant clusters, mostly reductions, for instance those that produced the usual modern pronunciations of such letter combinations as ⟨wr-⟩, ⟨kn-⟩ and ⟨wh-⟩.

The development of vowels has been much more complex. One of the most notable series of changes is that known as the Great Vowel Shift, which began around the late 14th century. Here the [iː] and [uː] in words like*price* and *mouth* became diphthongized, and other long vowels became higher: [eː] became [iː] (as in *meet*), [aː] became [eː] and later [eɪ] (as in *name*), [oː] became [uː] (as in *goose*), and [ɔː] became [oː] and later [oʊ] (in RP now [əʊ]; as in *bone*). These shifts are responsible for the modern pronunciations of many written vowel combinations, including those involving a silent final ⟨e⟩.

Many other changes in vowels have taken place over the centuries (see the separate articles on the low back, high back and high front vowels, short A, and diphthongs). These various changes mean that many words that formerly rhymed (and may be expected to rhyme based on their spelling) no longer do. For example, in Shakespeare's time, following the Great Vowel Shift, *food*, *good* and *blood* all had the vowel [uː], but in modern pronunciation *good* has been shortened to [ʊ], while *blood* has been shortened and lowered to [ʌ] in most accents. In other cases, words that were formerly distinct have come to be pronounced the same – examples of such mergers include *meet–meat*, *pane–pain* and *toe–tow*.

Controversial issues

Velar nasal

- The phonemic status of the velar nasal consonant [ŋ]is disputed; one analysis claims that the only nasal phonemes in English are /m/ and /n/, while [ŋ] is an allophone of /n/ found before velar consonants. Evidence in support of this analysis is found in accents of the north-west Midlands of England where [ŋ] is only found before /k/ or /ɡ/, with *sung* being pronounced as [sʌŋɡ]. However, in most other accents of English *sung* is pronounced [sʌŋ], producing a three-way phonemic contrast *sum – sun – sung*/sʌm sʌn sʌŋ/ and supporting the analysis of the phonemic status of /ŋ/. In support of treating the velar nasal as an allophone of /n/, Sapir (1925) claims on psychological grounds that [ŋ] did not form part of a series of three nasal consonants: "no naive speaker of English can be made to feel in his bones that it belongs to a single series with /m/ and /n/... it still *feels* like ŋg." More recent writers have indicated that analyses of [ŋ] as an allophone of /n/ may still have merit, even though [ŋ] may appear both with and without a following velar consonant; in such analyses, an underlying /ɡ/ that is deleted by a phonological rule would account for occurrences of [ŋ] not followed by a velar consonant. Thus the phonemic representation of *sing* would be /sɪnɡ/ and that of *singer* is /sɪnɡə/; in order to reach the phonetic form [sɪŋ] and [sɪŋə], it is necessary to apply

a rule that changes /n/ to [ŋ] before /k/ or /ɡ/, then a second rule that deletes /ɡ/ when it follows [ŋ].

Vowel system

It is often stated that English has a particularly large number of vowel phonemes and that there are 20 vowel phonemes in Received Pronunciation, 14–16 in General American, and 20–21 in Australian English. These numbers, however, reflect just one of many possible phonological analyses. A number of "biphonemic" analyses have proposed that English has a basic set of short (sometimes called "simple" or "checked") vowels, each of which can be shown to be a phoneme and which can be combined with another phoneme to form long vowels and diphthongs. One of these biphonemic analyses asserts that diphthongs and long vowels may be interpreted as comprising a short vowel linked to a consonant. The fullest exposition of this approach is found in Trager & Smith (1951), where all long vowels and diphthongs ("complex nuclei") are made up of a short vowel combined with either /j/ (for which the authors use the symbol ⟨y⟩), /w/ or /h/ (plus /r/ for rhotic accents), each thus comprising two phonemes. Using this system, the word *bite* would be transcribed /bajt/, *bout* as /bawt/, *bar* as /bar/ and *bra* as /brah/. One attraction that the authors claim for this analysis is that it regularizes the distribution of the consonants /j/, /w/, and /h/ (as well as /r/ in non-rhotic accents), which would otherwise not be found in syllable-final position. Trager & Smith (1951) suggest nine simple vowel phonemes to allow them to represent all the accents of American and British English they surveyed, symbolized /i, e, æ/ (front vowels); /ɨ, ə, a/ (central vowels); and /u, o, ɔ/ (back

vowels). The analysis from Trager & Smith (1951) came out of a desire to build an "overall system" to accommodate all English dialects, with dialectal distinctions arising from differences in the ordering of phonological rules, as well as in the presence or absence of such rules. Another category of biphonemic analyses of English treats long vowels and diphthongs as conjunctions of two vowels. Such analyses, as found in Sweet (1877) or Kreidler (2004) for example, are less concerned with dialectal variation. In MacCarthy (1957), for example, there are seven basic vowels and these may be doubled (geminated) to represent long vowels. Some of the short vowels may also be combined with /i/ (/ei/*bay*, /ai/*buy*, /oi/*boy*), with /u/ (/au/bough, /ou/*beau*) or with /ə/ (/iə/*peer*, /eə/*pair*, /uə/*poor*). The vowel inventory of English RP in MacCarthy's system therefore totals only seven phonemes. Analyses such as these could also posit six vowel phonemes, if the vowel of the final syllable in *comma* is considered to be an unstressed allophone of that of *strut*. These seven vowels might be symbolized /i/, /e/, /a/, /o/, /u/, /ʌ/ and /ə/. Six or seven vowels is a figure that would put English much closer to the average number of vowel phonemes in other languages.

A radically different approach to the English vowel system was proposed by Chomsky and Halle. Their *Sound Pattern of English* (Chomsky & Halle 1968) proposed that English has lax and tense vowel phonemes which are operated on by a complex set of phonological rules to transform underlying phonological forms into surface phonetic representations. This generative analysis is not easily comparable with conventional analyses, but the total number of vowel phonemes proposed falls well short of the figure of 20 often claimed as the number of English vowel phonemes.

Chapter 4

Stress and Vowel Reduction in English

Stress is a prominent feature of the English language, both at the level of the word *(lexical stress)* and at the level of the phrase or sentence *(prosodic stress)*. Absence of stress on a syllable, or on a word in some cases, is frequently associated in English with vowel reduction – many such syllables are pronounced with a centralized vowel (schwa) or with certain other vowels that are described as being "reduced" (or sometimes with a syllabic consonant as the syllable nucleus rather than a vowel). Various phonological analyses exist for these phenomena.

Lexical and prosodic stress

Lexical stress (word stress) is regarded as being phonemic in English; the position of the stress is generally unpredictable and can serve to distinguish words. For example, the words *insight* and *incite* are distinguished in pronunciation only by the syllable being stressed. In *insight*, the stress is placed on the first syllable; and in *incite,* on the second. Similarly, the noun and the verb *increase* are distinguished by the placement of the stress in the same way – this is an example of an initial-stress-derived noun. Moreover, even within a given letter sequence and a given part of speech, lexical stress may distinguish between different words or between different meanings of the same word (depending on differences in theory

about what constitutes a distinct word): For example, initial-stress pronunciations of *offense*/ˈɔfɛns/ and *defense*/ˈdifɛns/ in American English denote concepts specific to sports, whereas pronunciations with stress on the words' respective second syllables (*offense*/əˈfɛns/ and *defense*/dəˈfɛns/) denote concepts related to the legal (and, for *defense*, the military) field and encountered in sports only as borrowed from the legal field in the context of adjudicating rule violations. British English stresses the second syllable in both sports and legal use.

Some words are shown in dictionaries as having two levels of stress: primary and secondary.

For example, the RP pronunciation of *organization*may be given as /ˌɔːɡənaɪˈzeɪʃən/, with primary stress on the fourth syllable, secondary stress on the first syllable, and the remaining syllables unstressed. For different ways of analysing levels of stress in English, see § Degrees of lexical stress below.

English also has relatively strong prosodic stress—particular words within a phrase or sentence receive additional stress to emphasize the information they convey.

There is also said to be a natural "tonic stress" that falls on the last stressed syllable of a prosodic unit – for more on this, see below under § Descriptions with only one level of stress.

English is classified as a *stress-timed language*, which means that there is a tendency to speak so that the stressed syllables come at roughly equal intervals. See Isochrony § Stress timing.

Reduced vowels

Certain vowel sounds in English are associated strongly with absence of stress: they occur practically exclusively in unstressed syllables; and conversely, most (though not all) unstressed syllables contain one of these sounds. These are known as **reduced vowels**, and tend to be characterized by such features as shortness, laxness and central position. The exact set of reduced vowels depends on dialect and speaker; the principal ones are described in the sections below.

Schwa and r-coloured schwa

Schwa, [ə], is the most common reduced vowel in English. It may be denoted orthographically by any of the vowel letters, as the *a* in ***a**bout*, the *e* in *synth**e**sis*, the *o* in *harm**o**ny*, the *u* in *medi**u**m*, the *i* in *dec**i**mal* and the *y* in *s**y**ringe* (although the last two are pronounced as a near-close vowel by some speakers – see the following section).

In many rhotic dialects, an r-colored schwa, [ɚ], occurs in words such as *wat**er*** and *stand**ard.*** Non-rhotic dialects simply have schwa in these positions, except where the dialect has linking R. The r-colored schwa can be analyzed phonemically as /ər/.

Reduced vowels in the close unrounded area

In some dialects of English there is a distinction between two vowel heights of reduced vowels: in addition to schwa, there is a distinct near-close central unrounded vowel[ɪ̈] (or

equivalently [ɨ̞]). In the British phonetic tradition, the latter vowel is represented with the symbol ⟨ɪ⟩, and in the American tradition ⟨i⟩. An example of a minimal pair contrasting these two reduced vowels is *Rosa's* vs. *roses*: the *a* in *Rosa's* is a schwa, while the *e* in *roses* (for speakers who make the distinction) is the near-close vowel. See weak vowel merger.

This vowel is sometimes informally referred to as **schwi** in analogy with *schwa*.

Like schwa, [ï] does not correspond in spelling to any single vowel letter. It can be represented by *a* (for example, *message*[ˈmɛsïdʒ], *climate*[ˈklaɪmï̈t], *orange*[ˈɒɹïndʒ]), *e* (*puppet*), *i* (*limit*), *u* (*minute*), or *y* (*polyp*).

