THE CAMBRIDGE ENCYCLOPEDIA OF LANGUAGE

SECOND EDITION

DAVID CRYSTAL

CAMBRIDGE UNIVERSITY PRESS

1 · THE PRESCRIPTIVE TRADITION

At the beginning of any book on language, readers have a distinct advantage over the author. More than in most areas of enquiry, they already 'know' the subject, in the sense that they already speak and read a language. Moreover, because in modern societies linguistic skills are highly valued, many readers will have definite views about the nature of language and how it should function. This is not the usual state of mind of someone who opens an encyclopedia on, say, astronomy, Roman mythology, or physics.

We must therefore begin our investigation by looking at the main opinions and beliefs people already hold about language as a result of the normal processes of education and social development. These views will provide a frame of reference familiar to many readers, and they will also act as a point of departure for the detailed, systematic, and objective study of the subject in the following pages.

AN EMOTIONAL SUBJECT

It is not easy to be systematic and objective about language study. Popular linguistic debate regularly deteriorates into invective and polemic. Language belongs to everyone; so most people feel they have a right to hold an opinion about it. And when opinions differ, emotions can run high. Arguments can flare as easily over minor points of usage as over major policies of linguistic planning and education (§61).

Language, moreover, is a very public behaviour, so that it is easy for different usages to be noted and criticized. No part of society or social behaviour is exempt: linguistic factors influence our judgments of personality, intelligence, social status, educational standards, job aptitude, and many other areas of identity and social survival. As a result, it is easy to hurt, and to be hurt, when language use is unfeelingly attacked.

The American linguist Leonard Bloomfield (1887–1949) discussed this situation in terms of three levels of response people give to language. The 'primary response' is actual usage. 'Secondary responses' are the views we have about language, often expressed in some kind of terminology. 'Tertiary responses' are the feelings which flare up when anyone dares to question these views. Bloomfield tells the story of visiting a doctor who was quite firm in his view that the Amerindian language Chippewa had only a few hundred words (p. 6). When Bloomfield attempted to dispute the point, the doctor turned away and refused to listen. Irrational responses of this kind are unfortunately all too common; but everyone is prone to them – linguist and non-linguist alike.

PRESCRIPTIVISM

In its most general sense, prescriptivism is the view that one variety of language has an inherently higher value than others, and that this ought to be imposed on the whole of the speech community. The view is propounded especially in relation to grammar and vocabulary, and frequently with reference to pronunciation. The variety which is favoured, in this account, is usually a version of the 'standard' written language, especially as encountered in literature, or in the formal spoken language which most closely reflects this style. Adherents to this variety are said to speak or write 'correctly'; deviations from it are said to be 'incorrect'.

All the main European languages have been studied prescriptively, especially in the 18th century approach to the writing of grammars and dictionaries. The aims of these early grammarians were threefold: (a) they wanted to codify the principles of their languages, to show that there was a system beneath the apparent chaos of usage, (b) they wanted a means of settling disputes over usage, (c) they wanted to point out what they felt to be common errors, in order to 'improve' the language. The authoritarian nature of the approach is best characterized by its reliance on 'rules' of grammar. Some usages are 'prescribed', to be learnt and followed accurately; others are 'proscribed', to be avoided. In this early period, there were no half-measures: usage was either right or wrong, and it was the task of the grammarian not simply to record alternatives, but to pronounce judgment upon them.

These attitudes are still with us, and they motivate widespread concern that linguistic standards should be maintained. Nevertheless, there is an alternative point of view that is concerned less with 'standards' than with the *facts* of linguistic usage. This approach is summarized in the statement that it is the task of the grammarian to *describe,* not *prescribe* – to record the facts of linguistic diversity, and not to attempt the impossible tasks of evaluating language variation or halting language change. In the second half of the 18th century, we already find advocates of this view, such as Joseph Priestley, whose *Rudiments of English Grammar* (1761) insists that 'the custom of speaking is the original and only just standard of any language'. Linguistic issues, it is argued, cannot be solved by logic and legislation. And this view has become the tenet of the modern linguistic approach to grammatical analysis.

In our own time, the opposition between 'descriptivists' and 'prescriptivists' has often become extreme,

George Orwell (1903–50)

In *Politics and the English Language* (1947), Orwell lists six rules 'that one can rely on when instinct fails'. These rules were not written with literary or scientific language in mind, but with the everyday need to foster language 'as an instrument for expressing and not for concealing or preventing thought'. In this way, Orwell hoped, it would be possible to halt the decline in the language, which he saw as intimately connected with the 'political chaos' of the time.

1 Never use a metaphor, simile or other figure of speech which you are used to seeing in print.
2 Never use a long word when a short one will do.
3 If it is possible to cut a word out, always cut it out.
4 Never use the passive where you can use the active.
5 Never use a foreign phrase, a scientific word or a jargon word if you can think of an everyday English equivalent.
6 Break any of these rules sooner than say anything outright barbarous.

(See further, p. 382.)

with both sides painting unreal pictures of the other. Descriptive grammarians have been presented as people who do not care about standards, because of the way they see all forms of usage as equally valid. Prescriptive grammarians have been presented as blind adherents to a historical tradition. The opposition has even been presented in quasi-political terms – of radical liberalism vs elitist conservatism.

If these stereotypes are abandoned, we can see that both approaches are important, and have more in common than is often realized – involving a mutual interest in such matters as acceptability, ambiguity, and intelligibility. The descriptive approach is essential because it is the only way in which the competing claims of different standards can be reconciled: when we know the facts of language use, we are in a better position to avoid the idiosyncrasies of private opinions, and to make realistic recommendations about teaching or style. The prescriptive approach provides a focus for the sense of linguistic values which everyone possesses, and which ultimately forms part of our view of social structure, and of our own place within it. After 200 years of dispute, it is perhaps sanguine to expect any immediate rapport to be achieved, but there are some grounds for optimism, now that sociolinguists (p. 414) are beginning to look more seriously at prescriptivism in the context of explaining linguistic attitudes, uses, and beliefs.

Where traditional grammatical rules come from

	Example of a prescriptive rule	Descriptive comment
Latin and Greek The unchanging form of these languages, the high prestige they held in European education, and the undisputed brilliance of classical literature led to their adoption as models of linguistic excellence by grammarians of other languages.	You should say or write *It is I* and not *It is me,* because the verb *be* is followed by the nominative case in Latin, not the accusative.	The Latin rule is not universal. In Arabic, for example, *be* is followed by the accusative. In English, *me* is the educated informal norm; *I* is felt to be very formal. In French, only *moi* is possible (*c'est moi,* etc.)
The written language Writing is more careful, prestigious and permanent than speech, especially in the context of literature. People are therefore often told to speak as they would write.	You should say and write *whom* and not *who,* in such sentences as – *did you speak to?*	*Whom* is common in writing, and in formal styles of speech; but *who* is more acceptable in informal speech. The rules which govern acceptable speech and writing are often very different.
Logic Many people feel that grammar should be judged insofar as it follows the principles of logic. Mathematics, from this viewpoint, is the ideal use of language.	You shouldn't say *I haven't done nothing* because two negatives make a positive.	Here, two negatives do not make a positive, but a more emphatic negative – a construction which is found in many languages (e.g. French, Russian). The example is not acceptable in standard English, but this is the result of social factors, not the dictates of logic.

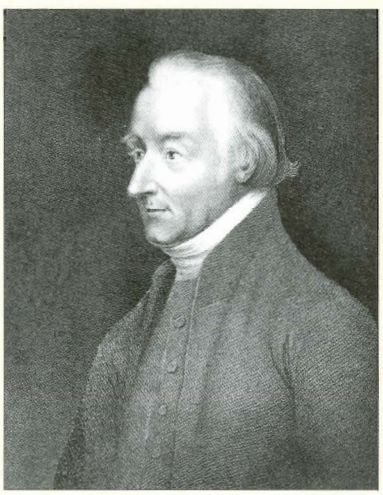
MURRAY'S *GRAMMAR*

One of the most influential grammars of the 18th century was Robert Lowth's *Short Introduction to English Grammar* (1762). This was the inspiration for Lindley Murray's widely used *English Grammar* (1794). Both grammars went through over 20 editions in the decades following publication.

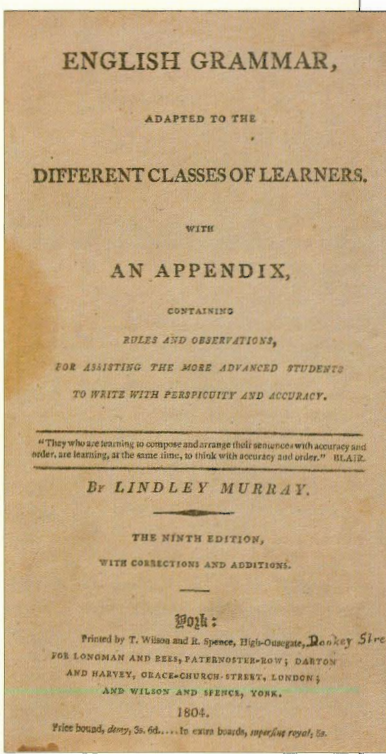
Murray's book had an enormous influence on school practice and popular attitudes, especially in the USA. His alliterative axiom contains several watchwords of prescriptivism: 'Perspicuity requires the qualities of purity, propriety and precision'.

Some of Murray's general linguistic principles were unexceptionable, such as 'Keep clear of double meaning or ambiguity' and 'Avoid unintelligible words or phrases.' But most of his analyses, and the detailed principles of his Appendix, 'Rules and observations for promoting perspicuity in speaking and writing', contain the kind of arbitrary rule and artificial, Latinate analysis which was to fuel two centuries of argument. In Rule 16, for example, we find the negation principle illustrated: 'Two negatives, in English, destroy one another, or are equivalent to an affirmative.'

Murray's rules were widely taught, and formed the basis for much of the linguistic purism still encountered today. However, they were also fiercely attacked. One writer in the *American Journal of Education* (in 1826) compares the grammar to a 'foreign rack on which our simple language has been stretched'. Another (in 1833) insists that grammarians should 'discover' and not 'invent' rules. Long before the advent of modern linguistics, the battle lines of both descriptivism and prescriptivism had been clearly established.



Right: **Lindley Murray (1745–1826)**



ENGLISH GRAMMAR,

ADAPTED TO THE

DIFFERENT CLASSES OF LEARNERS.

WITH

AN APPENDIX,

CONTAINING

RULES AND OBSERVATIONS,

FOR ASSISTING THE MORE ADVANCED STUDENTS

TO WRITE WITH PERSPICUITY AND ACCURACY.

"They who are learning to compose and arrange their sentences with accuracy and order, are learning, at the same time, to think with accuracy and order." BLAIR.

BY LINDLEY MURRAY.

THE NINTH EDITION,

WITH CORRECTIONS AND ADDITIONS.

York:

Printed by T. Wilson and R. Spence, High-Ousegate, Pavkey Stre

FOR LONGMAN AND REES, PATERNOSTER-ROW; DARTON AND HARVEY, GRACE-CHURCH-STREET, LONDON; AND WILSON AND SPENCE, YORK.

1804.

Price bound, demy, 3s. 6d.... In extra boards, imperial royal, 5s.

Above: **Murray's *English Grammar***

THE ACADEMIES

Some countries have felt that the best way to look after a language is to place it in the care of an academy. In Italy, the *Accademia della Crusca* was founded as early as 1582, with the object of purifying the Italian language. In France, in 1635, Cardinal Richelieu established the *Académie française,* which set the pattern for many subsequent bodies. The statutes of the *Académie* define as its principal function:

> to labour with all possible care and diligence to give definite rules to our language, and to render it pure, eloquent, and capable of treating the arts and sciences.

The 40 academicians were drawn from the ranks of the church, nobility, and military – a bias which continues to the present day. The *Académie's* first dictionary appeared in 1694.

Several other academies were founded in the 18th and 19th centuries. The Spanish Academy was founded in 1713 by Philip V, and within 200 years corresponding bodies had been set up in most South American Spanish countries. The Swedish Academy was founded in 1786; the Hungarian in 1830. There are three Arabic academies, in Syria, Iraq, and Egypt. The Hebrew Language Academy was set up more recently, in 1953.

Kippers *sur* toast? Menus like this could be found, with the appropriate language change, in almost any European city. They illustrate the way English has permeated public life, despite the efforts of many countries to stop it. The German post office, for example, insisted for many years that *Fernsprecher* should be used on phone booths, though *Telefon* was far more common in speech; but in 1981 they made the change. In 1975, the French went so far as to pass a law banning the use of English loan words in official contexts, if an equivalent word exists in French (the *loi Bas-Lauriol*): *a corner* (in football) was to be replaced by *jet de coin*, or *collapser* by *s'évanouir*. However, it was a law honoured more in the breach than in the observance; and when a further attempt to impose French in a range of public contexts was made in 1994 (the *loi Toubon*), parts of the proposal were rejected on the grounds that they were contrary to the principle of freedom of speech, and thus against the constitution . Whether one approves or not, the academies seem to be no match for Franglais, Angleutsch, Swedlish, Spanglish, and all the other hybrids which have become so noticeable in recent years (§§55, 61).

In England, a proposal for an academy was made in the 17th century, with the support of such men as John Dryden and Daniel Defoe. In Defoe's view, the reputation of the members of this academy

> would be enough to make them the allowed judges of style and language; and no author would have the impudence to coin without their authority ...There should be no more occasion to search for derivations and constructions, and it would be as criminal then to coin words as money.

In 1712, Jonathan Swift presented his *Proposal for Correcting, Improving and Ascertaining the English Tongue,* in which he complains to the Lord Treasurer of England, the Earl of Oxford, that

> our language is extremely imperfect; that its daily improvements are by no means in proportion to its daily corruptions; that the pretenders to polish and refine it have chiefly multiplied abuses and absurdities; and that in many instances it offends against every part of grammar.

His academy would 'fix our language for ever', for,

> I am of the opinion, it is better a language should not be wholly perfect, than it should be perpetually changing.

The idea received a great deal of support at the time, but nothing was done. And in due course, opposition to the notion grew. It became evident that the French and Italian academies had been unsuccessful in stopping the course of language change. Dr Johnson, in the Preface to his Dictionary, is under no illusion about the futility of an academy, especially in England, where he finds 'the spirit of English liberty' contrary to the whole idea:

> When we see men grow old and die at a certain time one after another, century after century, we laugh at the elixir that promises to prolong life to a thousand years; and with equal justice may the lexicographer be derided, who being able to produce no example of a nation that has preserved their words and phrases from mutability, shall imagine that his dictionary can embalm his language, and secure it from corruption, and decay, that it is in his power to change sublunary nature, or clear the world at once from folly, vanity, and affectation.

From time to time, the idea of an English Academy continues to be voiced, but the response has never been enthusiastic. A similar proposal in the USA was also rejected. By contrast, since the 18th century, there has been an increasing flow of individual grammars, dictionaries, and manuals of style in all parts of the English-speaking world.

Daniel Defoe (1660?–1731)

Jonathan Swift (1667–1745)

LANGUAGE CHANGE

The phenomenon of language change probably attracts more public notice and criticism than any other linguistic issue. There is a widely held belief that change must mean deterioration and decay. Older people

observe the casual speech of the young, and conclude that standards have fallen markedly. They place the blame in various quarters – most often in the schools, where patterns of language education have changed a great deal in recent years (§44), but also in state public broadcasting institutions, where any deviations from traditional norms provide an immediate focus of attack by conservative, linguistically sensitive listeners. The concern can even reach national proportions, as in the widespread reaction in Europe against what is thought of as the 'American' English invasion.

UNFOUNDED PESSIMISM

It is understandable that many people dislike change, but most of the criticism of linguistic change is misconceived. It is widely felt that the contemporary language illustrates the problem at its worst, but this belief is shared by every generation. Moreover, many of the usage issues recur across generations: several of the English controversies which are the focus of current attention can be found in the books and magazines of the 18th and 19th centuries – the debate over *it's me* and *very unique,* for example. In *The Queen's English* (1863), Henry Alford, the Dean of Canterbury, lists a large number of usage issues which worried his contemporaries, and gave them cause to think that the language was rapidly decaying. Most are still with us, with the language not obviously affected. In the mid-19th century, it was predicted that British and American English would be mutually unintelligible within 100 years!

There are indeed cases where linguistic change can lead to problems of unintelligibility, ambiguity, and social division. If change is too rapid, there can be major communication problems, as in contemporary Papua New Guinea – a point which needs to be considered in connection with the field of language planning (§§55, 61). But as a rule, the parts of language which are changing at any given time are tiny, in comparison to the vast, unchanging areas of language. Indeed, it is because change is so infrequent that it is so distinctive and noticeable. Some degree of caution and concern is therefore always desirable, in the interests of maintaining precise and efficient communication; but there are no grounds for the extreme pessimism and conservatism which is so often encountered – and which in English is often summed up in such slogans as 'Let us preserve the tongue that Shakespeare spoke'.

THE INEVITABILITY OF CHANGE

For the most part, language changes because society changes (§10). To stop or control the one requires that we stop or control the other – a task which can succeed to only a very limited extent. Language change is inevitable and rarely predictable, and those who try to plan a language's future waste their time if they think otherwise – time which would be better spent in devising fresh ways of enabling society to cope with the new linguistic forms that accompany each generation. These days, there is in fact a growing recognition of the need to develop a greater linguistic awareness and tolerance of change, especially in a multi-ethnic society. This requires, among other things, that schools have the knowledge and resources to teach a common standard, while recognizing the existence and value of linguistic diversity. Such policies provide a constructive alternative to the emotional attacks which are so commonly made against the development of new words, meanings, pronunciations, and grammatical constructions. But before these policies can be implemented, it is necessary to develop a proper understanding of the inevitability and consequences of linguistic change (§54).