Among speakers who make this distinction, the distributions of schwa and [ï] are quite variable, and in many cases the two are in free variation: the *i* in *decimal*, for example, may be pronounced with either sound.

A symbolization convention recently introduced by Oxford University Press for some of their English dictionaries uses the non-IPA "compound" symbol ⟨ɨ⟩ (ɨ) in words that may be pronounced with either [ï] or schwa. For example, the word *noted* is transcribed /ˈnəʊtɨd/.

The final vowel of words like *happy* and *coffee* is an unstressed front close unrounded vowel most commonly represented with [i], although some dialects (including more traditional Received Pronunciation) may have [ɪ]. This [i] used to be identified with the phoneme /iː/, as in FLEECE. See *happy* tensing. However, some contemporary accounts regard it as a symbol

representing a close front vowel that is neither the vowel of KIT nor that of FLEECE; it occurs in contexts where the contrast between these vowels is neutralized; these contexts include unstressed prevocalic position within the word, such as *react*/riˈækt/. For some speakers, however, there is a contrast between this vowel and /ɪ/ in such pairs as *taxis* vs. *taxes* and *studied* vs. *studded.* See English phonology: § *Unstressed syllables* under § *Vowels.*

Reduced vowels in the close rounded area

According to Bolinger (1986:347–360), there is a reduced rounded phoneme /ɵ/ as in *willow*/ˈwɪlɵ/, *omission*/ɵˈmɪʃən/, thus forming a three-way contrast with *Willa*/ˈwɪlə/ and *Willie*/ˈwɪlɨ/ or with *a mission*/əˈmɪʃən/ and *emission*/ɨˈmɪʃən/.

The vowel is sometimes informally referred to as **schwu** in analogy with *schwa.*

Analogously to the ⟨ɨ⟩ symbol mentioned above, Oxford University Press have devised the non-IPA symbol ⟨ʊ̈⟩ to represent a vowel that may be either /ʊ/ or /ə/ in free variation. For example, *awful*/ˈɔːfʊ̈l/ may be pronounced /ˈɔːfəl/ or /ˈɔːfʊl/. Phonologically, this vowel is an archiphoneme representing the neutralization of /ʊ/ and /ə/.

A rounded vowel [u], corresponding to the [i]*happY* vowel, is widely used in British works for words such as *infl**u**ence*/ˈɪnfluəns/, *int**o**/ˈɪntu/. Phonologically, this vowel is an archiphoneme representing the neutralization of /uː/ and /ʊ/.

Syllabic consonants

The other sounds that can serve as the peaks of reduced syllables are the syllabic consonants, which can result in syllables with no vowel sound. Alternative pronunciations of syllabic consonants are however also possible. For example, *cycle*may be pronounced as either /ˈsaɪkl/ with only a dark l sound or as /ˈsaɪkəl/ with a schwa and the dark l sound.

In other words, a syllabic consonant can be phonologically analyzed as consisting of either just the consonant or of an underlying schwa followed by the consonant. The consonants that can be syllabic in English are principally /l/, /m/, and /n/, for example in *cycle* (followed by a silent e), *pris**m***, and *pris**on***. In rhotic accents, /ɜr/ and /ər/are also pronounced as syllabic [ɹ] or [ɻ].

Unstressed full vowels

All full (unreduced) vowels may occur in unstressed position (except under theoretical approaches that routinely assign secondary or tertiary stress to syllables containing such vowels – see § Degrees of lexical stress below). Some examples of words with unstressed syllables that are often pronounced with full vowels in Received Pronunciation are given below (pronunciation may be different in other varieties of English).

- Unreduced short vowels: /ɛ/ in the final syllable of *document* when used as a verb (compare the /ə/ heard when the word is used as a noun); /æ/ in the first syllable of *ambition*; /ɒ/ in the second syllable

of *neon*; /ʌ/ in words with the negative prefix *un-*, such as *unknown* (compare /ə/ in *until*).

- Long vowels: /ɑː/ in the final syllable of *grandma*; /ɔː/ in the final syllable of *outlaw*; /uː/ in *tofu*; /ɜː/ in the noun *convert*; /iː/ in *manatee*. Note that this last may stand in contrast to the *happY* vowel found at the end of *humanity*. This contrast is further described under § Distinctions between reduced and unreduced vowels below.
- Diphthongs: /eɪ/ in *Monday*; /əʊ/ in *piano*; /aʊ/ in *discount*; /aɪ/ in *idea*; /ɔɪ/ in *royale*.

Full vowels can often be found in unstressed syllables in compound words, as in *bedsh**ee**t*, *moonl**i**t*, *tentp**e**g*, *snowm**a**n*, and *kettledr**u**m*. However, in some well-established compounds the vowel of the unstressed part may be reduced, as in *postman*/ˈpəʊstmən/.

Many other full unstressed vowels also derive historically from stressed vowels, due to shifts of stress over time (such as stress shifting away from the final syllable of French loan words, like *ballet* and *bureau*, in British English), or the loss or change of stress in compound words or phrases (as in *óverseas vóyage* from *overséas* or *óverséas* plus *vóyage*). There is a tendency, though, for such vowels to become reduced over time, especially in common words.

With vowels represented as ⟨ɪ⟩ and ⟨ʊ⟩, it may be hard to ascertain whether they represent a full vowel or a reduced vowel. A word that illustrates the contrast is *chauvinism*, where the first *i* is the reduced vowel /ɨ/, and the second is unreduced /ɪ/.

Degrees of lexical stress

Descriptions with primary and secondary stress

In many phonological approaches, and in many dictionaries, English is represented as having two levels of stress: primary and secondary. In every lexical word, and in some grammatical words, one syllable is identified as having primary stress, though in monosyllables the stress is not generally marked. In addition, longer words may have one or more syllables identified as having secondary stress.

Syllables that have neither primary nor secondary stress are called unstressed.

In International Phonetic Alphabet transcriptions, primary stress is denoted with ˈ and secondary stress with ˌ. IPA stress marks are placed before the stressed syllable. When citing words in English spelling, primary stress is sometimes denoted with an acute accent ´ and secondary stress with a grave accent `, placed over the vowel of the stressed syllable.

Secondary stress is frequently indicated in the following cases:

- In words where the primary stress falls on the third syllable or later, it is normal for secondary stress to be marked on one of the first two syllables of the word. In words where the primary stress falls on the third syllable, secondary stress usually falls on the first rather than the second syllable. For example, *ìnterjéction* and *èvolútion* have their primary stress on the third syllable, and secondary stress on the

first syllable. However, in certain words with primary stress on the third syllable, the second syllable may have secondary stress corresponding to the primary stress of a shorter related word or base. For example, *electricity* is pronounced by some speakers with secondary stress on the second syllable (*elèctrícity*), corresponding to the primary stress in *eléctric*. In words where the primary stress falls on the fourth syllable or later, the position of the secondary stress on either the first or second syllable often corresponds to the position of the primary stress in a shorter related word or base. For example, *òrganizátion* and *assòciátion*, which both have primary stress on the fourth syllable, have secondary stress on the first and second syllable respectively: the same positions as the primary stress on the first syllable of *organize* and the second syllable of *associate*.

- In words where the primary stress falls on the third or fourth syllable from the end, a following syllable may be marked with secondary stress.

- In many compound words, where one part of the compound is pronounced more prominently; here the stressed syllable of the prominent part of the compound is marked with primary stress, while the stressed syllable of the other part may be marked with secondary stress. For example, *còunterintélligence*[ˌkaʊntər.ɪnˈtɛlɪdʒəns], and *cóunterfòil*[ˈkaʊntərˌfɔɪl]. Dictionaries are not always consistent in this, particularly when the secondary stress would come after the primary – for instance the *foil* of *counterfoil* is transcribed with secondary

stress in Merriam-Webster dictionaries but not in the OED, although both of them assign secondary stress to the *counter* of *counterintelligence.*

- In some dictionaries (particularly American ones), all syllables that contain a full (unreduced) vowel are ascribed at least secondary stress, even when they come after the primary stress (as in the *counter**foil*** example above). Bolinger (1986:358–360) notes that such dictionaries make use of the secondary-stress mark to distinguish full vowels from reduced vowels in unstressed syllables, as they may not have distinct symbols for reduced vowels. John Wells remarks, "Some analysts (particularly Americans) argue […] that the presence of a strong [= full] vowel is sufficient evidence that the syllable in question is stressed. In the British tradition we regard them as unstressed."

Note that this last-mentioned group of syllables are those ascribed tertiary stress in the approach described in the next section.

Descriptions with primary, secondary and tertiary stress

In some theories, English has been described as having three levels of stress: primary, secondary, and tertiary (in addition to the unstressed level, which in this approach may also be called quaternary stress). For example, our examples would be *²coun.ter.³in.¹tel.li.gence* and *¹coun.ter.³foil.* Exact treatments vary, but it is common for tertiary stress to be assigned to those syllables that, while not assigned primary or secondary stress, nonetheless contain *full* vowels (unreduced vowels, i.e.,

those not among the reduced vowels listed in the previous section). Dictionaries do not generally mark tertiary stress, but as mentioned above, some of them treat all syllables with unreduced vowels as having at least secondary stress.

Descriptions with only one level of stress

Phoneticians such as Peter Ladefoged have noted that it is possible to describe English with only one degree of stress, as long as unstressed syllables are phonemically distinguished for vowel reduction. According to this view, the posited multiple levels, whether primary–secondary or primary–secondary–tertiary, are mere phonetic detail and not true phonemic stress. They report that often the alleged secondary (or tertiary) stress in English is not characterized by the increase in respiratory activity normally associated with primary stress in English or with all stress in other languages. In their analysis, an English syllable may be either stressed or unstressed, and if unstressed, the vowel may be either full or reduced. This is all that is required for a phonemic treatment.

The difference between what is normally called primary and secondary stress, in this analysis, is explained by the observation that the last stressed syllable in a normal prosodic unit receives additional intonational or "tonic" stress. Since a word spoken in isolation, in citation form (as for example when a lexicographer determines which syllables are stressed) acquires this additional tonic stress, it may appear to be inherent in the word itself rather than derived from the utterance in which the word occurs. (The tonic stress may also occur elsewhere than on the final stressed syllable, if the speaker uses contrasting or other prosody.)