Some people go a stage further, and see change in language as a progression from a simple to a complex state – a view which was common as a consequence of 19th-century evolutionary thinking. But there is no evidence for this view. Languages do not develop, progress, decay, evolve, or act according to any of the metaphors which imply a specific endpoint and level of excellence. They simply change, as society changes. If a language dies out, it does so because its status alters in society, as other cultures and languages take over its role: it does not die because it has 'got too old', or 'become too complicated', as is sometimes maintained. Nor, when languages change, do they move in a predetermined direction. Some are losing inflections; some are gaining them. Some are moving to an order where the verb precedes the object; others to an order where the object precedes the verb. Some languages are losing vowels and gaining consonants; others are doing the opposite. If metaphors must be used to talk about language change, one of the best is that of a system holding itself in a state of equilibrium, while changes take place within it; another is that of the tide, which always and inevitably changes, but never progresses, while it ebbs and flows.

WILLIAM CAXTON

One of the earliest English voices to complain about the problems of linguistic change was William Caxton (1422?–91). He was writing at a time when English had undergone its greatest period of change, which had resulted in a major shift in pronunciation, the almost total loss of Anglo-Saxon inflections, and an enormous influx of new vocabulary, mainly from French:

> And certaynly our language now used varyeth ferre from that whiche was used and spoken whan I was borne... And that comyn Englysshe that is spoken in one shyre varyeth from a nother. In so moche that in my dayes happened that certayne marchauntes were in a shippe in Temyse [Thames] for to have sayled over the see into Zelande, and for lacke of wynde thei taryed atte forlond, and wente to lande for to refreshe them. And one of theym named Sheffelde, a mercer, cam in to an hows and axed for mete, and specyally he axyd after 'eggys'. And the good wyf answerde that she coude speke no Frenshe. And the marchaunt was angry, for he also coude speke no Frenshe, but wold have hadde egges, and she understode hym not. And thenne at last a nother sayd that he wolde have 'eyren'. Then the good wyf sayd that she understod hym wel. Loo! What shold a man in thyse dayes now wryte, 'egges' or 'eyren'? Certaynly, it is harde to playse every man by cause of dyversite & chaunge of langage.
>
> (Preface to *Eneydos*, 1490; modernized punctuation)

Caxton's plaint echoes through the ages, though problems of linguistic change have never been so serious since, with the subsequent standardization of English, and the spread of the written language.

2 · THE EQUALITY OF LANGUAGES

It comes near to stating the obvious that all languages have developed to express the needs of their users, and that in a sense all languages are equal. But this tenet of modern linguistics has often been denied, and still needs to be defended. Part of the problem is that the word 'equal' needs to be used very carefully. We do not know how to quantify language, so as to be able to say whether all languages have the same 'amounts' of grammar, phonology, or semantic structure (§§16, 17, 28). There may indeed be important differences in the structural complexity of language, and this possibility needs to be investigated. But all languages are arguably equal in the sense that there is nothing intrinsically limiting, demeaning, or handicapping about any of them. All languages meet the social and psychological needs of their speakers, are equally deserving of scientific study, and can provide us with valuable information about human nature and society. This view is the foundation on which the whole of the present book is based.

'PRIMITIVE' LANGUAGES

There are, however, several widely held misconceptions about languages which stem from a failure to recognize this view. The most important of these is the idea that there are such things as 'primitive' languages – languages with a simple grammar, a few sounds, and a vocabulary of only a few hundred words, whose speakers have to compensate for their language's deficiencies through gestures. Speakers of 'primitive' languages have often been thought to exist, and there has been a great deal of speculation about where they might live, and what their problems might be. If they relied on gestures, how would they be able to communicate at night? Without abstract terms, how could they possibly develop moral or religious beliefs? In the 19th century, such questions were common, and it was widely thought that it was only a matter of time before explorers would discover a genuinely primitive language.

The fact of the matter is that every culture which has been investigated, no matter how 'primitive' it may be in cultural terms, turns out to have a fully developed language, with a complexity comparable to those of the so-called 'civilized' nations. Anthropologically speaking, the human race can be said to have evolved from primitive to civilized states, but there is no sign of language having gone through the same kind of evolution (§48). There are no 'bronze age' or 'stone age' languages, nor have any language types been discovered which correlate with recognized anthropological groups (pastoral, nomadic, etc.). All languages have a complex grammar: there may be relative simplicity in one respect (e.g. no word-endings), but there seems always to be relative complexity in another (e.g. word-position). People sometimes think of languages such as English as 'having little grammar', because there are few word-endings. But this is once again (§1) the unfortunate influence of Latin, which makes us think of complexity in terms of the inflectional system of that language.

Simplicity and regularity are usually thought to be desirable features of language; but no natural language is simple or wholly regular. All languages have intricate grammatical rules, and all have exceptions to those rules. The nearest we come to real simplicity with

Juanita, a Navaho woman in the 1870s.

The Roman goddess Fortuna, holding a cornucopia and a rudder – an appropriate deity to associate with the uncertain destinies of languages.

SIMPLE SAVAGES?

Edward Sapir was one of the first linguists to attack the myth that primitive people spoke primitive languages. In one study, he compared the grammatical equivalents of the sentence *he will give it* (a stone) *to you* in six Amerindian languages. (Hyphens separate the parts of the Indian sentences, and in the literal translations that follow they join words that are equivalent to a single Indian form. For phonetic symbols, see p. 442.)

Wishram
a-č-i-m-l-ud-a
will he him thee to give will

Takelma
ʔòk-t-xpi-nk
will-give to thee he-or-they-in-future

Southern Paiute
maɣa-vaania-aka-aŋa-'mi
give will visible-thing visible-creature thee

Yana
ba·-ˉa-ma-si-wa-ʔnuma
round-thing away to does-or-will done-unto thou-in-future

Nootka
oʔ-yi·-ʔa·qχ-ʔat-eʔic
that give will done-unto thou-art

Navaho
n-a·-yi-diho-ʔá·l
thee to transitive-marker will round-thing-in-future

Among many fascinating features of these complex grammatical forms, note the level of abstraction introduced by some languages (expressed by *round thing* and *visible*) – quite contrary to the claim that primitive peoples could only talk about concrete objects.

Sapir also gave part of the full Takelma verb paradigm:

ʔokúspi	gives / gave it to you
ʔòspink	will give to you
ʔòspi	can give to you
ʔòspik	evidently gave to you

He points out the similarity to the way the verb varies in Latin – a comparison which many traditional scholars would have considered to verge on blasphemy!

natural languages is in the case of pidgin languages (§55); and the desire for regularity is a major motivation for the development of auxiliary languages (§58). But these are the only exceptions. Similarly, there is no evidence to suggest that some languages are in the long term 'easier for children to learn' than others – though in the short term some linguistic features may be learned at different rates by the children of speakers of different languages (Part VIII).

None of this is to deny the possibility of linguistic differences which correlate with cultural or social features (such as the extent of technological development), but these have not been found; and there is no evidence to suggest that primitive peoples are in any sense 'handicapped' by their language when they are using it within their own community.

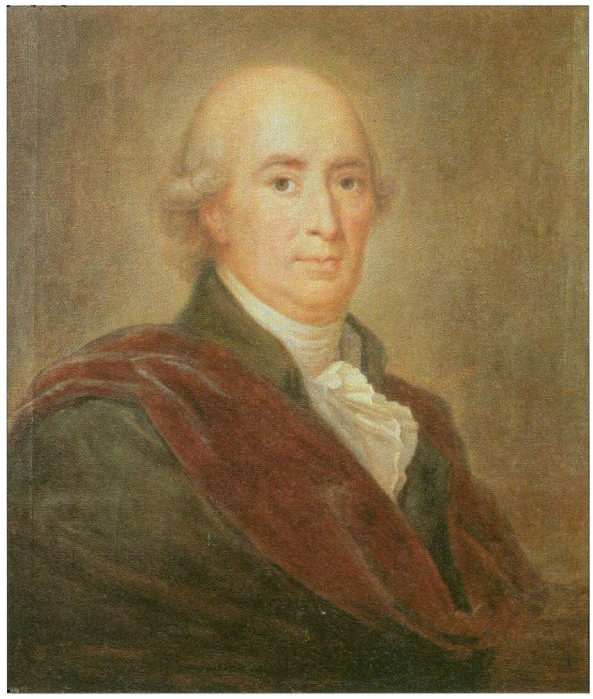
LANGUAGES OF EXCELLENCE

At the other end of the scale from so-called 'primitive' languages are opinions about the 'natural superiority' of certain languages. Latin and Greek were for centuries viewed as models of excellence in western Europe because of the literature and thought which these languages expressed; and the study of modern languages is still influenced by the practices of generations of classical linguistic scholars (p. 378).

The idea that one's own language is superior to others is widespread, but the reasons given for the superiority vary greatly. A language might be viewed as the oldest, or the most logical, or the language of gods, or simply the easiest to pronounce or the best for singing. Arabic speakers, for example, feel that their classical language is the most beautiful and logical, with an incomparable grammatical symmetry and lexical richness. Classical Arabic is strongly identified with religion (p. 388), as the language of the Qur'an is held to provide miraculous evidence of the truth of Islam. From this viewpoint, it would be self-evident that, as God chose Arabic as the vehicle of his revelation to his Prophet, this must be the language used in heaven, and thus must be superior to all others.

However, a similar argument has been applied to several other languages, such as Sanskrit and Classical Hebrew, especially in relation to claims about which language is the oldest (§49). For example, J. G. Becanus (1518–72) argued that German was superior to all other languages. It was the language Adam spoke in Eden, but it was not affected in the Babel event, because the early Germans (the Cimbrians) did not assist in the construction of the tower. God later caused the Old Testament to be translated from the original German (no longer extant) into Hebrew.

There have been many other spurious linguistic evaluations, reflecting the sociopolitical situation of the time. Charles V of Germany (who ruled from 1519 to 1558) is said to have spoken French to men, Italian to women, Spanish to God, and German to horses! The Swedish writer, Andreas Kempe (1622–89), satirized contemporary clerical attitudes in presenting the view that in Paradise Adam spoke Danish, God spoke Swedish, and the serpent spoke French.



Johann Herder (1744–1803)

A LINGUISTIC MYTH

A belief that some languages are intrinsically superior to others is widespread, but it has no basis in linguistic fact. Some languages are of course more useful or prestigious than others, at a given period of history, but this is due to the preeminence of the speakers at that time, and not to any inherent linguistic characteristics. The view of modern linguistics is that a language should not be valued on the basis of the political or economic influence of its speakers. If it were otherwise, we would have to rate the Spanish and Portuguese spoken in the 16th century as somehow 'better' than they are today, and modern American English would be 'better' than British English. Yet when we make such comparisons, we find only a small range of linguistic differences, and nothing to warrant such sweeping conclusions.

At present, it is not possible to rate the excellence of languages in linguistic terms. And it is no less difficult to arrive at an evaluation in aesthetic, philosophical, literary, religious, or cultural terms. How, ultimately, could we compare the merits of Latin and Greek with the proverbial wisdom of Chinese, the extensive oral literature of the Polynesian islands, or the depth of scientific knowledge which has been expressed in English? Perhaps one day some kind of objective linguistic evaluation measure will be devised; but until then, the thesis that some languages are intrinsically better than others has to be denied.

Nationalism In the 18th and 19th centuries, language evaluations were often tied to questions of national identity (§9), especially in Germany, in a school of thought which can be traced back to the view of Johann Herder: 'Has a nation anything more precious than the language of its fathers?' Johann Gottlieb Fichte (1762–1814) praised the German language, and dismissed others, in his *Addresses to the German Nation* (1807), even to the extent of claiming that the native German speaker 'can always be superior to the foreigner and understand him fully, even better than the foreigner understands himself'. But comparable claims were made for French and Spanish; and English was similarly lauded by Thomas Macaulay (1800–59): in his *Minute on Education* (1835), referring to the languages of India, he wrote that English 'stands preeminent even among the languages of the West... It may safely be said that the literature now extant in that language is of greater value than all the literature which three hundred years ago was extant in all the languages of the world together.'

5 · LANGUAGE AND THOUGHT

It seems evident that there is the closest of relationships between language and thought: everyday experience suggests that much of our thinking is facilitated by language (p. 13). But is there identity between the two? Is it possible to think without language? Or does our language dictate the ways in which we are able to think? Such matters have exercised generations of philosophers, psychologists, and linguists, who have uncovered layers of complexity in these apparently straightforward questions. A simple answer is certainly not possible; but at least we can be clear about the main factors which give rise to the complications.

KINDS OF THINKING

Many kinds of behaviour have been referred to as 'thinking', but not all of them require us to posit a relationship with language. Most obviously, there is no suggestion that language is involved in our emotional response to some object or event, such as when we react to a beautiful painting or an unpleasant incident: we may use language to explain our reaction to others, but the emotion itself is 'beyond words'. Nor do people engaged in the creative arts find it essential to think using language: composers, for example, often report that they 'hear' the music they wish to write. Also, our everyday fantasies, day-dreams, and other free associations can all proceed without language.

The thinking which seems to involve language is of a different kind: this is the reasoned thinking which takes place as we work out problems, tell stories, plan strategies, and so on. It has been called 'rational', 'directed', 'logical', or 'propositional' thinking. It involves elements that are both deductive (when we solve problems by using a given set of rules, as in an arithmetical task) and inductive (when we solve problems on the basis of data placed before us, as in working out a travel route). Language seems to be very important for this kind of thinking. The formal properties of language, such as word order and sentence sequencing, constitute the medium in which our connected thoughts can be presented and organized.

INDEPENDENCE OR IDENTITY?

But how close is this relationship between language and thought? It is usual to see this question in terms of two extremes. First, there is the hypothesis that language and thought are totally separate entities, with one being dependent on the other. At the opposite extreme, there is the hypothesis that language and thought are identical – that it is not possible to engage in any rational thinking without using language. The truth seems to lie somewhere between these two positions.

Within the first position, there are plainly two possibilities: language might be dependent upon thought, or thought might be dependent upon language. The traditional view, which is widely held at a popular level, adopts the first of these: people have thoughts, and then they put these thoughts into words. It is summarized in such metaphorical views of language as the 'dress' or 'tool' of thought. The view is well represented in the field of child language acquisition (§38), where children are seen to develop a range of cognitive abilities which precede the learning of language.

The second possibility has also been widely held: the way people use language dictates the lines along which they can think. An expressive summary of this is Shelley's 'He gave men speech, and speech created thought, /Which is the measure of the universe' (*Prometheus Unbound*). This view is also represented in the language acquisition field, in the argument that the child's earliest encounters with language are the main influence on the way concepts are learned. The most influential expression of this position, however, is found in the Sapir–Whorf hypothesis (see facing page).

A third possibility, which is also widely held these days, is that language and thought are interdependent – but this is not to say that they are identical. The identity view (for example, that thought is no more than an internalized vocalization) is no longer common. There are too many exceptions for such a strong position to be maintained: we need think only of the various kinds of mental operations which we can perform without language, such as recalling a sequence of movements in a game or sport, or visualizing the route from home to work. It is also widely recognized that pictorial images and physical models are helpful in problem-solving, and may at times be more efficient than purely verbal representations of a problem.

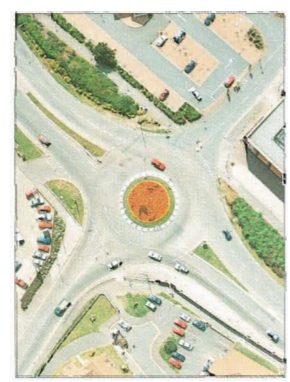
On the other hand, these cases are far outnumbered by those where language does seem to be the main means whereby successful thinking can proceed. To see language and thought as interdependent, then, is to recognize that language is a regular part of the process of thinking, at the same time recognizing that we have to think in order to understand language. It is not a question of one notion taking precedence over the other, but of both notions being essential, if we are to explain behaviour. Once again, people have searched for metaphors to express their views. Language has been likened to the arch of a tunnel; thought, to the tunnel itself. But the complex structure and function of language defies such simple analogies.

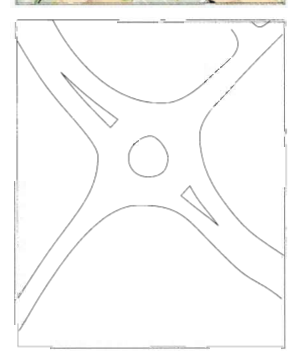
NON-VERBAL AND VERBAL THOUGHT

The two dimensions to rational thinking – linguistic and non-linguistic – can be discovered in a simple experiment, which anyone can perform.

1. Think of where you work. Now visualize the route you follow, as if you were driving along in a car, as you proceed from work to your home. The sequence of visual images which you bring to mind will be largely independent of language.

2. Now imagine you have to explain to a visitor how to reach your house from work. Think out the steps of your explanation, as you would present them, without saying anything aloud. The sequence of ideas will be expressed internally using language.





THE SAPIR–WHORF HYPOTHESIS

The romantic idealism of the late 18th century, as encountered in the views of Johann Herder (1744–1803) and Wilhelm von Humboldt (1762–1835), placed great value on the diversity of the world's languages and cultures. The tradition was taken up by the American linguist and anthropologist Edward Sapir (1884–1939) and his pupil Benjamin Lee Whorf (1897–1941), and resulted in a view about the relation between language and thought which was widely influential in the middle decades of this century.

The 'Sapir–Whorf hypothesis', as it came to be called, combines two principles. The first is known as *linguistic determinism*: it states that language determines the way we think. The second follows from this, and is known as *linguistic relativity*: it states that the distinctions encoded in one language are not found in any other language. In a much-quoted paragraph, Whorf propounds the view as follows:

> We dissect nature along lines laid down by our native languages. The categories and types that we isolate from the world of phenomena we do not find there because they stare every observer in the face; on the contrary, the world is presented in a kaleidoscopic flux of impressions which has to be organized by our minds – and this means largely by the linguistic systems in our minds. We cut nature up, organize it into concepts, and ascribe significances as we do, largely because we are parties to an agreement to organize it in this way – an agreement that holds throughout our speech community and is codified in the patterns of our language. The agreement is, of course, an implicit and unstated one, *but its terms are absolutely obligatory*; we cannot talk at all except by subscribing to the organization and classification of data which the agreement decrees.