This combination of lexical stress, phrase- or clause-final prosody, and the lexical reduction of some unstressed vowels, conspires to create the impression of multiple levels of stress. In Ladefoged's approach, our examples are transcribed phonemically as *cóunterintélligence*/ˈkaʊntər.ɪnˈtɛlɪdʒəns/, with two stressed syllables, and *cóunterfoil*/ˈkaʊntərfɔɪl/, with one. In citation form, or at the end of a prosodic unit (marked [‖]), extra stress appears from the utterance that is not inherent in the words themselves: *cóunterintélligence*[ˈkaʊntər.ɪnˈ ˈtɛlɪdʒəns‖] and *cóunterfoil*[ˈˈkaʊntərfɔɪl‖].

To determine where the actual lexical stress is in a word, one may try pronouncing the word in a phrase, with other words before and after it and without any pauses between them, to eliminate the effects of tonic stress: in *the còunterintèlligence commúnity,* for example, one can hear secondary (that is, lexical) stress on two syllables of *counterintelligence,* as the primary (tonic) stress has shifted to *community.*

Distinctions between reduced and unreduced vowels

As mentioned in the previous section, some linguists make a phonemic distinction between syllables that contain reduced vowels (as listed above – syllabic consonants are also included in this category), and those that, while being phonetically unstressed, nevertheless contain a full (unreduced) vowel. In some analyses syllables of the latter type are ascribed secondary stress (those of the former type being regarded as completely unstressed), while in others the reduced/unreduced distinction is regarded as one of vowel quality not involving

any difference in stress. This last approach is taken by linguists such as Ladefoged and Bolinger, who thus consider that there are two "tiers" of vowels in English, full and reduced.

A distinction of this type may become useful for the analysis of a potential contrast between words such as *humanity*, *chicory*, *shivery* and *manatee*, *chickaree*, *shivaree*. When assuming a separate set of reduced vowels, the former may end with /ɨ/, while the latter may end with an unreduced /iː/. Another example, for some speakers, is provided by the words *farrow* and *Pharaoh*; the former may end with a reduced /ɵ/ while the latter may end with the unreduced /oʊ/. Alternatively, these reduced vowels can be analyzed as instances of the same phonemes as full vowels. In that case, it may be the phonemic secondary stress that distinguishes these words.

Some linguists have observed phonetic consequences of vowel reduction that go beyond the pronunciation of the vowel itself. Bolinger (1986) observes that a preceding voiceless stop is likely to retain its aspiration before an unstressed full vowel, but not before a reduced vowel; and that flapping of /t/ and /d/ in American English is possible before a reduced vowel but not before a full vowel. Hence the /t/ in *manatee* would be an aspirated [tʰ], while that in *humanity* would be unaspirated [t] or a flap [ɾ]. Wells (1990) explains such phenomena by claiming that, in the absence of morpheme boundaries or phonotactical constraints, a consonant between a full and a reduced vowel generally belongs to the syllable with the full vowel, whereas a consonant between two reduced vowels belongs to the preceding syllable. According to this analysis, *manatee* is /ˈmæn.ə.tiː/ and *humanity* is /hjʊ.ˈmæn.ɨt.i/; it is then

asserted that voiceless stops are only aspirated at the beginning of syllables, and /t/ can only be flapped at the end of a syllable (as in *might I*/maɪt.aɪ/→[mʌɪɾaɪ] versus *my tie*/maɪ.taɪ/→[maɪtʰaɪ]).

Alternation between full and reduced vowels

It is a feature of English that reduced vowels frequently alternate with full vowels: a given word or morpheme may be pronounced with a reduced vowel in some instances and a full vowel in other instances, usually depending on the degree of stress (lexical or prosodic) given to it.

Alternation depending on lexical stress

When the stress pattern of words changes, the vowels in certain syllables may switch between full and reduced. For example, in *photograph* and *photographic*, where the first syllable has (at least secondary) stress and the second syllable is unstressed, the first *o* is pronounced with a full vowel (the diphthong of GOAT), and the second *o* with a reduced vowel (schwa). However, in *photography* and *photographer*, where the stress moves to the second syllable, the first syllable now contains schwa while the second syllable contains a full vowel (that of LOT).

Alternation depending on meaning

There are a number of English verb-adjective pairs that are distinguished solely by vowel reduction. For example, in some

dialects, *separate* as a verb (as in 'what separates nation from nation') has a full final vowel, [ˈsɛpəreɪt], whereas the corresponding adjective (as in 'they sleep in separate rooms') has a reduced vowel: [ˈsɛpərət] or [ˈsɛprət]. A distinction may be made in a similar way between a verb and a noun, as in the case of *document* (pronounced with a schwa in the noun's final syllable and sometimes pronounced with a full vowel /ɛ/ in the verb's final syllable). Finally, differences in syllabic stress and vowel reduction (or lack of the latter) may distinguish between meanings even within a given part of speech, with the best-known such pairs in American English being *offense* and *defense* (in each case with the first syllable accented in the context of sports and the second syllable accented in legal contexts).

Alternation depending on type of enunciation

In some words, the reduction of a vowel depends on how quickly or carefully the speaker enunciates the word. For example, the *o* in *obscene* is commonly reduced to schwa, but in more careful enunciation it may also be pronounced as a full vowel (that of LOT). Compare this with the *o* in *gallon*, which is never a full vowel, no matter how carefully one enunciates.

Weak and strong forms of words

Some monosyllabic English function words have a **weak form** with a reduced vowel, used when the word has no prosodic stress, and a phonemically distinct **strong form** with a full vowel, used when the word is stressed (and as the citation form or isolation form when a word is mentioned standing alone). In

the case of many such words the strong form is also used when the word comes at the end of a sentence or phrase.

An example of such a word is the modal verb *can*. When appearing unstressed within a sentence and governing a verb (as in *I can do it*), the weak form /kən/ is used. However the strong form /kæn/ is used:

- when the word is stressed: *I don't **have** to do it, but I **can** do it*
- when the word is phrase-final, i.e. without a governed verb: *we won't be doing it, but they can if they want*
- when the word is referred to in isolation: *The verb "can" is one of the English modals.*

In the case of most words with such alternative forms, the weak form is much more common (since it is relatively rare for function words to receive prosodic stress). This is particularly true of the English articles *the*, *a*, *an*, whose strong forms are used within normal sentences only on the rare occasions when definiteness or indefiniteness is being emphasized: *Did you find the cat? I found **a**[eɪ] cat.* (i.e. maybe not the one you were referring to). The weak form of *the* is typically [ði] before a vowel-initial word (*the apple*) but [ðə] before a consonant-initial word (*the pear*), although this distinction is being lost in the United States. A similar distinction is sometimes made with *to*: *to Oxford*[tu] vs. *to Cambridge*[tə].

The exact set of words that have weak forms depends on dialect and speaker; the following is a list of the chief words of this type in Received Pronunciation:

- Always reduced:
- *a*, *an*, *and*, *be*, *been*, *but*, *he*, *her*, *him*, *his*, *just*, *me*, *or*, *she*, *than*, *that* (as conjunction), *the*, *them*, *us*, *we*, *who*, *you*, *your*.
- Reduced, but stressed at the end of a sentence:
- *as*, *at*, *for*, *from*, *of*, *to*, *some*, *there*.
- Reduced, but stressed at the end of a sentence and when contracted with the negative *not*:
- *am*, *are*, *can*, *could*, *do*, *does*, *had*, *has*, *have*, *must*, *shall*, *should*, *was*, *were*, *will*, *would*.

In most of the above words the weak form contains schwa, or a syllabic consonant in the case of those ending /l/, /m/ or /n/. However, in *be*, *he*, *me*, *she*, *we*, *been*, *him* the vowel may be the reduced form of /ɪ/, or else [i]; and in *do*, *who*, *you* it may be the reduced form of /ʊ/, or [u]. (For *the* and *to*, see above.) These various sounds are described in the § Reduced vowels section above.

The weak form of *that* is used only for the conjunction or relative pronoun (*I said that you can; The man that you saw*), and not for the demonstrative pronoun or adjective (*Put that down; I like that colour*).

Another common word with a reduced form is *our*, but this is derived through smoothing rather than vowel reduction.

Other words that have weak forms in many varieties of English include *your* (weakly pronounced as [jə], or [jɚ] in rhotic accents), and *my* (pronounced [mɨ] or [mi]). These are sometimes given the eye dialect spellings *yer* and *me*.

In highly formal registers with exaggeratedly careful enunciation, weak forms may be avoided. An example is singing, where strong forms may be used almost exclusively, apart (normally) from *a*, although weak forms may be used more frequently as tempo increases and note-values shorten.

The vowel reduction in weak forms may be accompanied by other sound changes, such as h-dropping, consonant elision, and assimilation. For example, *and* may reduce to [ən] or just the syllabic consonant [n], or [ŋ] by assimilation with a following velar, as in *lock and key*. Compare also definite article reduction.

Synchronically, ˈ*em*[əm] functions as a weak form of *them*, though historically it is derived from a different pronoun, the Old English *hem*.

The homonymy resulting from the use of some of the weak forms can lead to confusion in writing; the identity of the weak forms of *have* and *of* sometimes leads to misspellings such as "would of", "could of", etc. for *would have*, *could have*, etc.

English weak forms are distinct from the clitic forms found in some languages, which are words fused with an adjacent word, as in Italian*mangiarla*, "to-eat-it".

Chapter 5

English Orthography

English orthography is the system of writing conventions used to represent spoken English in written form that allows readers to connect spelling to sound to meaning.

Like the orthography of most world languages, English orthography has a broad degree of standardisation. This standardization began to develop when movable type spread to England in the late 15th century. However, unlike with most languages, there are multiple ways to spell nearly every phoneme (sound), and most letters also have multiple pronunciations depending on their position in a word and the context.

This is partly due to the large number of words that have been borrowed from a large number of other languages throughout the history of English, without successful attempts at complete spelling reforms, and partly due to accidents of history, such as some of the earliest mass-produced English publications being typeset by highly trained, multilingual printing compositors, who occasionally used a spelling pattern more typical for another language. For example, the word *ghost* was previously spelled *gast* in English, until the Flemish spelling pattern was unintentionally substituted, and happened to be accepted. Most of the spelling conventions in Modern English were derived from the phonetic spelling of a variety of Middle English, and generally do not reflect the sound changes that have occurred since the late 15th century (such as the Great Vowel Shift). As a result of this, many words are spelled the

way that they were pronounced more than 600 years ago, instead of being spelled like they are pronounced in the 21st century.

Despite the various English dialects spoken from country to country and within different regions of the same country, there are only slight regional variations in English orthography, the two most recognised variations being British and American spelling, and its overall uniformity helps facilitate international communication. On the other hand, it also adds to the discrepancy between the way English is written and spoken in any given location.

Function of the letters

Phonemic representation

Letters in English orthography usually represent a particular sound (phoneme). For example, the word *cat*/ˈkæt/ consists of three letters ⟨c⟩, ⟨a⟩, and ⟨t⟩, in which ⟨c⟩ represents the sound /k/, ⟨a⟩ the sound /æ/, and ⟨t⟩ the sound /t/.

Sequences of letters may perform this role as well as single letters. Thus, in the word *ship* (pronounced /ˈʃɪp/), the digraph⟨sh⟩ (two letters) represents the sound /ʃ/. In the word *ditch*, the trigraph⟨tch⟩ represents the sound /tʃ/.

Less commonly, a single letter can represent multiple successive sounds. The most common example is the letter ⟨x⟩, which normally represents the consonant cluster /ks/ (for example, in the word *six*, pronounced /ˈsɪks/).

The same letter (or sequence of letters) may be pronounced in different ways when it occurs in different positions within a word. For instance, the digraph ⟨gh⟩ represents the sound /f/ at the end of some words, such as *rough*/ˈrʌf/, though not in others (*though*/ˈðoʊ/). At the beginning of syllables (i.e. the syllable onset), the digraph ⟨gh⟩ is pronounced /ɡ/, as in the word *ghost* (pronounced /ˈɡoʊst/). Conversely, the digraph ⟨gh⟩is never pronounced/f/ in syllable onsets other than in inflected forms, and is almost never pronounced /ɡ/ in syllable codas (the proper name *Pittsburgh* is an exception).

Some words contain silent letters, which do not represent any sound in modern English pronunciation. Examples include the ⟨b⟩ in *doubt*, *debt*, *dumb*, etc., the ⟨p⟩ in *psychology* and *pneumatic*, ⟨gh⟩ as mentioned above in numerous words such as *though*, *daughter*, *night*, *brought*, and the commonly encountered silent ⟨e⟩ (discussed further below).

Word origin

Another type of spelling characteristic is related to word origin. For example, when representing a vowel, the letter ⟨y⟩ represents the sound /ɪ/ in some words borrowed from Greek (reflecting an original upsilon), whereas the letter usually representing this sound in non-Greek words is the letter ⟨i⟩. Thus, the word *myth*/ˈmɪθ/ is of Greek origin, while *pith*/ˈpɪθ/ is a Germanic word.

Other examples include ⟨ph⟩ pronounced /f/ (which is usually spelled ⟨f⟩), and ⟨ch⟩ pronounced /k/ (which is usually spelled ⟨c⟩ or ⟨k⟩) – the use of these spellings for these sounds often mark words that have been borrowed from Greek.

Some researchers, such as Brengelman (1970), have suggested that, in addition to this marking of word origin, these spellings indicate a more formal level of style or register in a given text, although Rollings (2004) finds this point to be exaggerated as there would be many exceptions where a word with one of these spellings, such as ⟨ph⟩ for /f/ (like *telephone*), could occur in an informal text.

Homophone differentiation

Spelling may also be useful to distinguish between homophones (words with the same pronunciation but different meanings), although in most cases the reason for the difference is historical and was not introduced for the purpose of making a distinction.

For example, the words *heir* and *air* are pronounced identically in most dialects, but in writing they are distinguished from each other by their different spellings.

Another example is the pair of homophones *pain* and *pane*, where both are pronounced /peɪn/ but have two different spellings of the vowel /eɪ/.

Often this is because of the historical pronunciation of each word where, over time, two separate sounds become the same but the different spellings remain:*pain* used to be pronounced as /pain/, with a diphthong, and *pane* as /peːn/, but the diphthong /ai/ merged with the long vowel /eː/ in *pane*, making *pain* and *pane* homophones (*pane–pain* merger). Later /eː/ became a diphthong /eɪ/.

In written language, this may help to resolve potential ambiguities that would arise otherwise (cf. *He's breaking the car* vs. *He's braking the car*).

Nevertheless, many homophones remain that are unresolved by spelling (for example, the word *bay* has at least five fundamentally different meanings).

Marking sound changes in other letters

Some letters in English provide information about the pronunciation of *other* letters in the word. Rollings (2004) uses the term "markers" for such letters. Letters may mark different types of information.

For instance, the letter ⟨e⟩ in the word *cottage*/ˈkɒtɪdʒ/ indicates that the preceding ⟨g⟩ is pronounced /dʒ/, rather than the more common value of ⟨g⟩ in word-final position as the sound /ɡ/, such as in *tag*/ˈtæɡ/.

The letter ⟨e⟩ also often marks an altered pronunciation of a preceding vowel. In the pair *ban* and *bane*, the ⟨a⟩ of *ban* has the value /æ/, whereas the ⟨a⟩ of *bane*is marked by the ⟨e⟩ as having the value /eɪ/. In this context, the ⟨e⟩ is not pronounced, and is referred to as "silent e".

A single letter may even fill multiple pronunciation-marking roles simultaneously. For example, in the word *wage*, the ⟨e⟩ marks not only the change of the ⟨a⟩ from /æ/ to /eɪ/, but also of the ⟨g⟩ from /ɡ/ to /dʒ/. In the word *vague*, the ⟨e⟩ marks the long a sound, but the ⟨u⟩ keeps the g hard rather than soft.

Doubled consonants usually indicate that the preceding vowel is pronounced short. For example, the doubled ⟨t⟩ in *latter* indicates that the ⟨a⟩ is pronounced /æ/, while the single ⟨t⟩ of *later* gives /eɪ/. Doubled consonants only indicate any lengthening or gemination of the consonant sound itself when they come from different morphemes, as with the ⟨nn⟩ in *unnatural = un+natural.*

Multiple functionality

A given letter (or letters) may have dual functions. For example, the letter ⟨i⟩ in the word *cinema* has a sound-representing function (representing the sound /ɪ/) and a pronunciation-marking function (marking the ⟨c⟩ as having the value /s/ opposed to the value /k/).

Underlying representation

Like many other alphabetic orthographies, English spelling does not represent non-contrastive phonetic sounds (that is, minor differences in pronunciation which are not used to distinguish between different words).

Although the letter ⟨t⟩ is pronounced by some speakers with aspiration[tʰ] at the beginning of words, this is never indicated in the spelling, and, indeed, this phonetic detail is probably not noticeable to the average native speaker not trained in phonetics.

However, unlike some orthographies, English orthography often represents a very abstract underlying representation (or morphophonemic form) of English words.

[T]he postulated underlying forms are systematically related to the conventional orthography … and are, as is well known, related to the underlying forms of a much earlier historical stage of the language. There has, in other words, been little change in lexical representation since Middle English, and, consequently, we would expect … that lexical representation would differ very little from dialect to dialect in Modern English … [and] that conventional orthography is probably fairly close to optimal for all modern English dialects, as well as for the attested dialects of the past several hundred years.

In these cases, a given morpheme (i.e. a component of a word) has a fixed spelling even though it is pronounced differently in different words. An example is the past tensesuffix -⟨ed⟩, which may be pronounced variously as /t/, /d/, or /ɪd/ (for example, *dip*/ˈdɪp/, *dipped*/ˈdɪpt/, *boom*/ˈbuːm/, *boomed*/ˈbuːmd/, *loot*/ˈluːt/, *looted*/ˈluːtɪd/). As it happens, these different pronunciations of -⟨ed⟩can be predicted by a few phonological rules, but that is not the reason why its spelling is fixed.

Other examples of this type are the -⟨ity⟩ suffix (as in *agile* vs *agility*, *acid* vs *acidity*, *divine* vs *divinity*, *sane* vs *sanity*). See also: Trisyllabic laxing.

Another such class of words includes *sign*/ˈsaɪn/ and *bomb*/ˈbɒm/ with "silent" letters ⟨g⟩ and ⟨b⟩, respectively. However, in the related words *signature* and *bombard* these letters are pronounced /ˈsɪgnətʃər/ and /bɒmˈbɑːrd/, respectively. Here it could be argued that the underlying representation of *sign* and *bomb* is |saɪgn| and |bɒmb|, in which the underlying |g| and |b| are only pronounced in the surface forms when followed by certain suffixes (-⟨ature⟩, -

⟨ard⟩). Otherwise, the |ɡ| and |b| are not realised in the surface pronunciation (e.g. when standing alone, or when followed by suffixes like -⟨ing⟩ or -⟨er⟩). In these cases, the orthography indicates the underlying consonants that are present in certain words but are absent in other related words.

Other examples include the ⟨t⟩ in *fast*/ˈfɑːst/ and *fasten*/ˈfɑːsən/, and the ⟨h⟩ in *heir*/ˈɛər/ and *inherit*/ɪnˈhɛrɪt/.

Another example includes words like *mean*/ˈmiːn/ and *meant*/ˈmɛnt/. Here the vowel spelling ⟨ea⟩is pronounced differently in the two related words. Thus, again the orthography uses only a single spelling that corresponds to the single morphemic form rather than to the surface phonological form.

English orthography does not always provide an underlying representation; sometimes it provides an intermediate representation between the underlying form and the surface pronunciation.

This is the case with the spelling of the regular plural morpheme, which is written as either -⟨s⟩ (as in *tick, ticks* and *mite, mites*) or -⟨es⟩ (as in *box, boxes*). Here the spelling -⟨s⟩ is pronounced either /s/ or /z/ (depending on the environment, e.g. *ticks*/ˈtɪks/ and *pigs*/ˈpɪɡz/) while -⟨es⟩ is usually pronounced /ɨz/ (e.g. *boxes*/ˈbɒksɨz/). Thus, there are two different spellings that correspond to the single underlying representation |z| of the plural suffix and the three surface forms. The spelling indicates the insertion of /ɨ/ before the /z/ in the spelling -⟨es⟩, but does not indicate the devoiced/s/ distinctly from the unaffected /z/ in the spelling -⟨s⟩.