Whorf illustrated his view by taking examples from several languages, and in particular from Hopi, an Amerindian language. In Hopi, there is one word (*masa'ytaka*) for everything that flies except birds – which would include insects, aeroplanes and pilots. This seems alien to someone used to thinking in English, but, Whorf argues, it is no stranger than English-speakers having one word for many kinds of snow, in contrast to Eskimo, where there are different words for falling snow, snow on the ground, snow packed hard like ice, slushy snow (cf. English *slush*), and so on. In Aztec, a single word (with different endings) covers an even greater range of English notions – snow, cold, and ice. When more abstract notions are considered (such as time, duration, velocity), the differences become yet more complex: Hopi, for instance, lacks a concept of time seen as a dimension; there are no forms corresponding to English tenses, but there are a series of forms which make it possible to talk about various durations, from the speaker's point of view. It would be very difficult, Whorf argues, for a Hopi and an English physicist to understand each other's thinking, given the major differences between the languages.

Examples such as these made the Sapir–Whorf hypothesis very plausible; but in its strongest form it is unlikely to have any adherents now. The fact that successful translations between languages can be made is a major argument against it, as is the fact that the conceptual uniqueness of a language such as Hopi can nonetheless be explained using English. That there are some conceptual differences between cultures due to language is undeniable, but this is not to say that the differences are so great that mutual comprehension is impossible. One language may take many words to say what another language says in a single word, but in the end the circumlocution can make the point.

Similarly, it does not follow that, because a language lacks a word, its speakers therefore cannot grasp the concept. Several languages have few words for numerals: Australian aboriginal languages, for example, are often restricted to a few general words (such as 'all', 'many', 'few'), 'one' and 'two'. In such cases, it is sometimes said that the people lack the concept of number – that Aborigines 'haven't the intelligence to count', as it was once put. But this is not so, as is shown when these speakers learn English as a second language: their ability to count and calculate is quite comparable to that of English native speakers.

However, a weaker version of the Sapir–Whorf hypothesis is generally accepted. Language may not determine the way we think, but it does influence the way we perceive and remember, and it affects the ease with which we perform mental tasks. Several experiments have shown that people recall things more easily if the things correspond to readily available words or phrases. And people certainly find it easier to make a conceptual distinction if it neatly corresponds to words available in their language. A limited salvation for the Sapir–Whorf hypothesis can therefore be found in these studies, which are carried out within the developing field of psycholinguistics (p. 418).

HAVING A WORD FOR IT

There is nothing in everyday English to correspond to the many Arabic words for *horse* or *camel*, the Eskimo words for *snow*, or the Australian languages' words for *hole* or *sand*. Speakers of English have to resort to circumlocutions if they want to draw the distinctions which these languages convey by separate words – such as the size, breed, function, and condition of a camel. On the other hand, several languages cannot match the many words English has available to identify different sizes, types, and uses of vehicles – *car, lorry, bus, tractor, taxi, moped, truck*, and so on – and might have just one word for all of these.

There is in fact no single word in English for the driver of all kinds of motor vehicles – *motorist* being restricted to private cars, and *driver* being unacceptable for motorcycles – a lexical gap which greatly worried the British Automobile Association in 1961. It was felt that such a word would be useful, and they therefore asked for suggestions. Among the 500 they received were:

autoist	*autonaut*
roadist	*vehiclist*
chassimover	*murderist*
mobilist	*roadent*
wheelist	*vehicuwary*

doice (Driver Of Internal Combustion Engine)
pupamotor (Person Using Power-Assisted Means of Travel on Roads)
licentiat (Licensed Internal Combustion Engine Navigator Trained in Automobile Tactics)

However, none of these ingenious ideas has survived.

WORDS FOR *HOLE* IN PINTUPI

It takes between three and 14 English words to distinguish the various senses of *hole* in this Australian aboriginal language, but the distinctions can nonetheless be conveyed.

yarla a hole in an object

pirti a hole in the ground

pirnki a hole formed by a rock shelf

kartalpa a small hole in the ground

yulpilpa a shallow hole in which ants live

mutara a special hole in a spear

nyarrkalpa a burrow for small animals

pulpa a rabbit burrow

makarnpa a goanna burrow

katarta the hole left by a goanna when it has broken the surface after hibernation

8 · GEOGRAPHICAL IDENTITY

The most widely recognized features of linguistic identity are those that point to the geographical origins of the speakers – features of *regional dialect,* which prompt us to ask the question 'Where are they from?' But there are several levels of response to this question. We might have a single person in mind, yet all of the following answers would be correct 'America', 'The United States', 'East Coast', 'New York', 'Brooklyn'. People belong to regional communities of varying extent, and the dialect they speak changes its name as we 'place' them in relation to these communities.

Languages, as well as dialects, can convey geographical information about their speakers, but this information varies greatly, depending on the language of which we are thinking. The variation can be seen if we complete a test sentence using different language names: 'If they speak —, they must be from —.' If the first blank is filled by 'Swedish', the second blank will almost certainly be filled by 'Sweden'. But 'Portuguese' would not inevitably lead to 'Portugal': the second blank could be filled by 'Brazil', 'Angola', 'Mozambique', and several other countries. 'French' would give us the choice of about 40 countries, and 'English' well over 50. 'Dialect', by contrast with 'language', is a much more specific geographical term.

POPULAR NOTIONS OF DIALECT

It is sometimes thought that only a few people speak regional dialects. Many restrict the term to rural forms of speech – as when they say that 'dialects are dying out these days'. They have noticed that country dialects are not as widespread as they once were, but they have failed to notice that urban dialects are now on the increase (p. 32). Another view is to see dialects as substandard varieties of a language, spoken only by low-status groups – implicit in such comments as 'He speaks correct English, without a trace of dialect'. Comments of this kind fail to recognize that standard English is as much a dialect as any other variety – though a dialect of a rather special kind (p. 39). Or again, languages in isolated parts of the world, which may not have been written down, are sometimes referred to pejoratively as dialects, as when someone talks of a tribe speaking 'a primitive kind of dialect'. But this fails to recognize the true complexity and range of all the world's languages (§47).

In this encyclopedia, as is standard practice in linguistics, dialects are seen as applicable to all languages and all speakers. In this view, all languages are analysed into a range of dialects, which reflect the regional and social background of their speakers. The view maintains that everyone speaks a dialect – whether urban or rural, standard or non-standard, upper class or lower class. And no dialect is thought of as 'superior' to any other, in terms of linguistic structure – though several are considered prestigious from a social point of view.

WHERE ARE YOU FROM?

How easy is it to tell where someone is from? A few years ago, it would have been relatively straightforward for a specialist to work out from a sample of speech the features that identified someone's regional background. Some dialect experts have been known to run radio shows in which they were able to identify the general regional background of members of their audience with considerable success. But it is doubtful whether anyone has ever developed the abilities of Shaw's Henry Higgins: '*I* can place any man within six miles. I can place him within two miles in London. Sometimes within two streets' (*Pygmalion,* Act 1).

These days, dialect identification has become much more difficult, mainly because of increased social mobility. In many countries, it is becoming less common for people to live their whole lives in one place, and 'mixed' dialects are more the norm. Also, as towns and cities grow, once-distinct communities merge, with a consequent blurring of speech patterns. And nowadays, through radio and television, there is much more exposure to a wide range of dialects, which can influence the speech of listeners or viewers even within their own homes. A radio dialect show would be much less impressive today. On the other hand, meticulous analysis can bring results, and there have been several notable successes in the field of forensic linguistics (p. 69).

DIALECT OR ACCENT?

It is important to keep these terms apart, when discussing someone's linguistic origins. *Accent* refers only to distinctive pronunciation, whereas *dialect* refers to grammar and vocabulary as well. If we heard one person say *He done it* and another say *He did it,* we would refer to them as using different dialects, because a grammatical difference is involved. Similarly, the choice between *wee bairn* and *small child* is dialectal, because this is a contrast in vocabulary. But the difference between *bath* with a 'short *a*' [a] and *bath* with a 'long *a*' [ɑ:] is to do with accent, as this is solely a matter of pronunciation (or phonology, §28).

Usually, speakers of different dialects have different accents; but speakers of the same dialect may have different accents too. The dialect known as 'standard English' is used throughout the world, but it is spoken in a vast range of regional accents.

DIALECT, IDIOLECT, AND LECT

Probably no two people are identical in the way they use language or react to the usage of others. Minor differences in phonology, grammar, and vocabulary are normal, so that everyone has, to a limited extent, a 'personal dialect'. It is often useful to talk about the linguistic system as found in a single speaker, and this is known as an *idiolect.* In fact, when we investigate a language, we have no alternative but to begin with the speech habits of individual speakers: idiolects are the first objects of study. Dialects can thus be seen as an abstraction, deriving from an analysis of a number of idiolects; and languages, in turn, are an abstraction deriving from a number of dialects.

It is also useful to have a term for *any* variety of a language which can be identified in a speech community – whether this be on personal, regional, social, occupational, or other grounds. The term *variety* is itself often used for this purpose; but in recent years, many sociolinguists (p. 418) have begun to use *lect* as a general term in this way.

LANGUAGE VS DIALECT

One of the most difficult theoretical issues in linguistics is how to draw a satisfactory distinction between language and dialect. The importance of this matter will be repeatedly referred to in Part IX, where we have to make judgments about the number of languages in the world and how they are best classified.

At first sight, there may appear to be no problem. If two people speak differently, then, it might be thought, there are really only two possibilities. Either they are not able to understand each other, in which case they can be said to speak different languages; or they do understand each other, in which case they must be speaking different dialects of the same language. This criterion of *mutual intelligibility* works much of the time; but, unfortunately, matters are not always so simple.

MUTUAL INTELLIGIBILITY

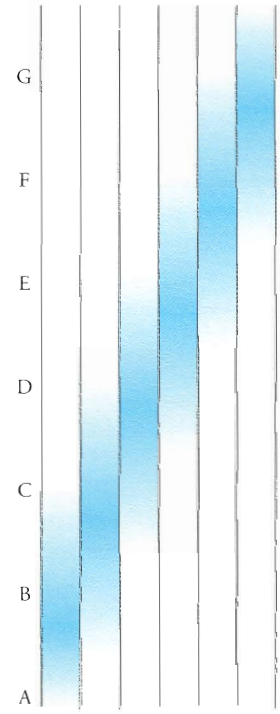
One common problem with this criterion is that dialects belonging to the same language are not always mutually intelligible in their spoken form. It can be very difficult for someone speaking a regional dialect in one part of Britain to understand some of the regional dialects of other areas; and the degree of intelligibility can be even worse when people attempt to communicate with English speakers from other countries. However, at least all of these speakers have one thing in common: they share a common written language. On this count, the varieties they speak could justly be called dialects of the same language.

A rather more serious problem arises in cases where there is a geographical *dialect continuum*. There is often a 'chain' of dialects spoken throughout an area. At any point in the chain, speakers of a dialect can understand the speakers of other dialects who live in adjacent areas to them; but they find it difficult to understand people who live further along the chain; and they may find the people who live furthest away completely unintelligible. The speakers of the dialects at the two ends of the chain will not understand each other; but they are nonetheless linked by a chain of mutual intelligibility.

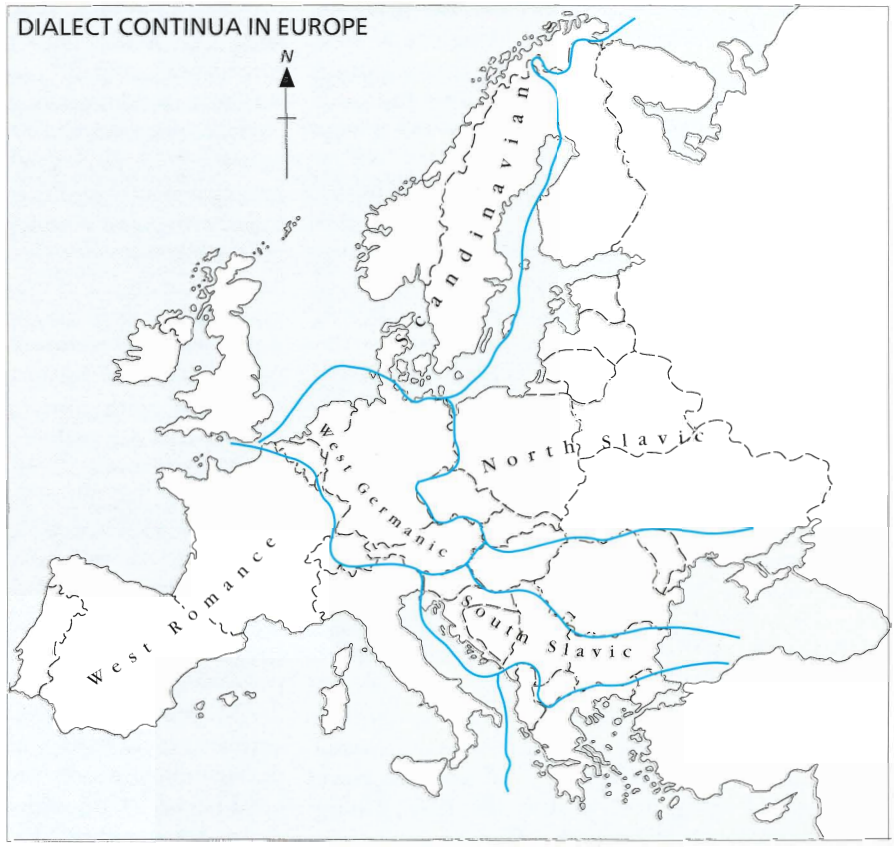
This kind of situation is very common. An extensive continuum links all the dialects of the languages known as German, Dutch, and Flemish. Speakers in eastern Switzerland cannot understand speakers in eastern Belgium; but they are linked by a chain of mutually intelligible dialects throughout the Netherlands, Germany, and Austria. Other chains in Europe include the Scandinavian continuum, which links dialects of Norwegian, Swedish, and Danish; the West Romance continuum, which links rural dialects of Portuguese, Spanish, Catalan, French, and Italian; and the North Slavic continuum, which links Slovak, Czech, Ukrainian, Polish, and Russian.

The theoretical problem should be clear. At what point in the chain can we say that one language ends and the next begins? On what basis can we draw boundary lines between Portuguese, Spanish, French, and so on? We are used to thinking of these languages as quite different from each other, but this is only because we are usually exposed to their standard varieties, which are not mutually intelligible. At the local level, it is not possible to make a clear decision on linguistic grounds.

But decisions are of course made on other grounds. As one crosses a well-established national boundary, the variety of speech will change its name: 'Dutch' will become 'German', 'Spanish' will become 'Portuguese', 'Swedish' will become 'Norwegian'. It is important to appreciate that the reasons are political and historical, not linguistic (§47). Arguments over language names often reduce to arguments of a political nature, especially when there is a dispute over national boundaries. For example, in the South Slavic continuum, varieties spoken on the western side of the border between the Former Yugoslav Republic of Macedonia and Bulgaria are called dialects of Macedonian by the former country, but dialects of Bulgarian by the latter – reflecting a claim to the territory. However, because there is a dialect chain in the area, linguistic criteria will never be able to solve conflicts of this kind.



A schematic dialect continuum between dialects A and G. The possible degrees of mutual intelligibility are represented by different shading, from maximum (dark) to zero (light).



DIALECTOLOGY

The systematic study of regional dialects is known variously as *dialectology*, *dialect geography*, or *linguistic geography*; but these terms are not exact equivalents. In particular, the latter terms suggest a much wider regional scope for the subject. Dialect specialists who spend their lives researching the local usage of a single Yorkshire village can hardly be called 'linguistic geographers', though they are certainly 'dialectologists'. By contrast, the 'geographer' designation would be quite appropriate for anyone involved in plotting the distribution of forms over a large area, such as Scotland, or the eastern United States.

There is another difference between these terms. Traditionally, dialectology has been the study of regional dialects, and for many people that is still its main focus. But in recent years, dialectologists have been paying more attention to social as well as geographical space, in order to explain the extent of language variation (§§9–10). Factors such as age, sex, social class, and ethnic group are now seen as critical, alongside factors of a purely regional kind.

But whatever the approach, the contemporary fascination with dialects seems no less than that shown by previous generations. Radio programmes on dialect variations are popular in several countries, and compilations of dialect data continue to be produced in the form of grammars, dictionaries, folk-lore collections, and guides to usage. Local dialect societies thrive in many parts of the world. Dialects continue to be seen as a major source of information about contemporary popular culture and its historical background; and dialect variation forms part of the study of change (§54).

Probably the most important application of dialectology these days is in education, where the development of dialect 'awareness' in children is widely recognized as a way of getting them to see the heterogeneity of contemporary society, and their place within it (§§44, 61). Teachers are often faced with a conflict between the child's spontaneous use of dialect forms and the need to instil a command of the standard language, especially in writing. The conflict can be resolved only by developing in children a sense of the relationships between the two kinds of language, so that the value of both can be better appreciated. There needs to be an awareness of the history, structure, and function of present-day dialects – and this is what dialectology can provide.

THE HISTORY OF REGIONAL DIALECTOLOGY

While there has been sporadic interest in regional dialects for centuries, the first large-scale systematic studies, in Germany and France, did not take place until the end of the 19th century. In 1876, Georg Wenker (1852–1911) began sending out questionnaires to all the school districts in the German Empire. It took him ten years to contact nearly 50,000 local teachers, who were asked to provide equivalents for 40 sentences in the local dialect. An enormous amount of data was received, and this led to the publication in 1881 of the first linguistic atlas, *Sprachatlas des Deutschen Reichs*. A larger series of works, based on Wenker's files, appeared between 1926 and 1956; but even today, much of the original material has not been published.