The abstract representation of words as indicated by the orthography can be considered advantageous since it makes etymological relationships more apparent to English readers. This makes writing English more complex, but arguably makes reading English more efficient. However, very abstract underlying representations, such as that of Chomsky & Halle (1968) or of underspecification theories, are sometimes considered too abstract to accurately reflect the communicative competence of native speakers. Followers of these arguments believe the less abstract surface forms are more "psychologically real" and thus more useful in terms of pedagogy.

Diacritics

English has some words that can be written with accent marks. These words have mostly been imported from other languages, usually French. As imported words become increasingly naturalised, there is an increasing tendency to omit the accent marks, even in formal writing. For example, words such as *rôle* and *hôtel*were first seen with accents when they were borrowed into English, but now the accent is almost never used.

The words were originally considered foreign – and some people considered that English alternatives were preferable – but today their foreign origin is largely forgotten. Words most likely to retain the accent are those atypical of English morphology and therefore still perceived as slightly foreign. For example, *café* and *pâté* both have a pronounced final *e*, which would otherwise be silent under the normal English pronunciation rules. However*café* is now sometimes facetiously pronounced "caff", while in *pâté*, the acute accent is helpful to distinguish

it from *pate.* Further examples of words sometimes retaining diacritics when used in English are: Ångström (partly because the scientific symbol for this unit of measurement is "Å"), *appliqué*, *attaché*, *blasé*, *bric-à-brac*, *Brötchen*, *cliché*, *crème*, *crêpe*, *façade*, *fiancé(e)*, *flambé*, *naïve*, *naïveté*, *né(e)*, *papier-mâché*, *passé*, *piñata*, *protégé*, *résumé*, *risqué*, *über-*, *voilà.* Italics, with appropriate accents, are generally applied to foreign terms that are uncommonly used in or have not been assimilated into English: for example, *adiós, crème brûlée, pièce de résistance, raison d'être, über, vis-à-vis* and *belles-lettres.*

It was formerly common in American English to use a diaeresis mark to indicate a hiatus: for example, *coöperate*, *daïs*, *reëlect.* *The New Yorker* and *Technology Review* magazines still use it for this purpose, even though it is increasingly rare in modern English. Nowadays the diaeresis is normally left out (*cooperate*), or a hyphen is used (*co-operate*) if the hiatus is between two morphemes in a compound word. It is, however, still common in monomorphemic loanwords such as *naïve* and *Noël.*

Written accents are also used occasionally in poetry and scripts for dramatic performances to indicate that a certain normally unstressed syllable in a word should be stressed for dramatic effect, or to keep with the metre of the poetry. This use is frequently seen in archaic and pseudoarchaic writings with the *-ed* suffix, to indicate that the *e* should be fully pronounced, as with *cursèd.*

The acute and grave accents are occasionally used in poetry and lyrics: the acute to indicate stress overtly where it might

be ambiguous (rébel vs. rebél) or nonstandard for metrical reasons (caléndar); the grave to indicate that an ordinarily silent or elided syllable is pronounced (warnèd, parlìament).

Ligatures

In certain older texts (typically British), the use of the ligatures æ and œ is common in words such as *archæology*, *diarrhœa*, and *encyclopædia*. Such words have Latin or Greek origin. Nowadays, the ligatures have been generally replaced in British English by the separated digraph*ae* and *oe* (*encyclopaedia*, *diarrhoea*); but usually *economy*, *ecology*, and in American English by *e* (*encyclopedia*, *diarrhea*; but usually *paean*, *amoeba*, *oedipal*, *Caesar*). In some cases, usage may vary; for instance, both *encyclopedia* and *encyclopaedia* are current in the UK.

Phonic irregularities

Partly because English has never had any official regulating authority for spelling, such as the Spanish *Real Academia Española*, the French *Académie française*, and the German *Rat für deutsche Rechtschreibung*, English spelling, compared to many other languages, is quite irregular and complex. Although French, among other languages, presents a similar degree of difficulty when *encoding* (writing), English is more difficult when *decoding* (reading), as there are clearly many more possible pronunciations of a group of letters. For example, in French, the /u/ sound (as in "food", but short), can be spelled *ou*, *ous*, *out*, or *oux* (*ou*, *nous*, *tout*, *choux*), but the pronunciation of each of those sequences is always the

same. In English, the /uː/ sound can be spelled in up to 18 different ways (see the Sound-to-spelling correspondences section below), including *oo*, *u*, *ui*, *ue*, *o*, *oe*, *ou*, *ough*, and *ew* (*food*, *truth*, *fruit*, *blues*, *to*, *shoe*, *group*, *through*, *grew*), but all of these have other pronunciations as well (e.g. as in *flood*, *trust*, *build*, *bluest*, *go*, *hoe*, *grout*, *rough*, *sew*). The Spelling-to-sound correspondences section below presents a summary of pronunciation variations. Thus, in unfamiliar words and proper nouns the pronunciation of some sequences, *ough* being the prime example, is unpredictable to even educated native English speakers.

Spelling irregularities

Attempts to regularise or reform the spelling of English have usually failed. However, Noah Webster popularised more phonetic spellings in the United States, such as *flavor* for British *flavour*, *fiber* for *fibre*, *defense* for *defence*, *analyze* for *analyse*, *catalog* for *catalogue* and so forth. These spellings already existed as alternatives, but Webster's dictionaries helped make them standard in the US. See American and British English spelling differences for details.

Besides the quirks the English spelling system has inherited from its past, there are other idiosyncrasies in spelling that make it tricky to learn. English contains, depending on dialect, 24–27 separate consonantphonemes and 13–20 vowels. However, there are only 26 letters in the modern English alphabet, so there is not a one-to-one correspondence between letters and sounds. Many sounds are spelled using different letters or multiple letters, and for those words whose pronunciation is predictable from the spelling, the sounds

denoted by the letters depend on the surrounding letters. For example, the digraph*th* represents two different sounds (the voiced dental fricative and the voiceless dental fricative) (see Pronunciation of English *th*), and the voiceless alveolar sibilant can be represented by the letters *s* and *c*.

It is, however, not the shortage of letters which makes English spelling irregular. Its irregularities are caused mainly by the use of many different spellings for some of its sounds, such as the sounds /uː/, /iː/ and /oʊ/ (*too*, tr*ue*, sh*oe*, *fl*ew, thr*ough*; sl*ee*ve, l*ea*ve, *e*ven, s*ei*ze, s*ie*ge; st*o*le, c*oa*l, b*ow*l, r*o*ll, *o*ld, m*ou*ld), and the use of identical sequences for spelling different sounds (*ove*r, *ove*n, m*ove*).

Furthermore, English no longer makes any attempt to anglicise the spellings of loanwords, but preserves the foreign spellings, even when they employ exotic conventions like the Polish*cz* in *Czech* (rather than **Check*) or the Norwegian*fj* in *fjord* (although *fiord* was formerly the most common spelling). In early Middle English, until roughly 1400, most imports from French were respelled according to English rules (e.g. *bataille–battle*, *bouton–button*, but not *double*, or *trouble*).

Instead of loans being respelled to conform to English spelling standards, sometimes the pronunciation changes as a result of pressure from the spelling. One example of this is the word *ski*, which was adopted from Norwegian in the mid-18th century, although it did not become common until 1900. It used to be pronounced/ʃiː/, which is similar to the Norwegian pronunciation, but the increasing popularity of the sport after the middle of the 20th century helped the /skiː/ pronunciation replace it.

There was also a period when the spelling of a small number of words was altered to make them conform to their perceived etymological origins. For example, the letter *b*was added to *debt* (originally *dette*) in an attempt to link it to the Latin *debitum*, and the letter *s* in *island* is a misplaced attempt to link it to Latin *insula* instead of the Old English word *īġland*, which is the true origin of the English word. The letter *p* in *ptarmigan* has no etymological justification whatsoever, only seeking to invoke Greek despite being a Gaelic word.

The spelling of English continues to evolve. Many loanwords come from languages where the pronunciation of vowels corresponds to the way they were pronounced in Old English, which is similar to the Italian or Spanish pronunciation of the vowels, and is the value the vowel symbols [a], [e], [i], [o], and [u] have in the International Phonetic Alphabet.

As a result, there is a somewhat regular system of pronouncing "foreign" words in English, and some borrowed words have had their spelling changed to conform to this system. For example, *Hindu* used to be spelled*Hindoo*, and the name *Maria* used to be pronounced like the name *Mariah*, but was changed to conform to this system.

Commercial advertisers have also had an effect on English spelling. They introduced new or simplified spellings like *lite* instead of *light*, *thru* instead of *through*, *smokey* instead of *smoky* (for "smokey bacon" flavour crisps), and *rucsac* instead of *rucksack*. The spellings of personal names have also been a source of spelling innovations: diminutive versions of women's names that sound the same as men's names have been spelled differently: *Nikki* and *Nicky*, *Toni* and *Tony*, *Jo* and *Joe*. The

differentiation in between names that are spelled differently but have the same phonetic sound may come from modernization or different countries of origin. For example, *Isabelle* and *Isabel* sound the same but are spelled differently; these versions are from France and Spain respectively.

As examples of the idiosyncratic nature of English spelling, the combination *ou* can be pronounced in at least nine different ways: /aʊ/ in *out*, /oʊ/ in *soul*, /uː/ in *soup*, /ʌ/ in *touch*, /ʊ/ in *could*, /ɔː/ in *four*, /ɜː/ in *journal*, /ɒ/ in *cough*, and /ə/ in *famous*. See the section Spelling-to-sound correspondences for a comprehensive treatment. In the other direction, the vowel sound /iː/ in *me* can be spelled in at least 18 or 21 different ways: *b**e** (c**e**d**e**), *sk**i** (mach**i**n**e**), *bologn**a***(GA), *alg**ae***, *qu**ay***, *b**ea**ch*, *b**ee***, *dec**ei**t*, *p**eo**ple*, *k**ey***, *voll**eye**d*, *f**ie**ld (hyg**ie**ne)*, *am**oe**ba*, *cham**oi**s*, *deng**ue***, *beg**ui**ne*, *g**uy**ot*, and *cit**y***. See the section Sound-to-spelling correspondences below. (These examples assume a more-or-less standard non-regional British English accent. Other accents will vary.)

Sometimes everyday speakers of English change a counterintuitive pronunciation simply because it is counterintuitive. Changes like this are not usually seen as "standard", but can become standard if used enough. An example is the word *miniscule*, which still competes with its original spelling of *minuscule*, though this might also be because of analogy with the word *mini*.

History

Inconsistencies and irregularities in English pronunciation and spelling have gradually increased in number throughout the

history of the English language. There are a number of contributing factors. First, gradual changes in pronunciation, such as the Great Vowel Shift, account for a tremendous number of irregularities. Second, relatively recent loan words from other languages generally carry their original spellings, which are often not phonetic in English. The Romanization of languages (e.g., Chinese) using alphabets derived from the Latin alphabet has further complicated this problem, for example when pronouncing Chinese proper names (of people or places).

The regular spelling system of Old English was swept away by the Norman Conquest, and English itself was supplanted in some spheres by Norman French for three centuries, eventually emerging with its spelling much influenced by French. English had also borrowed large numbers of words from French, which naturally kept their French spellings as there was no reason or mechanism to change them. The spelling of Middle English, such as in the writings of Geoffrey Chaucer, is very irregular and inconsistent, with the same word being spelled in different ways, sometimes even in the same sentence. However, these were generally much better guides to the then-pronunciation than modern English spelling is.

For example, the sound /ʌ/, normally written *u*, is spelled with an *o* in *son*, *love*, *come*, etc., due to Norman spelling conventions which prohibited writing *u* before *v*, *m*, *n* due to the graphical confusion that would result. (*v*, *u*, *n* were identically written with two minims in Norman handwriting; *w* was written as two *u* letters; *m* was written with three minims, hence *mm* looked like *vun*, *nvu*, *uvu*, etc.). Similarly, spelling conventions also prohibited final *v*. Hence the identical

spellings of the three different vowel sounds in *love*, *grove* and *prove* are due to ambiguity in the Middle English spelling system, not sound change.

In 1417Henry V began using English for official correspondence, which had no standardised spelling, instead of Latin or French which had standardised spelling. For example, for the word *right*, Latin had one spelling*, rectus;* Old French as used in English law had six spellings; Middle English had 77 spellings. English, now used as the official replacement language for Latin and French, motivated writers to standardise spellings, an effort which lasted about 500 years.

There was also a series of linguistic sound changes towards the end of this period, including the Great Vowel Shift, which resulted in the *i* in *mine*, for example, changing from a pure vowel to a diphthong.

These changes for the most part did not detract from the rule-governed nature of the spelling system; but in some cases they introduced confusing inconsistencies, like the well-known example of the many pronunciations of *ough* (*rough*, *through*, *though*, *trough*, *plough*, etc.). Most of these changes happened before the arrival of printing in England. However, the arrival of the modern printing press in 1476 froze the current system, rather than providing the impetus for a realignment of spelling with pronunciation. Furthermore, it introduced further inconsistencies, partly because of the use of typesetters trained abroad, particularly in the Low Countries. For example, the *h* in *ghost* was influenced by Flemish Dutch. The addition and deletion of a silent *e* at the ends of words was also sometimes used to make the right-hand margin line up more

neatly. By the time dictionarieswere introduced in the mid-17th century, the spelling system of English had started to stabilise. By the 19th century, most words had set spellings, though it took some time before they diffused throughout the English-speaking world. In *The Mill on the Floss* (1860), English novelist George Eliot satirised the attitude of the English rural gentry of the 1820s towards orthography:

Mr. Tulliver did not willingly write a letter, and found the relation between spoken and written language, briefly known as spelling, one of the most puzzling things in this puzzling world. Nevertheless, like all fervid writing, the task was done in less time than usual, and if the spelling differed from Mrs. Glegg's,–why, she belonged, like himself, to a generation with whom spelling was a matter of private judgment.

The modern English spelling system, with its national variants, spread together with the expansion of public education later in the 19th century.

"Ough" words

The most notorious group of letters in the English language, the *ough* tetragraph, can be pronounced in at least ten different ways, six of which are illustrated in the construct, *Though the tough cough and hiccough plough him through*, which is quoted by Robert A. Heinlein in *The Door into Summer* to illustrate the difficulties facing automated speech transcription and reading. The "ough" tetragraph, traditionally representing a pronunciation of roughly /ʊx/, is in fact a word in its own right, though rarely known or used: an exclamation of disgust similar to *ugh*. The following are typical

pronunciations of this string of letters throughout English dialects of the world:

- /oʊ/ (as in t*oe*) for *though* and *dough*
- /ʌf/ (as in c*uff*) for *tough*, *rough*, *enough*, and the name *Hough*
- /ɒf/ (as in *off*) for *trough*, *cough*, and *Gough*
- /uː/ (as in bl*ue*) for *through*
- /ɔː/ (as in c*augh*t) for *thought*, *ought*, *sought*, *nought*, *brought*, etc.
- /ə/ (as in comm*a*) for *thorough*, *borough*, and names ending in -*borough*; however, American English pronounces this as /oʊ/
- /aʊ/ (as in c*ow*) as in *bough*, *sough*, *drought*, *plough* (*plow* in North America), *doughty*, and the names *Slough* and *Doughty*

The following pronunciations are found in uncommon single words:

- *hough*: /ɒk/ (more commonly spelled "hock" now)
- *hiccough* (a now-uncommon variant of *hiccup*): /ʌp/ as in *up* (unique)
- *lough*: /ɒx/ with a velar fricative like the **ch** in *loch*, of which *lough* is an anglicised spelling

The place name Loughborough uses two different pronunciations of *ough*: the first *ough* has the sound as in c*uff* and the second rhymes with *thorough*.

For instance, the letter *a* can represent the lax vowel /æ/, tense /eɪ/, heavy /ɑː/, or (often allophonically) [ɛə] before |r|.

Heavy and tense-r vowels are the respective lax and tense counterparts followed by the letter *r*.

Tense vowels are distinguished from lax vowels with a "silent" *e* letter that is added at the end of words. Thus, the letter *a* in *hat* is lax /æ/, but when the letter *e* is added in the word *hate* the letter *a* is tense /eɪ/. Similarly, heavy and tense-r vowels pattern together: the letters *ar* in *car* are heavy /ɑːr/, the letters *ar* followed by silent *e* in the word *care* are /ɛər/. The letter *u* represents two different vowel patterns, one being /ʌ/, /juː/, /ə/, /jʊ/, the other /ʊ/, /uː/, /ʊ/. There is no distinction between heavy and tense-r vowels with the letter *o*, and the letter *u* in the /ʊ-uː-ʊ/ pattern does not have a heavy vowel member.

Besides silent *e*, another strategy for indicating tense and tense-r vowels, is the addition of another orthographic vowel forming a digraph. In this case, the first vowel is usually the main vowel while the second vowel is the "marking" vowel. For example, the word *man* has a lax *a* pronounced /æ/, but with the addition of *i* (as the digraph *ai*) in the word *main* the *a* is marked as tense and pronounced /eɪ/. These two strategies produce words that are spelled differently but pronounced identically, as in *mane* (silent *e* strategy), *main* (digraph strategy) and *Maine* (both strategies). The use of two different strategies relates to the function of distinguishing between words that would otherwise be homonyms.

Besides the 20 basic vowel spellings, Rollings (2004) has a reduced vowel category (representing the sounds /ə, ɪ/) and a miscellaneous category (representing the sounds /ɔɪ, aʊ, aɪ/ and /j/+V, /w/+V, V+V).

Combinations of vowel letters

To reduce dialectal difficulties, the sound values given here correspond to the conventions at Help:IPA/English. This table includes H, W and Y when they represent vowel sounds. If no information is given, it is assumed that the vowel is in a stressed syllable.

Deriving the pronunciation of an English word from its spelling requires not only a careful knowledge of the rules given below (many of which are not explicitly known even by native speakers: speakers merely learn the spelling of a word along with its pronunciation) and their many exceptions, but also:

- a knowledge of which syllables are stressed and which are unstressed (not derivable from the spelling: compare *hallow* and *allow*)
- which combinations of vowels represent monosyllables and which represent disyllables (ditto: compare *waive* and *naive*, *creature* and *creator*)

Chapter 6

English-language Spelling Reform

For centuries, there has been a movement to reform the spelling of the English language. It seeks to change English orthography so that it is more consistent, matches pronunciation better, and follows the alphabetic principle. Common motives for spelling reform include quicker learning, cheaper learning, and making English more useful as an international auxiliary language.

Reform proposals vary in terms of the depth of the linguistic changes and by their implementations. In terms of writing systems, most spelling reform proposals are moderate; they use the traditional English alphabet, try to maintain the familiar shapes of words, and try to maintain common conventions (such as silent e).

More radical proposals involve adding or removing letters or symbols, or even creating new alphabets. Some reformers prefer a gradual change implemented in stages, while others favor an immediate and total reform for all.

Some spelling reform proposals have been adopted partially or temporarily. Many of the spellings preferred by Noah Webster have become standard in the United States, but have not been adopted elsewhere (see American and British English spelling differences). Harry Lindgren's proposal, SR1, was once popular in Australia. However, spelling reform has rarely attracted widespread public support.

History

Modern English spelling developed from about 1350 onwards, when—after three centuries of Norman French rule—English gradually became the official language of England again, although very different from before 1066, having incorporated many words of French origin (battle, beef, button, etc.). Early writers of this new English, such as Geoffrey Chaucer, gave it a fairly consistent spelling system, but this was soon diluted by Chancery clerks who re-spelled words based on French orthography. English spelling consistency was dealt a further blow when William Caxton brought the printing press to London in 1476. Having lived in mainland Europe for the preceding 30 years, his grasp of the English spelling system had become uncertain. The Belgian assistants whom he brought to help him set up his business had an even poorer command of it.

As printing developed, printers began to develop individual preferences or "house styles". Furthermore, typesetters were paid by the line and were fond of making words longer. However, the biggest change in English spelling consistency occurred between 1525, when William Tyndale first translated the New Testament, and 1539, when King Henry VIII legalized the printing of English Bibles in England. The many editions of these Bibles were all printed outside England by people who spoke little or no English. They often changed spellings to match their Dutch orthography. Examples include the silent *h* in *ghost* (to match Dutch *gheest*, which later became *geest*), *aghast*, *ghastly* and *gherkin*. The silent *h* in other words—such as *ghospel*, *ghossip* and *ghizzard*—was later removed.

There have been two periods when spelling reform of the English language has attracted particular interest.