The postal questionnaire method enables a large amount of data to be accumulated in a relatively short time, but it has several limitations – chiefly that dialect pronunciations cannot be accurately recorded. The alternative, to send out trained field workers to observe and record the dialect forms, was first used in the linguistic survey of France, which began in 1896. The director, Jules Gilliéron (1854–1926), appointed Edmond Edmont (1849–1926) – a grocer with a very sharp ear for phonetic differences – to do the field work. For four years, Edmont went around France on a bicycle, conducting interviews with 700 informants using a specially devised questionnaire of nearly 2,000 items. The *Atlas linguistique de la France* was subsequently published in 13 volumes between 1902 and 1910. It stands as the most influential work in the history of dialectology.

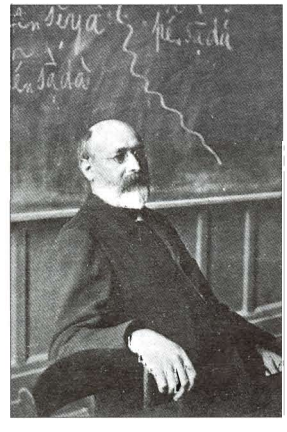
In the first half of this century, major projects were initiated in many parts of Europe, such as Romania, Italy, Holland, Spain, and Denmark, and there have been several impressive publications. In due course the large-scale dialect surveys of the United States and England began (p. 30). A great deal of dialect work has also been undertaken in Japan and China, as well as in parts of Africa, Australia, Canada, and South America. In some countries, even, surveys leading to a 'second generation' of linguistic atlases have begun. Direct interviewing and postal questionnaires continue to be used today, as does the tradition of presenting the linguistic material in the form of maps; and in recent years, dialectology has benefited enormously from the development of techniques using tape recorders. The field is also now being influenced by the electronic revolution, with computers helping to 'crunch' the data provided by questionnaires, and making large databases of regional variants more available, accessible, and analysable – and even more visible, using computer graphic techniques.

However, nowadays there are fewer big regional dialect projects, and some of those that have begun may never be completed. This is mainly because of the large costs involved in collecting, analysing, and publishing dialect data; but it is also partly because of the new direction dialect studies have taken. Younger scholars are these days more likely to be attracted by the sociolinguistically inspired approaches that developed in the 1970s. with their focus on social factors, and on urban rather than on rural dialects (p. 32).

THE EARLIEST USE OF DIALECTOLOGY?

Then Gilead cut Ephraim off from the fords of the Jordan, and whenever an Ephraimite fugitive said 'Let me cross', the men of Gilead asked him, 'Are you an Ephraimite?'. If he answered 'No', they said, 'Then say "Shibboleth".' He would say 'Sibboleth', since he could not pronounce the word correctly. Thereupon they seized and slaughtered him by the fords of the Jordan.
(Judges XII, 4–6)

The Ephraimites were betrayed by their regional pronunciation. As a result of this story, *shibboleth*, which then meant 'ear of corn' or 'flowing stream', has in modern use come to mean 'distinguishing mark' or 'criterion'.



Jules Gillieron (1854–1926)

THE FARM

THE FARMSTEAD

Show an aerial photograph of a farmstead and surrounding fields @.

1 ... these? **Fields**
2 ... this? **Farmstead.**
3 ... this? **Farmyard.**
4 ... this? **Stackyard.**
... the various buildings?

If necessary, ask the relevant question below.

5 ... the place where you keep pigs? **Pigsty.**–April 1953, *the animals that go* (*i. grunting*) replaced *pigs.*
6 ... the place where you keep hens? **Hen-house.**–April 1953, *the birds that lay eggs for you* replaced *hens.*
7 ... *the place where you keep pigeons?* **Dove-cote.**–April 1953, *the birds that go (i. cooing)* replaced *pigeons.*
8 ... the place where you keep cows? **Cow-house.**–April 1953, *the animals that give you milk* replaced *your cows.*
9 ... the yard in which cattle are kept, especially during the winter, for fattening, and for producing dung? **Straw-yard.** (Verify the kind of cattle and the purpose).
10 ... the small enclosed piece of pasture near the farmhouse, the place where you might put a cow or a pony that's none too well? **Paddock.**
11 What's the **barn** for and where is it?

COW-HOUSE

Q. *What do you call the place where you keep your cows?* – April 1953, *the animals that give you milk* replaced *your cows.*

Rr. BEEF-HOUSE (COW-)BYRE, COW-HOLE/HOUSE/HULL/SHADE/SHED, LATHE, MISTALL, SHIPPON

1 Nb 1 baɪɔᵏ 2 baɪɔᵏ [baɪɔᵏmən[1] *byre-man* (= *cowman*) I.2.3] 3 ku:baɪɔᵏ 4–5 baɪɔᵏ 6 baɪəᵏ 7 baɪɔᵏ, ○□baɪɔᵏz[1] 8 baɪɔᵏ 9 baɪə

2 Cu 1 baɪər 2 baɪər 3 baɪə, ku:əs 4 baɪə, ○□baɪəz[1] 5 ku:baɪəɹ 6 baɪə, kᵊu:əs ["old name"]

3 Du 1 ku:baɪɔᵏ, ○ku:ʃɪəɖ[3] 2 bɛ̈ɪəᵏ 3 baɪər 4–5 baɪə 6 baɪə ○□baɪəz[1]

4 We 1 baɪəɹ, ○baɪə[1] 2–3 baɪə 4 ʃɷpm

5 La 1–3 ʃɷpm 4 ʃɪpn, ○ʃɷpn[1] 5 ʃɪpn 6 ʃɪpm, ʃɷpm ["older"], ○ʃɪpən[1] III.11.3, ○□ʃɪpmz[1] 7 ʃɪpn, ○□ʃɪpənz[1] 8–9 ʃɪpn 10 ʃɪpən 11 ʃɪppɷn 12 ʃɪpɷn, ○ʃɷpɷn[1] 13 ʃɪpən, ○ʃɪpɪn[2] 14 ʃɪpɷn

QUESTIONNAIRES

In a large dialect survey, there will be many informants and several investigators. One way of ensuring that the results of all the interviews will be comparable, while also saving a great deal of time, is through the use of questionnaires. On the other hand, unless the questions are particularly ingenious, the responses will lack the spontaneity of informal speech. Results thus have to be interpreted with caution.

Opposite is an extract from the questionnaire used in the English Dialect Survey (p. 30). The dots at the beginning of each line stand for 'What do you call ...', *i* = imitate. The second extract illustrates the depth of phonetic detail recorded by the field workers. Abbreviations after each number stand for the different northern counties of England.

PAUSY, *adj.* n.Lin.[1] [**pō̤zi.**] Slightly intoxicated. Slightly the worse for drink; said of persons who combine an amiable desire to impart information with an incapacity to call to mind all the necessary words. 'Drunk ! naw he was n't what you'd call drunk, nobbud he was pausy like.'

PAUT, *v.* and *sb.* Sc. Nhb. Dur. Lakel. Yks. Lan. Chs. Der. Not. Lin. Wor. Suf. Also written **pawt** Sc. Lakel.[2] Cum.[14] n.Yks.[2] e.Yks.[1] m.Yks.[1] w.Yks. ne.Lan.[1] Der.[1] Not.[13] n.Lin.[1] sw.Lin.[1]; **pawte** w.Yks.; port w.Yks. Not.[3]; and in forms **paat** Cai.[1] Nhb.[1] Cum.[14]; **paout** se.Wor.[1]; **pout** Sc. (JAM.) N.Cy.[1] s.Wor.; **powt** Sc. (JAM.) Bnff.[1] n.Cy Suf.[1] [**pōt, pǭət, pāt.**] 1. *v.* To poke or push with the hand or a stick; to stir up; to paw, handle, or finger things. Cf. **pote.**

Sc. To search with a rod or stick in water, or in a dark or confined place. To make a noise when searching or poking in water (JAM.). **n.Cy.** GROSE (1790). **Nhb.**[1] Divent paat on wi'd, or ye'll spoil'd. **Cum.** Children pawt when they make repeated attempts to get things with their hands (E.W.P); **Cum.**[4] A dog pawts at the door when it wants to get in, and children pawt when they make repeated attempts to get hold of things with their hands. **n.Yks.**[1]; **n.Yks.**[2] Kneading with the fingers into a soft mass. **n.Lin.** SUTTON *Wds.* (1881); **n.Lin.**[1] I wish we hed n't noä cats, really, thaay're alus pawtin' at one, when one's gettin' one's meät. **sw.Lin.**[1] Some lasses are always pawting things about they've no business with. **s.Wor.** To beat down apples, PORSON *Quaint Wds.* (1875) 15.

Hence (1) **Pouting,** *vbl. sb.* the practice of spearing salmon ; also used *attrib.* ; (2) **Pout-net,** *sb.* a net fastened

An extract from the *English Dialect Dictionary*
Joseph Wright (1855–1930), published this dictionary in six volumes between 1898 and 1905; it contained 100,000 entries. Wright was largely self-taught, and did not learn to read until he was a teenager – a fact that may have been an advantage to him in his later studies, as his early awareness of dialect differences would not have been influenced by the forms of the standard written language.

FROM STRINE TO SCOUSE

The contrast between regional dialect and standard English usage has been a source of humour the world over. In *Let Stalk Strine* (1965). Afferbeck Lauder (said to be Professor of Strine Studies at the University of Sinny) uses standard spellings to represent the popular impression of an Australian accent, with bizarre results:

Egg Nishner: A mechanical device for cooling and purifying the air of a room.
Jezz: Articles of furniture. As in: 'Set the tible, love, and get a coupler jezz'.
Money: The day following Sunny. (Sunny, Money, Chewsdy, Wensdy, Thursdy, Fridy, Sairdy.)
Scone: A meteorological term. As in: 'Scona rine'.
Sly Drool: An instrument used by engineers for discovering Kew brutes and for making other calculations.
Tiger: Imperative mood of the verb to take. As in: 'Tiger look at this, Reg...'
X. The twenty-fourth letter of the Strine alphabet; also plural of egg; also a tool for chopping wood.

Some of the colloquial pronunciations here are found in many dialects. For example, *Gissa* ('Please give me ...') is a feature of Strine, but it is also well known in Liverpool, as can be seen from the section on 'Forms of Address' in *Lern Yerself Scouse* (1966), by Frank Shaw, Fritz Spiegl, and Stan Kelly (whose standard English translations are given in parentheses):

Ullo dur! ('Greetings; I am pleased to make your acquaintance.')
Gisalite ('Could you oblige me with a match, please?')
Ay-ay ('I say!')
La ('I say, young man.')
Ere, tatty-head! ('I say, young woman!')

In the Appendix to this work, selected verses from *The Rubáiyat of Omár Khayyám* are translated into Scouse by Stan Kelly:

Gerrup dere La! De knocker-up sleeps light;
Dawn taps yer winder, ends anudder night;
And Lo! de dog-eared moggies from next-door
Tear up de jigger fer an early fight.

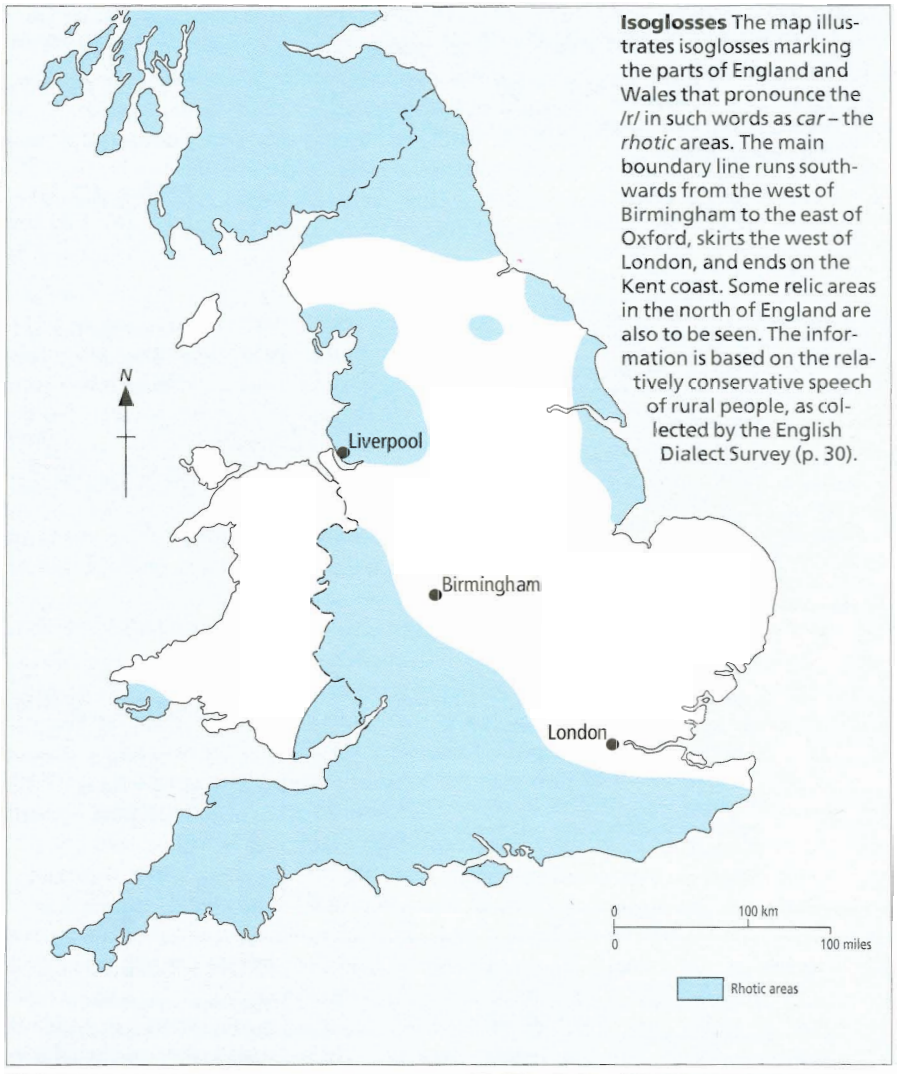
LINES ON MAPS

Once the speech of dialect informants has been collected, it is analysed, and the important features are marked on a map of the area in which the informants live. When several points on the map have been located, it is then possible to see whether there is a pattern in the way these features are used. The usual way of identifying dialect patterns is to draw lines around the places where the people use a linguistic feature in the same way. These boundary lines are known as *isoglosses*. For example, one famous isogloss runs across England, from the Severn to the Wash: it distinguishes northern speakers who pronounce a rounded *u* /ʊ/ in words like *cup* from southern speakers who keep the vowel open and unrounded, /ʌ/. A series of lexical isoglosses, identifying various words for *snack*, is illustrated on p. 30.

When isoglosses were first introduced (in 1892), it was expected that they would provide a clear method for identifying dialect areas. Because people from a particular part of a country 'speak in the same way', it was assumed that the isoglosses for many linguistic features would coincide, and form a neat 'bundle', demarcating one dialect from another. However, early dialectology studies soon discovered that the reality was very different. Isoglosses crisscrossed maps in all directions, and very few actually coincided. There seemed to be no clear dialect boundaries at all – a finding which made some scholars go so far as to argue that the whole idea of a dialect was meaningless.

In due course, however, supplementary notions were developed to make sense of the data. It was noted that, while isoglosses rarely coincided, they did often run in the same general direction. Some areas, called *focal areas*, were seen to be relatively homogeneous, containing few isoglosses. Where focal areas merged, there was a great deal of linguistic variation, with many isoglosses present: these became known as *transition areas*. Often, a feature might be left isolated, as a result of linguistic change affecting the areas around it: these 'islands' of more conservative usage were called *relic areas*.

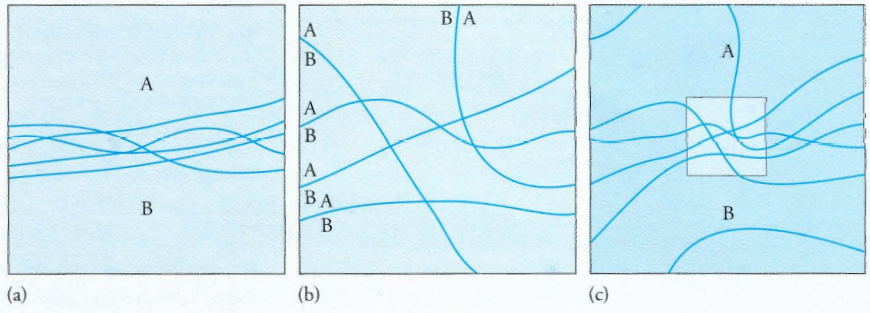
Dialectologists have mixed feelings about isoglosses. There is often too much variability in the way a linguistic feature is used for the data to be easily summarized in a single isogloss. Also, the relative significance of different isoglosses remains to be interpreted. Some isoglosses mark distinctions that are considered to be more important than others (such as the contrast between short and long *a* in words like *bath* in British English, which has long been the focus of special comment). Isoglosses are an important visual guide, but they need to be supplemented by other criteria if they are to display, and not to obscure, the true complexity of regional variation.



Isoglosses The map illustrates isoglosses marking the parts of England and Wales that pronounce the /r/ in such words as *car* – the *rhotic* areas. The main boundary line runs southwards from the west of Birmingham to the east of Oxford, skirts the west of London, and ends on the Kent coast. Some relic areas in the north of England are also to be seen. The information is based on the relatively conservative speech of rural people, as collected by the English Dialect Survey (p. 30).

The main kinds of isogloss

Term	Separates	Examples
isolex	lexical items	*nunch* vs *luncheon* (p. 30)
isomorph	morphological features	*dived* vs *dove*
isophone	phonological features	*put* /pʊt/ vs /pʌt/
isoseme	semantic features	*dinner* (mid-day meal) vs (evening meal)



The expectation Isoglosses will form neat bundles, demarcating dialect A from dialect B.