16th and 17th centuries

The first of these periods was from the middle of the 16th to the middle of the 17th centuries AD, when a number of publications outlining proposals for reform were published. Some of these proposals were:

- *De recta et emendata linguæ angliæ scriptione* (*On the Rectified and Amended Written English Language*) in 1568 by Sir Thomas Smith, Secretary of State to Edward VI and Elizabeth I.
- *An Orthographie* in 1569 by John Hart, Chester Herald.
- *Booke at Large for the Amendment of English Orthographie* in 1580 by William Bullokar.
- *Logonomia Anglica* in 1621 by Dr. Alexander Gill, headmaster of St Paul's School in London.
- *English Grammar* in 1634 by Charles Butler, vicar of Wootton St Lawrence.

These proposals generally did not attract serious consideration because they were too radical or were based on an insufficient understanding of the phonology of English.

However, more conservative proposals were more successful. James Howell in his *Grammar* of 1662 recommended minor changes to spelling, such as changing *logique* to *logic*, *warre* to *war*, *sinne* to *sin*, *toune* to *town* and *tru* to *true*. Many of these spellings are now in general use.

From the 16th century AD onward, English writers who were scholars of Greek and Latin literature tried to link English words to their Graeco-Latin counterparts. They did this by adding silent letters to make the real or imagined links more obvious. Thus *det* became *debt* (to link it to Latin *debitum*), *dout* became *doubt* (to link it to Latin *dubitare*), *sissors* became *scissors* and *sithe* became *scythe* (as they were wrongly thought to come from Latin *scindere*), *iland* became *island* (as it was wrongly thought to come from Latin *insula*), *ake* became *ache* (as it was wrongly thought to come from Greek *akhos*), and so forth.

William Shakespeare satirized the disparity between English spelling and pronunciation. In his play *Love's Labour's Lost*, the character Holofernes is "a pedant" who insists that pronunciation should change to match spelling, rather than simply changing spelling to match pronunciation. For example, Holofernes insists that everyone should pronounce the unhistorical *B* in words like *doubt* and *debt*.

19th century

The second period started in the 19th century and appears to coincide with the development of phonetics as a science. In 1806, Noah Webster published his first dictionary, *A Compendious Dictionary of the English Language*. It included an essay on the oddities of modern orthography and his proposals for reform. Many of the spellings he used, such as *color* and *center*, would become hallmarks of American English. In 1807, Webster began compiling an expanded dictionary. It was published in 1828 as *An American Dictionary of the English*

Language. Although it drew some protest, the reformed spellings were gradually adopted throughout the United States.

In 1837, Isaac Pitman published his system of phonetic shorthand, while in 1848 Alexander John Ellis published *A Plea for Phonetic Spelling.* These were proposals for a new phonetic alphabet. Although unsuccessful, they drew widespread interest.

By the 1870s, the philological societies of Great Britain and America chose to consider the matter. After the "International Convention for the Amendment of English Orthography" that was held in Philadelphia in August 1876, societies were founded such as the English Spelling Reform Association and American Spelling Reform Association.

That year, the American Philological Society adopted a list of eleven reformed spellings for immediate use. These were *are→ar, give→giv, have→hav, live→liv, though→tho, through→thru, guard→gard, catalogue→catalog, (in)definite→(in)definit, wished→wisht.* One major American newspaper that began using reformed spellings was the *Chicago Tribune*, whose editor and owner,

Joseph Medill, sat on the Council of the Spelling Reform Association. In 1883, the American Philological Society and American Philological Association worked together to produce 24 spelling reform rules, which were published that year. In 1898, the American National Education Association adopted its own list of 12 words to be used in all writings: *tho, altho, thoro, thorofare, thru, thruout, catalog, decalog, demagog, pedagog, prolog, program.*

20th century onward

The Simplified Spelling Boardwas founded in the United States in 1906. The SSB's original 30 members consisted of authors, professors and dictionary editors. Andrew Carnegie, a founding member, supported the SSB with yearly bequests of more than US$300,000. In April 1906, it published a list of 300 words, which included 157 spellings that were already in common use in American English. In August 1906, the SSB word list was adopted by Theodore Roosevelt, who ordered the Government Printing Office to start using them immediately. However, in December 1906, the U.S. Congress passed a resolution and the old spellings were reintroduced. Nevertheless, some of the spellings survived and are commonly used in American English today, such as *anaemia/anæmia→anemia* and *mould→mold*. Others such as *mixed→mixt* and *scythe→sithe* did not survive. In 1920, the SSB published its *Handbook of Simplified Spelling*, which set forth over 25 spelling reform rules. The handbook noted that every reformed spelling now in general use was originally the overt act of a lone writer, who was followed at first by a small minority. Thus, it encouraged people to "point the way" and "set the example" by using the reformed spellings whenever they could. However, with its main source of funds cut off, the SSB disbanded later that year.

In Britain, spelling reform was promoted from 1908 by the Simplified Spelling Society and attracted a number of prominent supporters. One of these was George Bernard Shaw (author of *Pygmalion*) and much of his considerable willwas left to the cause. Among members of the society, the conditions of his will gave rise to major disagreements, which hindered the development of a single new system.

Between 1934 and 1975, the *Chicago Tribune*, then Chicago's biggest newspaper, used a number of reformed spellings.

Over a two-month spell in 1934, it introduced 80 respelled words, including *tho, thru, thoro, agast, burocrat, frate, harth, herse, iland, rime, staf* and *telegraf.* A March 1934 editorial reported that two-thirds of readers preferred the reformed spellings. Another claimed that "prejudice and competition" was preventing dictionary makers from listing such spellings. Over the next 40 years, however, the newspaper gradually phased out the respelled words. Until the 1950s, Funk & Wagnalls dictionaries listed many reformed spellings, including the SSB's 300, alongside the conventional spellings.

In 1949, a Labour MP, Dr Mont Follick, introduced a private member's bill in the House of Commons, which failed at the second reading. In 1953, he again had the opportunity, and this time it passed the second reading by 65 votes to 53. Because of anticipated opposition from the House of Lords, the bill was withdrawn after assurances from the Minister of Education that research would be undertaken into improving spelling education.

In 1961, this led to James Pitman's Initial Teaching Alphabet, introduced into many British schools in an attempt to improve child literacy. Although it succeeded in its own terms, the advantages were lost when children transferred to conventional spelling. After several decades, the experiment was discontinued.

In his 1969 book *Spelling Reform: A New Approach*, the Australian linguist Harry Lindgren proposed a step-by-step reform. The first, *Spelling Reform step 1* (SR1), called for the

short /ɛ/ sound (as in *bet*) to always be spelled with <e> (for example *friend→frend, head→hed*). This reform had some popularity in Australia.

In 2013, University of Oxford Professor of English Simon Horobin proposed that variety in spelling be acceptable. For example, he believes that it does not matter whether words such as "accommodate" and "tomorrow" are spelled with double letters. This proposal does not fit within the definition of spelling reform used by, for example, Random House Dictionary.

Arguments for reform

It is argued that spelling reform would make it easier to learn to read (decode), to spell, and to pronounce, making it more useful for international communication, reducing educational budgets (reducing literacy teachers, remediation costs, and literacy programs) and/or enabling teachers and learners to spend more time on more important subjects or expanding subjects.

Another argument is the sheer amount of resources that are wasted using the current spelling. For example, Cut Spelling can reduce spelling up to 15%. According to that figure, for every 100 letters being used on a daily basis there are 15 letters being used unnecessarily. That amounts to 15 pages for every 100 pages of a book, or about 1 in 7 trees. This applies to all aspects of daily living including shopping receipts, office documents, newspapers and magazines, and internet traffic. This is taxing on time, energy, money, and other resources.

Advocates note that spelling reforms have taken place already, just slowly and often not in an organized way. There are many words that were once spelled un-phonetically but have since been reformed. For example, *music* was spelled *musick* until the 1880s, and *fantasy* was spelled *phantasy* until the 1920s.

For a time, almost all words with the *-or* ending (such as *error*) were once spelled *-our* (*errour*), and almost all words with the *-er* ending (such as *member*) were once spelled *-re* (*membre*). In American spelling, most of them now use *-or* and *-er*, but in British spelling, only some have been reformed.

In the last 250 years, since Samuel Johnson prescribed how words ought to be spelled, pronunciations of hundreds of thousands of words (as extrapolated from Masha Bell's research on 7000 common words) have gradually changed, and the alphabetic principle in English has gradually been corrupted.

Advocates argue that if we wish to keep English spelling regular, then spelling needs to be amended to account for the changes.

Reduced spelling is currently practiced on informal internet platforms and is common in text messaging.

The way vowel letters are used in English spelling vastly contradicts their usual meanings. For example, ⟨o⟩, expected to represent [o], may stand for [u], while ⟨u⟩, expected to represent [u], may represent [ɐ]. This makes English spelling even less intuitive for foreign learners than it is for native speakers, which is of importance for an international auxiliary language.

Ambiguity

Unlike many other languages, English spelling has never been systematically updated and thus today only partly holds to the alphabetic principle. As an outcome, English spelling is a system of weak rules with many exceptions and ambiguities.

Most phonemes in English can be spelled in more than one way. E.g. the words fear and peer contain the same sound in different spellings. Likewise, many graphemes in English have multiple pronunciations and decodings, such as *ough* in words like *through*, *though*, *thought*, *thorough*, *tough*, *trough*, *plough*, and *cough*.

There are 13 ways of spelling the schwa (the most common of all phonemes in English), 12 ways to spell /ei/ and 11 ways to spell /ɛ/. These kinds of incoherences can be found throughout the English lexicon and they even vary between dialects. Masha Bell has analyzed 7000 common words and found that about 1/2 cause spelling and pronunciation difficulties and about 1/3 cause decoding difficulties.

Such ambiguity is particularly problematic in the case of heteronyms (homographs with different pronunciations that vary with meaning), such as *bow*, *desert*, *live*, *read*, *tear*, *wind*, and *wound*. In reading such words one must consider the context in which they are used, and this increases the difficulty of learning to read and pronounce English.

A closer relationship between phonemes and spellings would eliminate many exceptions and ambiguities, making the language easier and faster to master.