The reality Isoglosses criss-cross an area, with no clear boundary between A and B.

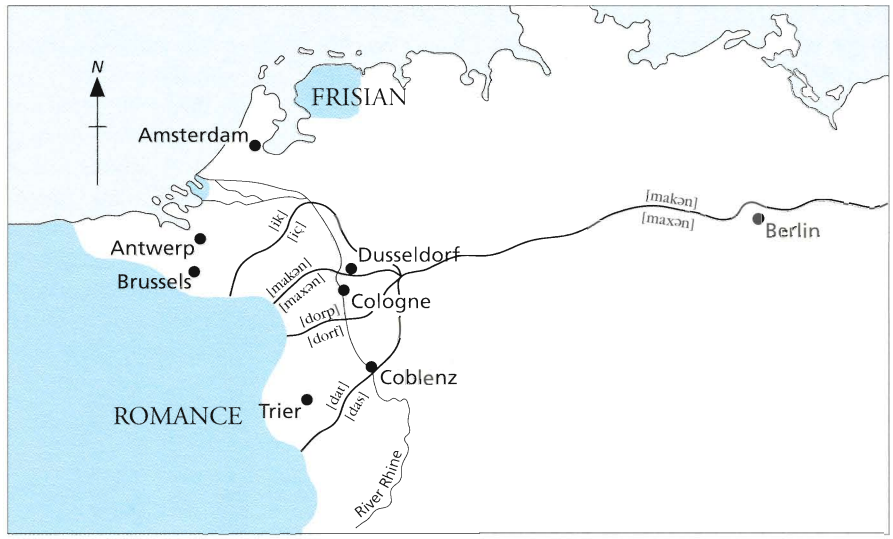
Focal and transitional On a larger scale, the isoglosses are seen to constitute a transitional area between the focal areas A and B.

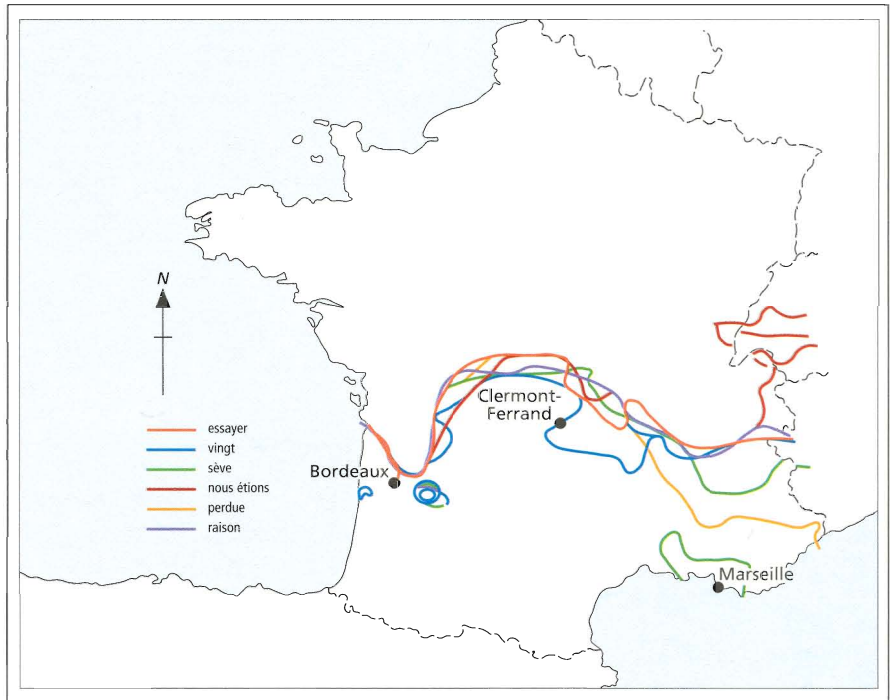
THE RHENISH FAN

One of the best examples of the way isoglosses fail to group themselves into bundles is in northern Europe. A set of isoglosses runs east–west across Germany and Holland, separating Low German, in the north, from High German, in the south. They reflect the different ways in which these dialects have developed the voiceless plosive consonants of Indo-European (p. 330). In Low German, the sounds have remained plosives (/p, t, k/); but in High German, these have generally become fricatives. For example, 'village' is [dorp] in the north, [dorf] in the south; 'that' is [dat], as opposed to [das]; 'make' and 'I' are [makən] and [ik] respectively, rather than [maxən] and [iç].

The map shows the location of the isoglosses that distinguish these words. Through most of Germany, they are close together, displaying only minor variations; but where they meet the River Rhine, the isoglosses move in quite different directions, in a pattern that resembles the folds in a fan. It thus becomes impossible to make simple generalizations about dialect differences in this area. A speaker in a village near Cologne, for example, would say [iç] and [maxən], as in High German, but say [dorp] and [dat], as in Low German.

What accounts for the Rhenish fan? It has been suggested that several of the linguistic features could be explained with reference to certain facts of social history. For example, the area between the [dorp/dorf] and [dat/das] isoglosses was coextensive with the old diocese of Trier; the area immediately north was coextensive with the old diocese of Cologne. The linguistic innovations seem to have spread along the Rhine from southern Germany to the cities, and then 'fanned out' throughout the administrative areas these cities controlled. Rural speakers were naturally influenced most by the speech of their own capital cities, and political and linguistic boundaries gradually came to coincide. (After L. Bloomfield, 1933.)





The two halves of France One of the main findings of the *Atlas linguistique de la France* (p. 26) was the bundle of isoglosses that runs across France from east to west, dividing the country into two major dialect areas. The areas are traditionally known as *langue d'oil* (in the north) and *langue d'oc* (in the south) – names based in the words for 'yes' current in these areas during the 13th century, when the division was first recognized. The map shows six items that are used differently on either side of an isogloss (J. K. Chambers & P. Trudgill, 1980, p. 111).

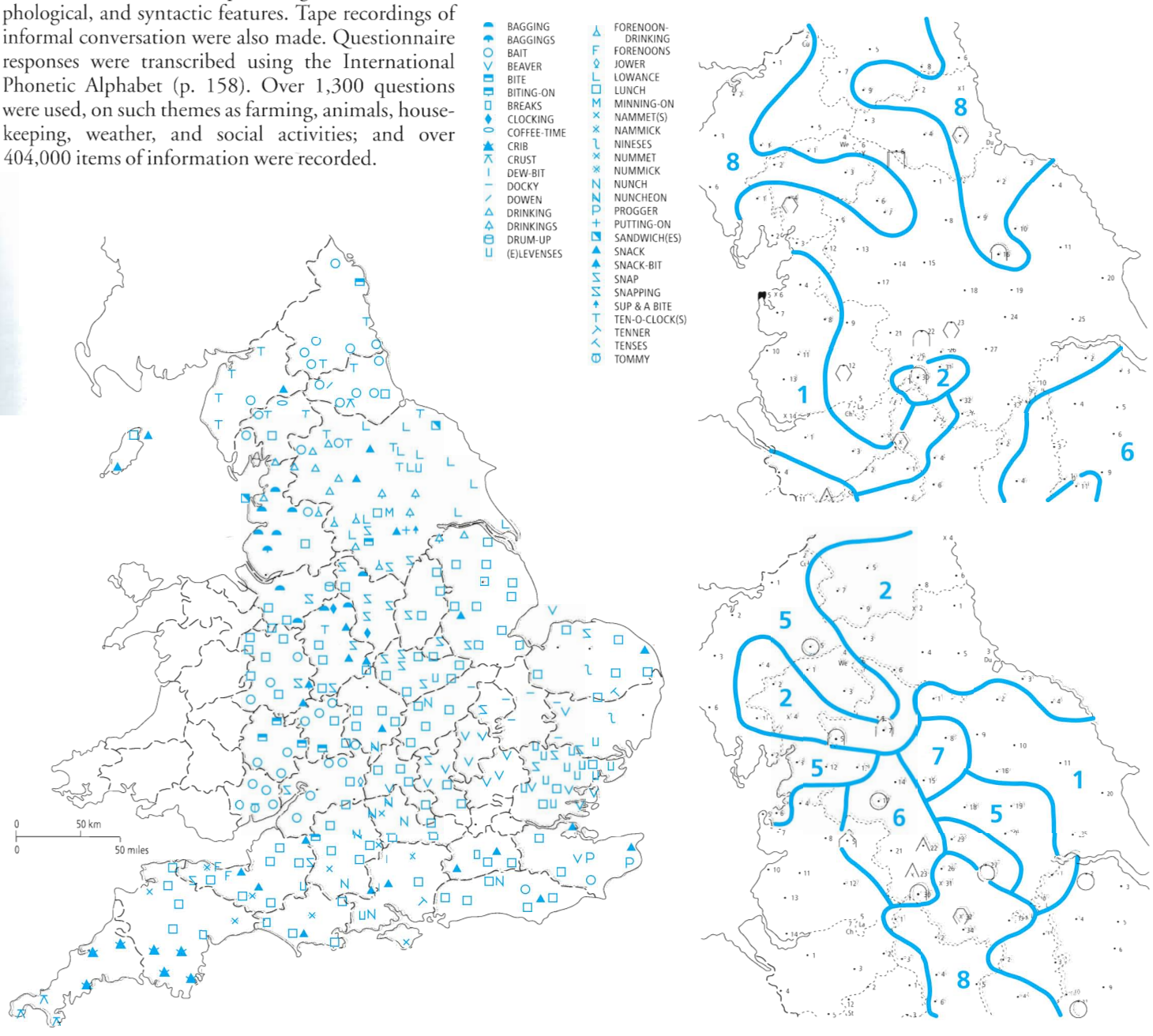
The distinction corresponds to several important social and cultural differences, some of which can still be observed today. For example, to the south of the isogloss bundle (roughly where the Provençal region begins), a biennial (as opposed to a triennial) method of crop rotation is traditionally used. A different legal system existed until the early 19th century, using a written code inspired by Roman traditions. And there is a major difference in architectural style, the roofs being generally flat, and not steeply pitched (as they are to the north of the bundle). Such clear correlations between language and cultural identity illustrate the way in which dialect studies form an important part of the study of social history.

THE LINGUISTIC ATLAS OF ENGLAND

Three of the maps from the English Dialect Survey, carried out by Harold Orton (1898–1975) and Eugene Dieth (1893–1956), are illustrated here. The field survey was undertaken between 1950 and 1961 in 313 localities throughout England. The localities were usually not more than 15 miles apart, and generally consisted of villages with a fairly stable population. The informants were natives of the locality, mainly male agricultural workers, with good mouths, teeth, and hearing, and over 60 years of age.

The principal method was a questionnaire that elicited information about phonological, lexical, morphological, and syntactic features. Tape recordings of informal conversation were also made. Questionnaire responses were transcribed using the International Phonetic Alphabet (p. 158). Over 1,300 questions were used, on such themes as farming, animals, housekeeping, weather, and social activities; and over 404,000 items of information were recorded.

Between 1962 and 1971 the basic material of the survey was published in an introduction and four separate volumes; in 1977 the *Linguistic Atlas of England* was published, containing an interpretation of a selection of the data. The maps below provide an example of the Survey's basic material for the item *snack* and two interpretive maps, based on this material. The first map is a display of all the responses obtained, which are listed in the top right-hand corner. The other maps pick out various trends in usage, and are a considerable simplification. (After H. Orton, S. Sanderson & J. Widdowson, 1978.)

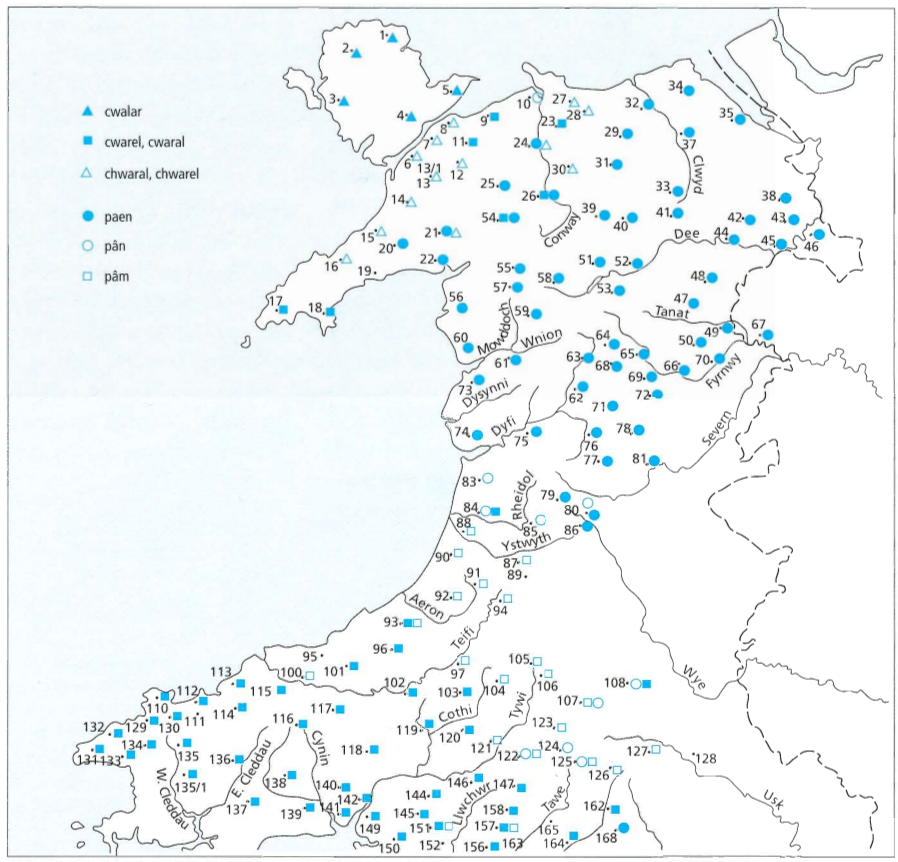


THE LINGUISTIC GEOGRAPHY OF WALES

One of the most recent dialect surveys was carried out in Wales in the 1960s under the direction of Alan R. Thomas (1935–) and published in 1973. It was based on 180 points of enquiry in the Welsh-speaking areas, the localities being selected on the basis of their position relative to the physical geography of the country and to the main communication routes.

The survey was based on a postal questionnaire, with questions using both Welsh and English. There were over 500 questions, which dealt largely with domestic, rural, and farming vocabulary; about 130,000 responses were received. The questionnaire was sent to a person of educated background, who supervised its completion by local informants, using spelling that reflected regional pronunciation. Informants were of the older generation, with little formal education, and had spent no prolonged periods away from their native area.

The main part of the atlas discusses the distribution of regional words for around 400 items, on the basis of which the main Welsh speech areas are drawn up. The illustration (right) shows the distribution of Welsh words for *pane of glass*, an item in which two distinct patterns of use can be clearly seen: *paen* and its variants in the north-east and the midlands, *cwalar* and its variants in most other places. (After A. R. Thomas, 1973.)

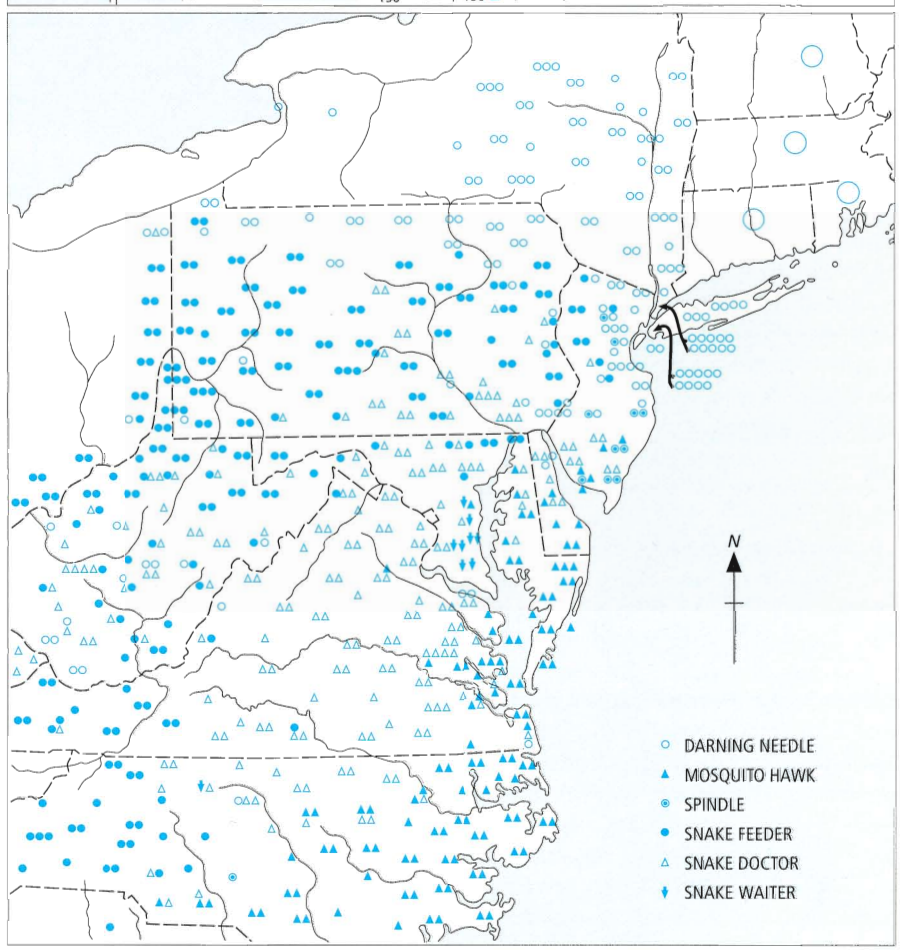


THE LINGUISTIC ATLAS OF THE UNITED STATES

This survey began in 1931, under the direction of Hans Kurath (1891–1992), as part of an ambitious programme to establish a linguistic atlas of the United States and Canada. The region was divided into survey areas, and the first atlas to appear, dealing with New England, was published in 1939–43. The project is ongoing, with informant interviews complete in many areas, but the amount of work involved means that publication is a slow and irregular process.