Undoing the changes

Some proposed simplified spellings already exist as standard or variant spellings in old literature. As noted earlier, in the 16th century, some scholars of Greek and Latin literature tried to make English words look more like their Graeco-Latin counterparts, at times even erroneously. They did this by adding silent letters, so *det* became *debt*, *dout* became *doubt*, *sithe* became *scythe*, *iland* became *island*, *ake* became *ache*, and so on. Some spelling reformers propose undoing these changes. Other examples of older spellings that are more phonetic include *frend* for *friend* (as on Shakespeare's grave), *agenst* for *against*, *yeeld* for *yield*, *bild* for *build*, *cort* for *court*, *sted* for *stead*, *delite* for *delight*, *entise* for *entice*, *gost* for *ghost*, *harth* for *hearth*, *rime* for *rhyme*, *sum* for *some*, *tung* for *tongue*, and many others. It was also once common to use *-t* for the ending *-ed* where it is pronounced as such (for example *dropt* for *dropped*). Some of the English language's most celebrated writers and poets have used these spellings and others proposed by today's spelling reformers. Edmund Spenser, for example, used spellings such as *rize, wize* and *advize* in his famous poem *The Faerie Queene*, published in the 1590s.

Redundant letters

The English alphabet has several letters whose characteristic sounds are already represented elsewhere in the alphabet. These include X, which can be realised as "ks", "gz", or z; soft G (/dʒ/), which can be realised as J; hard C (/k/), which can be realised as K; soft C (/s/), which can be realised as S; and Q ("qu", /kw/ or /k/), which can be realised as "kw" (or simply

K in some cases). However, these spellings are usually retained to reflect their often-Latin roots.

Arguments against reform

Spelling reform faces many arguments against the development and implementation of a reformed orthography for English. Public acceptance to spelling reform has been consistently low, at least since the early 19th century, when spelling was codified by the influential English dictionaries of Samuel Johnson (1755) and Noah Webster (1806). The irregular spelling of very common words, such as *is*, *are*, *have*, *done* and *of* makes it difficult to fix them without introducing a noticeable change to the appearance of English text.

English is the only one of the top ten major languages with no associated worldwide regulatory body with the power to promulgate spelling changes.

English is a West Germanic language that has borrowed many words from non-Germanic languages, and the spelling of a word often reflects its origin. This sometimes gives a clue as to the meaning of the word. Even if their pronunciation has strayed from the original pronunciation, the spelling is a record of the phoneme.

The same is true for words of Germanic origin whose current spelling still resembles their cognates in other Germanic languages. Examples include *light*, German *Licht*; *knight*, German *Knecht*; *ocean*, French *océan*; *occasion*, French *occasion.* Critics argue that re-spelling such words could hide those links, although not all spelling reforms necessarily

require significantly re-spelling them. Another criticism is that a reform may favor one dialect or pronunciation over others, creating a standard language. Some words have more than one acceptable pronunciation, regardless of dialect (e.g. *economic*, *either*). Some distinctions in regional accents are still marked in spelling. Examples include the distinguishing of *fern*, *fir* and *fur* that is maintained in Irish and Scottish English or the distinction between *toe* and *tow* that is maintained in a few regional dialects in England and Wales. However, dialectal accents exist even in languages whose spelling is called phonemic, such as Spanish. Some letters have allophonic variation, such as how the letter *a* in *bath* currently stands for both /æ/ and /ɑ/ and speakers pronounce it as per their dialect. Some words are distinguished only by non-phonetic spelling (as in *knight* and *night*).

Spelling reform proposals

Most spelling reforms attempt to improve phonemic representation, but some attempt genuine phonetic spelling, usually by changing the basic English alphabet or making a new one. All spelling reforms aim for greater regularity in spelling.

Using the basic English alphabet

- Cut Spelling
- *Handbook of Simplified Spelling*
- SoundSpel
- Spelling Reform 1 (SR1)
- Wijk's Regularized Inglish

Extending or replacing the basic English alphabet

These proposals seek to eliminate the extensive use of digraphs (such as "ch", "gh", "kn-", "-ng", "ph", "qu", "sh", voiced and voiceless "th", and "wh-") by introducing new letters and/or diacritics.

Each letter would then represent a single sound. In a digraph, the two letters represent not their individual sounds but instead an entirely different and discrete sound, which can lengthen words and lead to mishaps in pronunciation.

Notable proposals include:

- Benjamin Franklin's phonetic alphabet
- Deseret alphabet
- Interspel
- Shavian alphabet (revised version: Quikscript)
- SaypU (Spell As You Pronounce Universally)
- Simpel-Fonetik Method of Writing
- Unifon

Some speakers of non-Latin script languages occasionally write English phonetically in their respective writing systems, which may be perceived as an ad hoc spelling reform by some.

Historical and contemporary advocates of reform

A number of respected and influential people have been active supporters of spelling reform.

- Orm/Orrmin, 12th century Augustine canon monk and eponymous author of the *Ormulum*, in which he stated that, since he dislikes that people are mispronouncing English, he will spell words exactly as they are pronounced, and describes a system whereby vowel length and value are indicated unambiguously. He distinguished short vowels from long by doubling the following consonants, or, where this is not feasible, by marking the short vowels with a superimposed breve accent.

- Thomas Smith, a Secretary of State to Queen Elizabeth I, who published his proposal *De recta et emendata linguæ angliæ scriptione* in 1568.

- William Bullokar was a schoolmaster who published his book *English Grammar* in 1586, an early book on that topic. He published his proposal *Booke at large for the Amendment of English Orthographie* in 1580.

- John Milton, poet.

- John Wilkins, founder member and first secretary of the Royal Society, early proponent of decimalisation and a brother-in-law to Oliver Cromwell.

- Charles Butler, British naturalist and author of the first natural history of bees: *Ðe Feminin' Monarƙi'*, 1634. He proposed that 'men should write altogeđer according to đe sound now generally received,' and espoused a system in which the h in digraphs was replaced with bars.

- James Howell was a documented, successful (if modest) spelling reformer, recommending, in his *Grammar* of 1662, minor spelling changes, such as 'logique' to 'logic', 'warre' to 'war', 'sinne' to 'sin',

'toune' to 'town' and 'tru' to 'true', many of which are now in general use.

- Benjamin Franklin, American innovator and revolutionary, added letters to the Roman alphabet for his own personal solution to the problem of English spelling.

- Samuel Johnson, poet, wit, essayist, biographer, critic and eccentric, broadly credited with the standardisation of English spelling into its pre-current form in his Dictionary of the English Language (1755).

- Noah Webster, author of the first important American dictionary, believed that Americans should adopt simpler spellings where available and recommended it in his 1806 *A Compendious Dictionary of the English Language.*

- Charles Dickens

- Isaac Pitman developed the most widely used system of shorthand, known now as Pitman Shorthand, first proposed in *Stenographic Soundhand* (1837).

- U.S. President Theodore Roosevelt commissioned a committee, the Columbia Spelling Board, to research and recommend simpler spellings and tried to require the U.S. government to adopt them; however, his approach, to assume popular support by executive order, rather than to garner it, was a likely factor in the limited change of the time.

- Alfred Tennyson, 1st Baron Tennyson was a vice-president of the English Spelling Reform Association, precursor to the (Simplified) Spelling Society.

- Charles DarwinFRS, originator of the Theory of Evolution by Natural Selection, was also a vice-

president of the English Spelling Reform Association, his involvement in the subject continued by his physicist grandson of the same name.

- John Lubbock, 1st Baron Avebury, close friend, neighbour and colleague of Charles Darwin, also involved in the Spelling Reform Association.
- H.G. Wells, science fiction writer and one-time Vice President of the London-based Simplified Spelling Society.
- Andrew Carnegie, celebrated philanthropist, donated to spelling reform societies on the US and Britain, and funded the Simplified Spelling Board.
- Daniel Jones, phonetician. professor of phonetics at University College London.
- George Bernard Shaw, playwright, willed part of his estate to fund the creation of a new alphabet now called the Shavian alphabet.
- Ronald Kingsley Read, creator of the Shavian alphabet, Quikscript and Readspel.
- Mark Twain, a founding member of the Simplified Spelling Board.
- Robert Baden-Powell, 1st Baron Baden-Powell
- Upton Sinclair
- Melvil Dewey, inventor of the Dewey Decimal System, wrote published works in simplified spellings and even simplified his own name from *Melville* to *Melvil*.
- Israel Gollancz
- James Pitman, a publisher and ConservativeMember of Parliament, grandson of Isaac Pitman, invented the Initial Teaching Alphabet.
- Charles Galton Darwin, KBE, MC, FRS, grandson of Charles Darwin and director of Britain's National

Physical Laboratory (NPL) in World War II, was also a wartime vice-president of the Simplified Spelling Society.

- Mont Follick, LabourMember of Parliament, linguist (multi-lingual) and author who preceded Pitman in drawing the English spelling reform issue to the attention of Parliament. Favoured replacing w and y with u and i.

- Isaac Asimov

- HRHPrince Philip, Duke of Edinburgh, one-time Patron of the Simplified Spelling Society. Stated that spelling reform should start outside of the UK, and that the lack of progress originates in the discord amongst reformers. However, his abandonment of the cause was coincident with literacy being no longer an issue for his own children, and his less than lukewarm involvement may have ended as a result of the Society's rejection of attempts to 'pull strings' behind the scenes.

- Robert R. McCormick (1880–1955), publisher of the *Chicago Tribune*, employed reformed spelling in his newspaper. The *Tribune* used simplified versions of some words, such as "altho" for "although".

- Edward Rondthaler (1905–2009), commercial actor, chairman of the American Literacy Council and vice-president of the Spelling Society.

- John C. Wells, London-based phonetician, Esperanto teacher and former professor of phonetics at University College London: past President of The English Spelling Society.

- Valerie Yule, a fellow of the Galton Institute, Vice-president of The English Spelling Society and

founder of the Australian Centre for Social Innovations.

- Doug Everingham, doctor, former Australian Labor politician, health minister in the Whitlam government, and author of *Chemical Shorthand for Organic Formulae* (1943), and a proponent of the proposed SR1, which he used in ministerial correspondence.

- Allan Kiisk, professor of engineering, linguist (multi-lingual), author of *Simple Phonetic English Spelling* (2013) and *Simpel-Fonetik Dictionary for International Version of Writing in English* (2012).

- Anatoly Liberman, professor in the Department of German, Scandinavian and Dutch at the University of Minnesota advocates spelling reforms at his weekly column on word origins at the Oxford University Press blog. Current President of the English Spelling Society.