The illustration (right) is taken from Kurath's *Word Geography of the Eastern United States* (1949) – a survey area that included the coastal Atlantic states from Maine to Georgia, Pennsylvania, West Virginia, and eastern Ohio. Dialectologists went to nearly every county in these states and interviewed two people in each – one older-generation and unschooled, the other a member of the middle class with some degree of education. In the larger cities, people with a more cultured background were also interviewed. All were natives of their area, and had not moved much outside it. Interviewers spent from 10 to 15 hours with each informant, dealing with over 1,000 points of usage. More than 1,200 people were interviewed, and information was obtained about the diffusion of around 400 regional expressions for domestic and agricultural items.

The map records the distribution of words for *dragonfly*.



MODERN DIALECT STUDIES

Traditional dialectology studied geographical variation, generally using elderly, untravelled, and uneducated speakers from rural areas. Modern dialectology has moved in other directions.

Social factors now provide the focus of investigation. Speech variation can be partly understood with reference to regional location and movement, but social background is felt to be an equally if not more important factor in explaining linguistic diversity and change. Modern dialectologists therefore take account of socio-economic status, using such indicators as occupation, income, or education, alongside age and sex. Ideally informants are found in all social groups, and the traditional focus on the language of older people of working-class backgrounds has been replaced by the study of speakers of all ages and from all walks of life (§10).

Dialect studies have moved from the country to the city. The description of rural dialects led to fascinating results, but only a small proportion of a country's population was represented in such studies. In many countries, over 80% of the population live in towns and cities, and their speech patterns need to be described too – especially as linguistic change so often begins when people from the country imitate those from urban areas. This approach, accordingly, is known as *urban dialectology*.

Informants are now randomly selected. In the older studies, small numbers of speakers were carefully chosen to represent what were thought of as 'pure' forms of dialect. Today, larger numbers of people are chosen from the whole population of a city – perhaps using the electoral register or a telephone directory. Also, the earlier approach generally asked for one-word responses to a range of carefully chosen questions. This produced useful data, but these speech patterns were unlikely to have been typical. When people have their attention drawn to the way they speak, they usually adopt a more careful and unnatural style. Attempts are therefore now made to elicit speech that is more spontaneous in character by engaging informants in topics of conversation that they find interesting or emotionally involving (p. 334). The questionnaire has been largely replaced by the tape recorder.

LINGUISTIC VARIABLES

Traditional dialectology studied the fact that different people do not speak in the same way. Contemporary dialectology adds to this study the fact that the same person does not speak in the same way all the time. Individuals vary in their pronunciation, grammar, and vocabulary. Is there a reason for this variation, or is it random – 'free' variation, as it is often called? The current belief is that most of the variation is systematic, the result of the interplay between linguistic and social factors.

In the 1970s, the notion of the *linguistic variable* was developed, as a means of describing this variation. A linguistic variable is a unit with at least two variant forms, the choice of which depends on other factors, such as sex, age, social status, and situation. For example, in New York City, speakers sometimes pronounce /r/ in words like *car* and sometimes they do not. This unit can thus be seen as a variable, (r), with two variant forms, /r/ and zero. (It is usual to transcribe linguistic variables in parentheses.) It is then possible to calculate the extent to which individual speakers, or groups of speakers, use /r/, and to determine whether there is a correlation between their preferences and their backgrounds. Several interesting correlations have in fact been found (see also p. 334).

DROPPING THE /h/

In British English, the accent which carries most prestige (p. 39) pronounces /h/ at the beginnings of words such as *head*. But in most other accents of England and Wales, it is common to omit /h/ in this position. Regions do not pronounce or omit /h/ with total consistency, however, as can be seen from the results of two studies of this variable carried out in Norwich and Bradford.

The speakers were grouped into five social classes, based on such factors as their occupation, income, and education. The proportion of /h/-dropping was calculated, with the following results:

Class	*Bradford*	*Norwich*
Middle middle (MMC)	12%	6%
Lower middle (LMC)	28%	14%
Upper working (UWC)	67%	40%
Middle working (MWC)	89%	60%
Lower working (LWC)	93%	60%

The correlation is clear. In both areas, there is more /h/-dropping as one moves down the social scale. Moreover, the proportion is always greater in Bradford, suggesting that the phenomenon has been longer established in that area. (After J. K. Chambers & P. Trudgill, 1980.)

READING ALOUD IN NORWICH

People of different social levels were asked to read aloud a list of isolated words (A) and a piece of continuous text (B), and their pronunciations when reading were compared with their formal (C) and casual (D) speech.

The table shows whether the variable (ng) in such words as *walking* was pronounced /ŋ/ or /n/. (0 = no use of /n/; 100 = 100% use of /n/.)

Class	A	B	C	D
MMC	0	0	3	28
LMC	0	10	15	42
UWC	5	15	74	87
MWC	23	44	88	95
LWC	29	66	98	100

The consistency with which speakers increase their use of /n/ as their language becomes more spontaneous and casual is reflected at every social level. (After P. Trudgill, 1974.)

/l/ -DROPPING IN MONTREAL

The consonant /l/ is often dropped in the pronunciation of *il* ('he, it'), *elle* ('she, it'), *ils* ('they'), *la* ('her, it, the'), and *les* ('the, them'). The prestige forms retain the /l/. When usage is analysed by sex of speaker, a clear pattern emerges. (The numbers represent the percentage of /l/-dropping.)

	Male	*Female*
il (impersonal)	99	97
ils	94	90
il (personal)	94	84
elle	67	59
les (pronoun)	53	41
la (article)	34	25
la (pronoun)	31	23
les (article)	25	15

Women are much more likely to use the higher-prestige variant than men – a pattern of differentiation that has often been found in studies of urban dialectology. (After G. Sankoff & H. Cedergren, 1971.)

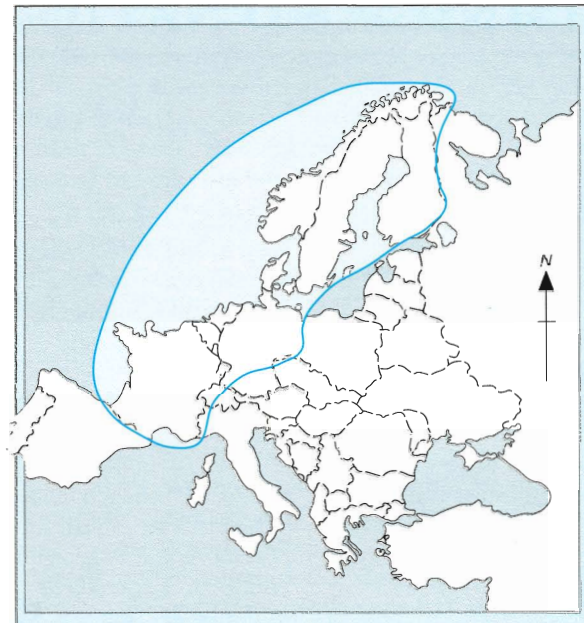
LINGUISTIC AREAS

Geographical identity can sometimes be established within a broader context than that provided by rural or urban dialectology. Certain features of speech can identify someone as coming from a particular part of the world, but the area involved may extend over several countries, languages, or even language families (§50). The study of 'areal features' of this kind is sometimes referred to as *areal linguistics.*

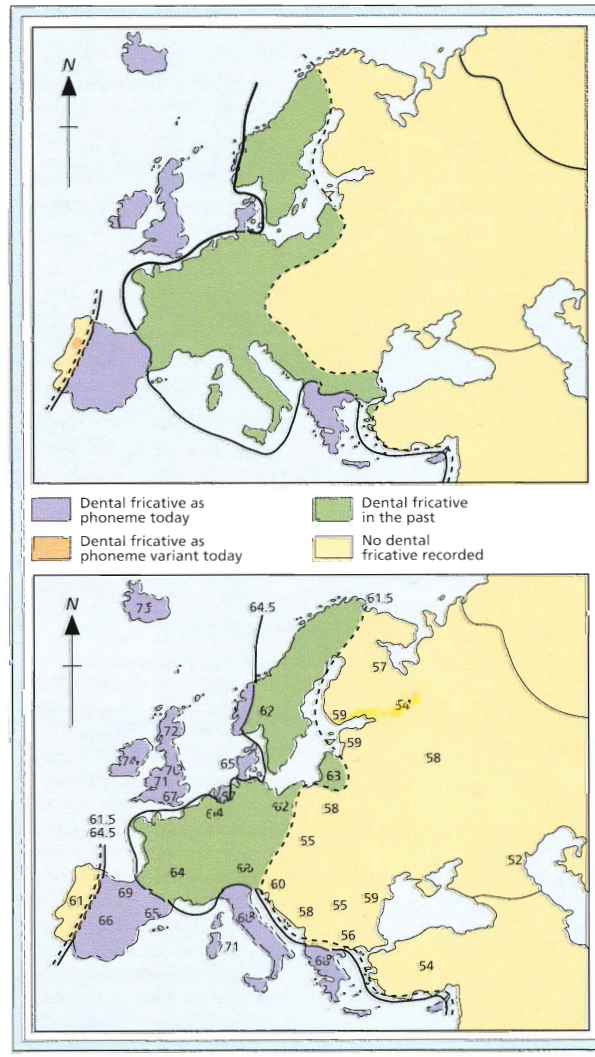
Features of pronunciation are often shared by adjacent, but historically-unrelated languages. In the indigenous languages of southern Africa (p. 317), the use of click sounds in speech identifies speakers of the Khoisan languages as well as of local Bantu languages, such as Zulu and Xhosa. In the Indian sub-continent (p. 310), languages that belong to different families (such as Indo-European and Dravidian) have several important phonological features in common – the use of retroflex consonants (p. 157) is particularly widespread, for example. In Europe the distribution of the affricate [ʧ] is interesting: it is found in many of the languages on the periphery of the area, such as Lapp, Romanian, Hungarian, Spanish, Galician, Basque, Italian, Gaelic, English, and the Slavic languages. The languages within this periphery, such as Danish, German, and French, do not use it.

Grammatical features can also cross linguistic and national boundaries. The use of particles to mark different semantic classes of nouns (§16) can be found throughout South-east Asia. In Europe, the Balkans constitutes a particularly well-defined linguistic area. For example, Albanian, Romanian, Bulgarian, and Macedonian all place the definite article *after* the noun, as in Romanian *lup* ('wolf') and *lupul* ('the wolf'), whereas historically-related languages outside of the Balkans area (such as Italian) do not.

How do areal features develop? In some areas, dialect chains (p. 25) have probably helped to diffuse a linguistic feature throughout an area. Concentrations of bilingual speakers along lines of communication would also play a part, and political factors will have exercised their influence. Sometimes, the progress of an areal feature can be traced – an example being the uvular pronunciation of /r/. Originally, speakers of European languages pronounced /r/ with the front of their tongue; but, in the 17th century, Parisians began to use a uvular variant. The variant caught on, spreading first throughout most of France, then to parts of Italy, Switzerland, Luxembourg, Belgium, Holland, Germany, Denmark, and (by the end of the 19th century) to southern Norway and Sweden. Spain, Austria, England, and other countries were not affected. The historical reasons for this complex state of affairs are little understood, and require investigation on several fronts. In such cases, the facts of dialectology, social history, and political history merge.



Front-rounded vowels
These vowels, such as in German *müde* ('tired') or French *soeur* ('sister'), are found along an axis which runs diagonally across northern Europe. They are heard in French, Dutch, German, Danish, Norwegian, Swedish, and Finnish. The feature cannot be explained on historical grounds: German and English are closely related, but the latter does not have front-rounded vowels; nor does Spanish, which is closely related to French. The main factor seems to be geographical proximity – as further illustrated by the way in which many south German dialects lack these vowels, whereas they are found in north-west Italy.
(J. K. Chambers & P. Trudgill, 1980, p. 185.)



A genetic explanation?
The distinctive European distribution of such sounds as front-rounded vowels, affricates, and dental fricatives has been studied from a genetic point of view. The geneticist C. D. Darlington (1903–) proposed in the 1940s that the genetic composition of a community would partly determine its preferences for types of sound. The maps show the distribution of dental fricatives in western Europe (above, left), and the frequency with which the O blood-group gene is distributed in the population (below, left). There seems to be an intriguing correlation: in populations where fewer than 60% have the gene, there is no history of these sounds; and in those where more than 65% have the gene, the sounds are well represented. Unfortunately, proposals of this kind have not been followed up, and remain only suggestive. There are also exceptions (e.g. /θ/ is used in Galician, in NW Spain). Social explanations of such distributions are currently felt to be far more likely. (After L. F. Brosnahan, 1961.)

10 · SOCIAL IDENTITY

In addition to the questions 'Who are you?' and 'Where are you from?', which have been addressed from a linguistic viewpoint in §§6–9, there is also '*What* are you, in the eyes of the society to which you belong?' It is a complex and multi-faceted question, to which there is no easy answer. People acquire varying status as they participate in social structure; they belong to many social groups; and they perform a large variety of social roles. As a consequence, no single system of classification is likely to do justice to the task of defining a person's social identity in linguistic terms, especially when the vast range of the world's cultural patterns is taken into account. This section, therefore, has to be extremely selective, in order to represent the range of sociolinguistic and ethnolinguistic variables involved.

SOCIAL STRATIFICATION

One of the chief forms of sociolinguistic identity derives from the way in which people are organized into hierarchically ordered social groups, or *classes*. Classes are aggregates of people with similar social or economic characteristics. Within sociology, the theoretical basis of social class has been a controversial subject, and it has not always proved easy to work consistently with the notion, especially when cross-cultural comparisons are involved. Factors such as family lineage, rank, occupation, and material possessions often conflict or are defined with reference to different criteria. But for most sociolinguistic purposes to date, it has been possible to make progress by recognizing only the broadest distinctions (such as high vs low, or upper vs middle vs lower) in order to determine the significant correlations between social class background and language. Examples of some of these correlations are given below and also on p. 32.

One does not need to be a sociolinguist to sense that the way people talk has something to do with their social position or level of education. Everyone has developed a sense of values that make some accents seem 'posh' and others 'low', some features of vocabulary and grammar 'refined' and others 'uneducated'. We have a large critical vocabulary for judging other people's language in this way. But one does need to be a sociolinguist to define precisely the nature of the linguistic features that are the basis of these judgments of social identity. And it is only as a result of sociolinguistic research that the pervasive and intricate nature of these correlations has begun to be appreciated.

CASTES

Probably the clearest examples of social dialects are those associated with a caste system. Castes are social divisions based solely on birth, which totally restrict a person's way of life – for example, allowing only certain kinds of job, or certain marriage partners (p. 405). The best-known system is that of Hindu society in India, which has four main divisions, and many sub-divisions – though in recent years, the caste barriers have been less rigidly enforced. The Brahmins (priests) constitute the highest class; below them, in descending order, are the Kshatriyas (warriors), Vaisyas (farmers and merchants), and Sudras (servants). The so-called 'untouchables', whose contact with the other castes is highly restricted, are the lowest level of the Sudra caste.

Linguistic correlates of caste can be found at all levels of structure. For example, in Tamil, there are several clear-cut distinctions between the phonology, vocabulary, and grammar of Brahmin and non-Brahmin speech. The former also tends to use more loan words, and to preserve non-native patterns of pronunciation.

Brahmin		**Non-Brahmin**
	Vocabulary	
tūngu	'sheep'	orangu
alambu	'wash'	kaḷuyu
jalō	'water'	taṇṇi
	Phonology	
krāfu	'haircut'	krāppu
jīni	'sugar'	cīni
vārepparo	'banana'	vāreppolo vāleppolo
	Grammar	
–du	'it'	–ccu
vandudu	'it came'	vanduccu
paṇra	'he does'	pannuhã

(After W. Bright & A. K. Ramanujan, 1964.)

SPEECH AND SILENCE IN KIRUNDI

In the Central African kingdom of Burundi, age and sex combine with caste to constrain the nature of linguistic interaction in several ways. Seniority (*ubukuru*) governs all behaviour. There are clear caste divisions; older people precede younger; and men precede women. The order in which people speak in a group is strictly governed by the seniority principle. Males of highest rank must speak first, regardless of age. Females do not speak at all, in the presence of outsiders, unless spoken to.

Upper-caste speakers seem never to raise their voices, or allow emotion to show. In group discussion, for the senior person to be silent implies disapproval. As others must then also stay silent, any further proceedings are effectively negated.

To speak well is considered a mark of good breeding in men. From their tenth year, boys in the upper castes are given formal speech training – how to use social formulae, talk to superiors and inferiors, and make speeches for special occasions. Upper-caste girls do not take part in public speaking, but they do develop effective bargaining skills, for use behind the scenes. They are also trained to listen with great care, so that they can accurately recount to the men of the family what has been said by visitors. (After E. M. Albert, 1964.)

The John Betjeman poem, 'How to get on in society', originally set as a competition in *Time and Tide*, was included in the book *Noblesse Oblige* as part of the U/non-U debate (see facing page).

HOW TO GET ON IN SOCIETY

Phone for the fish-knives, Norman,
As Cook is a little unnerved;
You kiddies have crumpled the serviettes
And I must have things daintily served.

Are the requisites all in the toilet?
The frills round the cutlets can wait
Till the girl has replenished the cruets
And switched on the logs in the grate.

It's ever so close in the lounge, dear,
But the vestibule's comfy for tea,
And Howard is out riding on horseback
So do come and take some with me.

Now here is a fork for your pastries
And do use the couch for your feet;
I know what I wanted to ask you –
Is trifle sufficient for sweet?

Milk and then just as it comes, dear?
I'm afraid the preserve's full of stones;
Beg pardon, I'm soiling the doilies
With afternoon tea-cakes and scones.

SOME ENGLISH MARKERS OF SOCIAL CLASS

Long before the days of 20th-century linguistics and phonetics, English novelists and dramatists, especially in the 18th and 19th centuries, were observing the relationship between language and social class in Britain and using it as a basis for characterization and social comment.

- George Gissing, about Mrs Yule, in his *New Grub Street* (1891, Chapter 7).

Mrs Yule's speech was seldom ungrammatical, and her intonation was not flagrantly vulgar, but the accent of the London poor, which brands as with hereditary baseness, still clung to her words, rendering futile such propriety of phrase as she owed to years of association with educated people.

- Mrs Waddy, about Harry Richmond's father, in George Meredith's *The Adventures of Harry Richmond* (1871, Chapter 3).

'More than his eating and his drinking, that child's father worrits about his learning to speak the language of a British gentleman ... Before that child your "h's" must be like the panting of an engine – to please his father ... and I'm to repeat what I said, to make sure the child haven't heard anything ungrammatical ...'

- Pip to Biddy, in Charles Dickens' *Great Expectations* (1861, Chapter 35).

'Biddy', said I, in a virtuously self-asserting manner, 'I must request to know what you mean by this?'
'By this?' said Biddy.
'No, don't echo,' I retorted. 'You *used not* to echo, Biddy.'
'*Used not!*' said Biddy. 'O Mr Pip! *Used!*'

- Elfride Swancourt to Mrs Swancourt, in Thomas Hardy's *A Pair of Blue Eyes* (1873, Chapter 14).

'I have noticed several ladies and gentlemen looking at me.'
'My dear, you mustn't say "gentlemen" nowadays ...
We have handed "gentlemen" to the lower classes, where the word is still to be heard at tradesmen's balls and provincial tea-parties, I believe. It is done with here.'
'What must I say then?'
'"Ladies and men" always.'

Dropping the *g*

'Where on earth did Aunt Em learn to drop her g's?'
'Father told me once that she was at a school where an undropped "g" was worse that a dropped "h". They were bringin' in a country fashion then, huntin' people, you know.'

This conversation between Clare and Dinny Cherrel, in John Galsworthy's *Maid in Waiting* (1931, Chapter 31), illustrates a famous linguistic signal of social class in Britain – the two pronunciations of final *ng* in such words as running, [n] and [ŋ]. But it also brings home very well the arbitrary way in which linguistic class markers work. The [n] variant is typical of much working-class speech today (p. 32), but a century ago this pronunciation was a *desirable* feature of speech in the upper middle class and above – and may still occasionally be heard. The change to [ŋ] came about under the influence of the written form: there was a *g* in the spelling, and it was felt (in the late 19th century) that it was more 'correct' to pronounce it. As a result, 'dropping the *g*' in due course became stigmatized.

U AND NON-U

In 1954, the British linguist A. S. C. Ross published an article entitled 'Linguistic class-indicators in present-day English' in a Finnish philological journal. It was read by Nancy Mitford , who wrote an *Encounter* article based upon it. The result was an enormous public reaction, with immediate recognition for the terms *U* and *non-U*. Two years later, Ross's essay was reprinted, with some modifications and a new title ('U and Non-U: an essay in sociological linguistics'), in *Noblesse Oblige,* which included contributions on the same subject by Nancy Mitford, Evelyn Waugh, and John Betjeman.

The essay's aim was to investigate the linguistic demarcation of the British upper class. *U* stood for 'upper class' usage; *non-U* stood for other kinds of usage. It looked at distinctive pronunciation and vocabulary, as well as written language conventions, such as how to open and close letters. It was a personal account containing many subjective judgments and disregarding the subtle gradations in usage intermediate between the two extremes; but it was also highly perceptive, drawing attention to a large number of distinguishing features. The nature of upper-class language has changed over 30 years later, but the terms *U* and *non-U* are still well known.

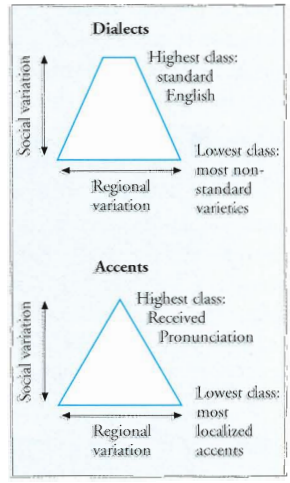
Some of the lexical oppositions proposed by Ross:

U	non-U
have a bath	take a bath
bike, bicycle	cycle
luncheon	dinner
riding	horse riding
sick	ill
knave	jack
mad	mental
looking-glass	mirror
writing-paper	note-paper
jam	preserve
wireless	radio
table-napkin	serviette
lavatory-paper	toilet-paper
rich	wealthy
vegetables	greens
pudding	sweet
telegram	wire
England	Britain
Scotch	Scottish

SOCIAL IDENTITY AND OTHER FACTORS

It is never possible to make a simple statement about language variation and social class because other influential factors are involved, such as the sex of the speaker, and the formality of the situation (p. 42). There is also an important interaction between social and regional factors (§8), as illustrated below for British English.

The two pyramids deal with differences of accent and dialect, and represent the relationship between 'where' a speaker is, both socially (the vertical dimension) and geographically (the horizontal dimension). At the top are the speakers of the highest social class: they speak the standard dialect with very little regional variation. Also at the top are those who speak Received Pronunciation (RP), the educated accent which signals no regional information at all (within Britain). The further we move down the class scale, the more we encounter regional accent and dialect variation. And when we reach the lowest social class, we encounter the widest range of local accents and dialects.



Thus, for example, speakers from the top social class will all use the same word *headache*, and give it the same (RP) pronunciation, but speakers from the lowest class will use *skullache, headwarch, sore head,* and other forms, in a variety of pronunciations, depending on where they are from. (After P. Trudgill, 1983.)

RESTRICTED AND ELABORATED CODES

Do people from different social classes display different abilities in their use of language? This was one of the questions widely discussed in the 1970s, as a result of a distinction proposed by the sociologist Basil Bernstein (1924–). The concepts of 'elaborated code' and 'restricted code' attempt to explain how a society's distribution of power and its principles of control shape and enter different modes of communication which carry the cultures of different social classes and that of the school, and so reproduce unequal educational advantages. The theory proposes that the sets of social relationships in which people are embedded act selectively on the production of meanings, and so upon choices within common linguistic resources.

Codes are said to have their origins in different family structures, associated (but not inevitably) with social classes, and are relayed through crucial socializing contexts, instructional and regulative, which differently orient children to the roles, meanings, and values of the school. Restricted codes arise where meanings are particular to and embedded in a local context, and the need to make meanings specific and explicit is reduced by the foregrounding of shared understandings, values, and identifications. By contrast, the forms of elaborated codes arise out of social relations where less is taken for granted, where shared understandings, values, and identifications are less foregrounded, and so where explicitness and specificity are more likely to be demanded. Middle-class children are said to have access to both codes, whereas lower-working-class children are more likely to be initially limited to a restricted code, and to experience difficulty in acquiring the form of the elaborated code required by the school, and thus the meanings and pedagogical practices regulated by that code.

The complexities of this theory were sometimes reduced to the proposition that middle-class children are able to abstract, but working-class children are not; this difference was then attributed to differences in the children's linguistic resources. Bernstein argues strongly that there is no basis for either of these propositions in his theory. Misreadings of the theory can also occur through a too-ready association of codes with language varieties or registers. Certainly, it is possible to show that a lower-class speaker can handle abstract concepts in restricted code. For example, in one of the recordings made by William Labov (1927–), a black 15-year-old was asked why he thought a God would be white. He replied: 'Why? I'll tell you why! Cause the average whitey out here got everything, you dig? And the nigger ain't got shit, y'know? Y'understan'? So – um – for in order for *that* to happen, you know it ain't no black God that's doin' that bullshit.' There is plainly abstract reasoning here, despite the non-standard language, and the restricted code.

Studies of this kind show that the correlation between the use of language and social class is evidently not simple: other factors intervene, such as the context in which learning takes place, and the way family life is structured. These factors always need to be borne in mind when debating levels of linguistic 'deficiency' or 'difference' between people of different social classes.

THE LANGUAGE OF RESPECT

Many communities make use of a complex system of linguistic levels in order to show respect to each other. The levels will partly reflect a system of social classes or castes, but the choice of forms may be influenced by several other factors, such as age, sex, kinship relationships, occupation, religious affiliation, or number of possessions. In Javanese, for example, choice of level can in addition be affected by the social setting of a conversation, its subject matter, or the history of contact between the participants. Other things being equal, people would use a higher level at a council meeting than in the street; in talking about religious matters than about buying and selling; and when addressing someone with whom they had recently quarrelled. Similar constraints have been noted for several languages, such as Japanese (p. 99), Korean, Tibetan, Samoan, and Sundanese.

Devices for conveying relative respect and social distance can be found in all languages. What is distinctive about 'respect' languages is the way differences of social level have been so extensively coded in the grammar and vocabulary. In Javanese, the differences between levels are so great that equivalent sentences may seem to have very little in common.

WOLOF GREETINGS

Greeting behaviour has a special place among the Wolof of Senegal, and well illustrates the link between language and social identity. Every interaction *must* begin with a greeting.

In the country, a greeting occurs between any two persons who are visible to each other – even if one person has to make a detour to accomplish it. In crowded areas, everyone close to the speaker must be greeted. In a conversational gathering, everyone must be greeted at the outset; and if, in the course of the conversation, someone leaves and then returns, it is often necessary to pause while all are greeted individually again.

Wolof society is divided into several castes, and a person's social identity is involved in every greeting. The most senior people present are greeted before those of lower rank; and in any meeting, those of lower rank must speak first. When two people meet, they must reach a tacit agreement about their relative status: the one who talks first accepts the lower role. Variations in status also occur. For example, an upper-caste person may not wish to adopt the higher-ranking position, because that would oblige him to support the lower-ranking person with a gift at some future point. He would therefore attempt to lower himself by speaking first in a conversation.

A Wolof proverb sums up this principle of social inequality: *sawaa dyi, sawaa dyi, gatyangga tya, ndamangga ca*, 'When two persons greet each other, one has shame, the other has glory'. (After J. T. Irvine, 1974.)

Level	are	you	going	to eat	rice	and	cassava	now	Complete
krama inggil	*menapa*	*pandjenengan*	*badé*	*ḍahar*	*sekul*	*kalijan*	*kaspé*	*samenika*	*Menapa pandjenengan badé ḍahar sekul kalijan kaspé samenika?*
krama biasa		*sampéjan*		*neḍa*					*Menapa sampéjan badé neda sekul kalijan kaspé samenika?*
madya	*napa*		*adjeng*			*lan*		*saniki*	*Napa sampéjan adjeng neḍa sekul lan kaspé saniki?*
ngoko madya	*apa*		*arep*		*sega*			*saiki*	*Apa sampéjan arep neḍa sega lan kaspé saiki?*
ngoko biasa		*kowé*		*mangan*					*Apa kowé arep mangan sega lan kaspé saiki?*

Five status levels, in one Javanese dialect (after C. Geertz, 1968), using the sentence *Are you going to eat rice and cassava now?* The names *krama*, *madya*, and *ngoko* refer to 'high', 'middle', and 'low' respectively. In addition, the high and low levels each have two divisions, depending on whether honorific words are used, to produce *krama inggil* vs *krama biasa*, and *ngoko madya* vs *ngoko biasa*.

SOCIAL STATUS AND ROLE

'Status' is the position a person holds in the social structure of a community – such as a priest, an official, a wife, or a husband. 'Roles' are the conventional modes of behaviour that society expects a person to adopt when holding a particular status. Public roles often have formal markers associated with them, such as uniforms; but among the chief markers of social position is undoubtedly language. People exercise several roles: they have a particular status in their family (head of family, first-born, etc.), and another in their place of work (supervisor, apprentice, etc.); they may have a third in their church, a fourth in a local sports centre, and so on. Each position will carry with it certain linguistic conventions, such as a distinctive mode of address, an 'official' manner of speech, or a specialized vocabulary. During the average lifetime, people learn many such linguistic behaviours.

It is only occasionally that the adoption of a social role requires the learning of a completely different language. For instance, a knowledge of Latin is required in traditional Roman Catholic practice; a restricted Latin vocabulary was once prerequisite for doctors in the writing out of prescriptions; students in some schools and colleges still have to speak a Latin grace at mealtimes; and Latin may still be heard in some degree ceremonies. More usually, a person learns a new *variety* of language when taking up a social role – for example, performing an activity of special significance in a culture (such as at a marriage ceremony or council meeting), or presenting a professional image (as in the case of barristers, the police, and drill sergeants). The use of new kinds of suprasegmental feature (§29) is particularly important in this respect. One of the most distinctive indications of professional role is the intonation, loudness, tempo, rhythm, and tone of voice in which things are said.

In many cases, the linguistic characteristics of social roles are fairly easy to identify; but often they are not, especially when the roles themselves are not clearly identifiable in social terms. With unfamiliar cultures and languages, too, there is a problem in recognizing what is really taking place in social interaction or realizing how one should behave when participating in an event. How to behave linguistically as a guest varies greatly from culture to culture. In some countries, it is polite to comment on the excellence of a meal, as one eats it; in others, it is impolite to do so. In some countries, a guest is expected to make an impromptu speech of thanks after a formal meal; in others there is no such expectation. Silence, at times, may be as significant as speech (p. 38).

CEREMONIAL LANGUAGE

Probably all communities have developed special uses of language for ritual purposes. Distinctive forms are employed by those who have official status in the ceremony, as well as by those who participate. This may extend to the use of totally different languages (without regard for listener intelligibility), or be no more than selective modifications of everyday speech – such as prayers and speeches that are distinguished only by a more careful articulation, abnormal prosody, and the occasional use of exceptional vocabulary and grammatical forms.

Among the Zuñi, for example, 'sacred words' (*téwusu péna· we*), usually prayers, are pronounced in rhythmical units, resembling the lines of written poetry, with a reversal of the expected patterns of stress and intonation: strongly stressed syllables become weak, and the weakest syllable in the unit is pronounced most strongly. Ceremonial speech among the Kamsá Indians of Colombia also involves distinctive intonation and timing, reminiscent of chant, but in addition there are grammatical and lexical changes. They use many more Spanish loan words than in everyday speech (60%, compared with 20%), and there is a marked increase in the number of affixes in a word (as many as 11 attached to a root, compared to the six or fewer heard in ordinary use).

Often, ceremonial genres are marked by considerable verbal ingenuity. For example, among the Ilongot of the northern Philippines there is a speech style known as 'crooked language' (*qambàqan*), used in oratory, play, song, riddles, and public situations, such as debates. It is a style rich in witty repartee, puns, metaphor, elaborate rhythms, and changes in words. In Malagasy, there is a contrast between everyday talk (*resaka*) and oratorial performance (*kabary*), which is used in ceremonial situations such as marriages, deaths, and bone-turnings, and also in formal settings, such as visits. An obligatory feature of *kabary* is 'winding' speech, in which male speakers perform a dialogue in a roundabout, allusive manner, using many stylistic devices, such as metaphors, proverbs, and comparisons. The genre uses traditional ways of speech, handed down from ancestors. To speak Malagasy well means to be in command of this style; and it is common to hear speakers' abilities discussed and evaluated.

In a marriage request ceremony, for example, the girl's family gather in her village, and await the arrival of the boy's family. Each is represented by a speech-maker. As the boy's family approaches, no official notice is taken of them until their speech-maker makes a series of requests to enter the village. Unless the girl's speech-maker judges that these speeches are performed adequately, according to the traditional standards of the *kabary*, they will not be allowed to proceed to the formal marriage request, and the speech-maker must redouble his efforts. Subsequent steps in the ceremony are evaluated in the same way. (After E. Keenan, 1974.)

Kabary in progress An orator at a Malagasy bone-turning ceremony.

SOCIAL SOLIDARITY AND DISTANCE

One of the most important functions of language variation is to enable individuals to identify with a social group or to separate themselves from it. The markers of solidarity and distance may relate to family, sex (p. 46), ethnicity, social class (p. 38), or to any of the groups and institutions that define the structure of society. They may involve tiny sections of the population, such as scout groups and street gangs, or complete cross-sections, such as religious bodies and political parties. The signals can be as small as a single word, phrase, or pronunciation, or as large as a whole language.

DIFFERENT LANGUAGES

Probably the clearest way people have of signalling their desire to be close to or different from those around them is through their choice of languages. Few societies are wholly monolingual, and it is thus possible for different languages to act as symbols of the social structure to which their speakers belong. The test sentence 'If they speak LANGUAGE NAME, they must be — ' can be completed using geographical terms (p. 24), but social answers are available as well: the blank can be filled by such phrases as 'my tribe', 'my religion', 'immigrants', 'well educated', 'rich', 'servants', and 'the enemy'.

The use of a different language is often a sign of a distinct religious or political group – as in the cases of Basque, Latin, Welsh, the many official languages of the Indian sub-continent, and the pseudolinguistic speech known as glossolalia (p. 11). Switching from one language to another may also be a signal of distance or solidarity in everyday circumstances, as can be seen in strongly bilingual areas, such as Paraguay. Here, the choice of Spanish or Guaraní is governed by a range of geographical and social factors, among which intimacy and formality are particularly important. In one study (J. Rubin, 1968), bilingual people from Itapuami and Luque were asked which language they would use in a variety of circumstances (e.g. with their spouse, sweetheart, children, boss, doctor, priest, etc.). For most, Guaraní was the language of intimacy, indicating solidarity with the addressee. The use of Spanish would indicate that the speaker was addressing a mere acquaintance or a stranger. Spanish was also the language to use in more formal situations, such as patient–doctor, or student–teacher. Jokes would tend to be in Guaraní. Courtship often began in Spanish, and ended in Guaraní.

The adoption of a local language as an emblem of group identity is well illustrated by the Vaupés Indians of Colombia, who live in more than 20 tribal units, each of which is identified by a separate language. Despite the existence of a lingua franca (Tukano), a homogeneous culture throughout the region, and the small numbers of speakers (around 5,000 in total, in the early 1960s), the Indians all learn at least three languages – some, as many as ten. The identity of the different languages is sharply maintained – for instance, several places have separate names in all the languages, and the Indians themselves emphasize their mutual unintelligibility. In such circumstances, the languages act as badges of membership of the tribal units. An Indian will often speak initially in his own father language to acknowledge publicly his tribal affiliation. And language acts as a criterion for all kinds of social behaviour. For example, when the investigator asked a Bará Indian about marriage sanctions, she was told: 'My brothers are those who share a language with me. Those who speak other languages are not my brothers, and I can marry their sisters.' On another occasion, when she asked an Indian why they spoke so many languages instead of using the lingua franca, she received the reply: 'If we were all Tukano speakers, where would we get our women?' (After J. Jackson, 1974.)

DIFFERENT VARIETIES

In monolingual communities, a major way of marking factors such as solidarity, distance, intimacy, and formality is to switch from one language variety to another. A Berlin business manager may use standard German at the office and lapse into local dialect on returning home. A conference lecturer in Paris may give a talk in formal French, and then discuss the same points with colleagues in an informal variety. A London priest may give a sermon in an archaic, poetic style, and talk colloquially to the parishioners as they leave. During the service, the priest might have used a modern English translation of the Bible, or one which derives from the English of the 16th century.

Languages have developed a wide range of varieties for handling the different kinds and levels of relationship which identify the social structure of a community. These varieties are discussed in other sections (§§11, 63), because they partly reflect such factors as occupation, subject matter, social status, and setting; but it is important to note that they may also be used as symbols of social identity. In English, for example, forms such as *liveth and reigneth, givest, vouchsafe,* and *thine* have long been distinctive in one variety of religious language; but in the 1960s, as proposals for the modernization of Christian liturgical language were debated, this variety came to be seen as a symbol of traditional practice with which people chose to identify or from which they dissociated themselves. The case is worth citing because the world-wide status of Christianity meant that many speech communities were involved, and over a quarter of the world's population was affected. No other linguistic change can ever have raised such personal questions of linguistic identity on such a global scale.

AVOIDANCE LANGUAGES

Among Australian aborigines, it is common for a man to 'avoid' certain relatives – often his wife's mother and maternal uncles, sometimes her father and sisters as well. Brothers and sisters, too, may not be allowed to converse freely, once they grow up. In some tribes, avoidance of taboo relatives means total lack of contact; in others, a degree of normal speech is tolerated; but the most interesting cases are those where special languages have developed to enable communication to take place. These are usually referred to as 'mother-in-law' languages, but all taboo relatives are included under this heading.

In Dyirbal (now almost extinct), the everyday language is known as Guwal, and the mother-in-law language is called Dyalnguy. The latter would be used whenever a taboo relative was within earshot. The two languages have virtually the same grammar, but no vocabulary in common. Dyalnguy also has a much smaller vocabulary than Guwal.

In Guugu-Yimidhirr, there is no contact at all with the mother-in-law, and a strong taboo also affects speech to brothers- and fathers-in-law. There are important differences in vocabulary, style, and prosody. Sexual topics are proscribed. One must speak to these relatives slowly, in a subdued tone, without approaching closely or facing them. The style is sometimes described as *dani-man-aarnaya,* 'being soft/slow', or *diili yirrgaalga,* speaking 'sideways'. (After J. B. Haviland, 1979.)

The avoidance languages of Australia illustrate yet another means of marking social distance. The people turn away, linguistically and physically, from their taboo relatives. Similar taboos have also been observed in many other parts of the world, such as among the Plains Indians of North America. These languages can therefore be contrasted with those (in South-east Asia, for example) where social relations are expressed by adding complexity to ordinary speech (p. 40).

Diglossia

Perhaps the clearest use of varieties as markers of social structure is in the case of *diglossia* – a language situation in which two markedly divergent varieties, each with its own set of social functions, coexist as standards throughout a community. One of these varieties is used (in many localized variant forms) in ordinary conversations; the other variety is used for special purposes, primarily in formal speech and writing. It has become conventional in linguistics to refer to the former variety as 'low' (L), and the latter as 'high' (H).

Diglossic situations are widespread, some of the better-known ones including Arabic, Modern Greek, and Swiss German. These speech communities recognize the H/L distinction and have separate names for the two varieties:

	High	Low
Greek	Katharévousa	Dhimotiki (Demotic)
Arabic	ʾal-fuṣḥā (Classical)	ʾal-ʿāmmiyyah (Colloquial)
Swiss German	Hochdeutsch (High German)	Schweitzer-deutsch (Swiss German)

The functional distinction between H and L is generally clear-cut. H is used in such contexts as sermons, lectures, speeches, news broadcasts, proverbs, newspaper editorials, and traditional poetry. It is a language that has to be learned in school. L is used in everyday conversation and discussion, radio 'soap operas', cartoon captions, folk literature, and other informal contexts.

H and L varieties can display differences in phonology, grammar, and vocabulary. For example, the sound systems of the two Swiss German varieties are strikingly different. Classical Arabic has three noun cases, whereas Colloquial Arabic has none. And in Greek there are many word pairs, such as *ínos* (H) and *krasí* (L) ('wine'): the H word would be written on Greek menus, but diners would ask for their wine using the L word. All three kinds of distinctiveness are illustrated in the following sentence given first in Hochdeutsch (H) and then in Schweitzerdeutsch (L): *Nicht nur die Sprache hat den Ausländer verraten, sondern auch seine Gewohnheiten;* and *Nüd nu s Muul häd de Ussländer verraate, au syni Möödeli.* 'It was not only his language that showed he was a foreigner, his way of life showed it too.' (After P. Trudgill, 1983.)

In diglossic situations, the choice of H vs L can easily become an index of social solidarity. A Swiss German speaker who used Hochdeutsch in everyday conversation would be considered snobbish or artificial – and if the context were a political discussion, it could even raise questions of national loyalty, as Hochdeutsch is used as the everyday language by people outside the country. Religious as well as political attitudes may be involved. The H form is often believed to be the more beautiful and logical, and thus the more appropriate for religious expression – even if it is less intelligible. In Greece, there were serious riots in 1903, when the New Testament was translated into Dhimotiki. And strong views are always expressed by Arabic speakers about Classical Arabic, which, as the language of the Qur'an, belongs to God and heaven (p. 388).

Diglosssic situations become unstable in the face of large-scale movements for a single standard – such as might be found in programmes of political unification, national identity, or literary reform. In such circumstances, there are arguments in favour of either H or L varieties becoming the standard. Supporters of H stress its link with the past, and its claimed excellence, and they contrast its unifying function with the diversity of local dialects. Supporters of L stress the need to have a standard which is close to the everyday thoughts and feelings of the people, and which is a more effective tool of communication at all levels. 'Mixed' positions, setting up a modified H or L, are also supported; and the steady emergence of L-based standards has been noted in Greece, China, Haiti, and several other areas.

A personal column from the Basel daily newspaper *Basler Zeitung* This item shows an interesting contrast between High German and Swiss German. The rest of the newspaper is written in High German, but in the *Perseenlig* column (High German *persönlich),* the last two items are entirely in Swiss German (apart from the words in English). One is a humorous announcement of the opening of a medical practice; the other is a birthday greeting.

Why are the remaining ads not in Swiss German? This is probably because of their content and level: the first item expresses the thanks of an old married couple to their neighbours for all they did at their golden wedding celebration; the second announces the assembly point and time for a meeting of the fire service association. Even so, the second item has one distinctive feature: *Besammlung* ('meeting') is an example of 'Swiss High German', midway between High German (*Versammlung)* and Swiss German (*Besammlig).*

DIFFERENT WORDS AND PHRASES

We recognize varieties of language as a result of perceiving several distinctive linguistic features being used together in a social situation. But often a single linguistic feature is enough to indicate social distance – such as the particular words or phrases used when people meet, address each other by name, or select pronouns for talking to or about each other.

Modes of address

One of the most significant ways of signalling social intimacy and distance is through the use of a person's name in direct address. In English, the basic choice is between first name (FN) or title with last name (TLN), but several other conventions are possible in certain contexts, such as the use of LN in business or military settings (*Come in, Smith ...*), or the use of abbreviations (*Is JM in?*). The range of possible forms is easy to state; but the factors that govern the choice of forms are often complex and difficult to summarize. When would two people use FNs or TLNs reciprocally to each other? When would one speaker use FN and the other TLN?

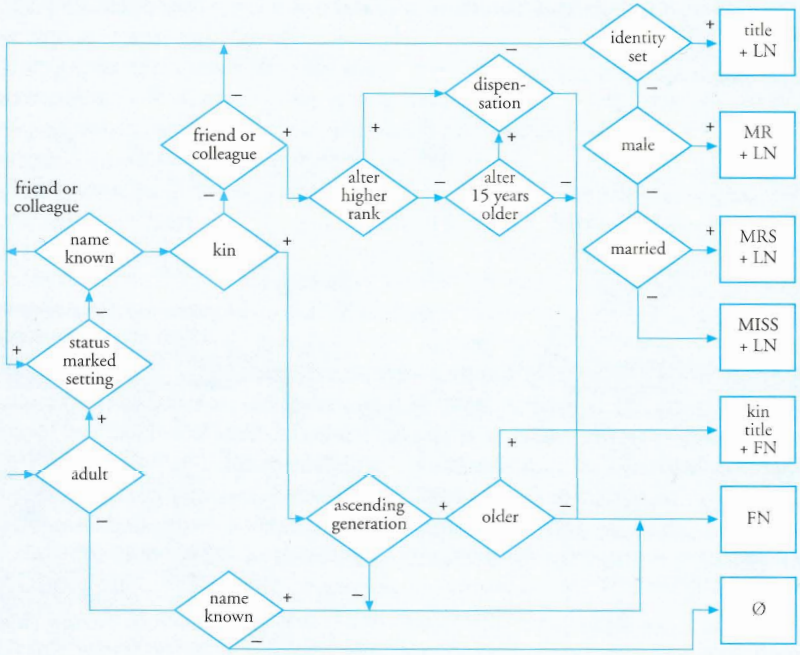
Charting address relationships Several studies have attempted to explicate these factors. The flowchart (right) was devised by Susan Ervin-Tripp (1927–) as a means of specifying the factors that condition a speaker's choice of address in American English. The chart is simply a logical statement of the various possibilities, given a context such as 'Look, — , it's time to leave'; it is not an account of what goes on in the speaker's mind. The knowledge structure represented is that of an American academic; but dialect differences, idiosyncratic preferences, and other variants are not taken into account.

The entrance point to the diagram is at the bottom left. Each path through the diagram leads to one of the possible modes of address, listed vertically at the right. Alternative realizations of these address modes are not given (e.g. a first name may alternate with a nickname). For example, as one enters the diagram, the first choice which has to be made is whether the addressee is a child (– Adult) or an adult (+ Adult). If the former, one follows the line downwards, where the only distinction drawn is that between name known (+) or not (–). If the child's name is known, one uses the first name; if not, one does not use a name at all (Ø). The diagram does not give criteria for deciding when a child becomes an adult.

Along the adult path, several decisions have to be made. 'Status-marked setting' refers to special occasions (such as a courtroom) where forms of address are rigidly prescribed (e.g. *your honour, Mr Chairman*). The 'identity set' refers to the list of occupational or courtesy titles that may be used alone to mark social identity (e.g. *Father, Doctor, Mr, Miss*).

In addressing people whose names are known, kinship is a major criterion. If the speaker is related to the addressee ('alter'), two factors are relevant: 'ascending generation' (e.g. aunt as opposed to cousin) and age. If the speaker is not related to alter, the factor of familiarity is relevant: whether or not alter is a friend or colleague. If familiarity applies, the next factor is social rank, here defined with reference to a professional hierarchy. A senior alter has the option of offering or accepting FN, instead of TLN ('dispensation' – *Call me Mike*), though this situation is often ambiguous. Age difference is not significant until there is a gap of nearly a generation.

American address system (after S. Ervin-Tripp, 1972).



NUER MODES OF ADDRESS

Address systems vary greatly from culture to culture. Among the Nuer (Sudan), a system of multiple names and titles marks a person's place in social structure. Every Nuer is given a personal name, shortly after birth, which he retains through life; but as an adult, it is used only by close relatives and friends. These names usually refer to the place of birth, or to events that took place at the time, such as *Nhial* 'rain', *Duob* 'path'. Maternal grandparents often give the child a second personal name, which is used by kinsfolk on the maternal side. Twins are given special personal names, which immediately identify their status, such as *Both* 'the one who goes ahead' and *Duoth* 'the one who follows'.

The social setting is an important factor in the selection of a mode of address. Every child inherits an honorific, or clan name, which tends to be used only in ceremonies or on special occasions (such as a return after a long absence). When a boy is initiated to manhood, he is given an ox, and from the distinctive features of this animal he takes his 'ox-name', which is used only by people of the same or similar ages. There are also 'dance-names' – more elaborate versions of ox-names that are used only at dances.

Kinship roles also play their part. A man would normally be addressed using the name of his father (his patronymic). But a man visiting maternal relatives will be greeted primarily by his mother's name (his matronymic). The naming of people after their eldest child (teknonymy) is also heard, especially when talking to in-laws. For example, a woman's status in her husband's home is based on her having borne him a child, and this is the link that binds her to her husband's social group. It is therefore natural for that group to address her using the child's name. (After E. E. Evans-Pritchard, 1948.)

T or V?

A well-studied example of address is the use of the familiar and polite pronouns found in many languages, as in French *tu/vous*, German *du/Sie,* Welsh *ti/chwi,* and so on. These forms (generally referred to as *T* forms and *V* forms, respectively, from Latin *tu* and *vos*) follow a complex set of rules that foreigners never find easy to master. Terms such as 'familiar' and 'polite' capture aspects of their use, but are inadequate summaries of all their social functions, and ignore important differences between languages.

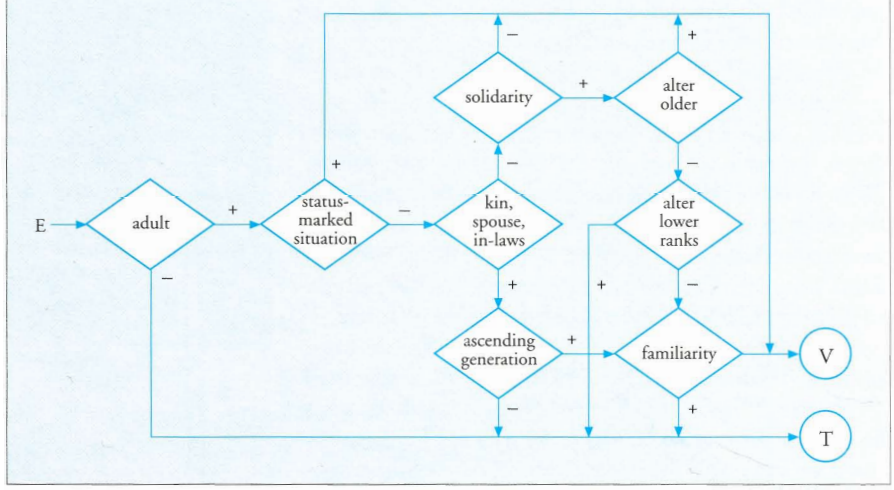
In Latin, the T forms were used for addressing one person, and the V forms for more than one; but from around the 4th century AD, the convention developed of referring to the Roman Emperor using the plural form *vos*. Gradually, this 'royal *you*' extended to others who exercised power, so that by medieval times, the upper classes were showing mutual respect through the use of V forms only. The historical picture is complicated and not entirely understood, but medieval nobles would generally address each other as V, whether talking to one person or more than one, and would address the lower classes as T. By contrast, the lower classes would use T to each other, and V to their superiors.

Later the V forms began to be used in other circumstances, not simply as a mark of respect due to those with power but as a sign of any kind of social distance. T forms, correspondingly, began to be used as markers of social closeness and intimacy. Thus, between equals, it became possible to use either T or V, depending on the degree of solidarity one wished to convey. Lower-class friends would address each other as T, and use V to strangers or acquaintances. Upper-class people would do likewise.

In these circumstances, where there is a power relationship motivating one usage (T = lack of respect), and a solidarity relationship motivating another (T = social closeness), situations of uncertainty would often arise. For example, during a meal, should diners address servants as T or V? The diners are more 'powerful' (and so should use T), but they are also socially distant from the servants (and so should use V). Similarly, should children address their parents as T (because they are intimates) or V (because there is a power difference)? By the 20th century, such conflicts had in most cases been resolved by following the dictates of the solidarity dimension: these days, diners address waiters as V, and children address parents as T.

But some fascinating differences remain. In the first systematic T/V study, male students from different linguistic backgrounds were asked about their pronoun preferences. The sample was relatively small, but it clearly emerged that Italians used T more than the French, and the French more than the Germans. There were several interesting points of detail: for example, Germans used T more to distant relations than did the French; Italians were more likely to use T to fellow female students than either French or Germans. There were psychological as well as geographical differences. Radical students used more T forms than did conservatives. One of the conclusions of the study was that 'a Frenchman could, with some confidence, infer that a male university student who regularly said T to female fellow students would favour the nationalization of industry, free love, trial marriage, the abolition of capital punishment, and the weakening of nationalistic and religious loyalties'. Inferences like these are difficult to confirm on a larger scale, partly because of the speed of linguistic change (since the early 1960s, when this study was done, student use of T has become much more widespread). But hypotheses of this kind are well worth following up, as they bear directly on the task of establishing the basis of sociolinguistic identity. (After R. Brown & A. Gilman, 1968.)

Flow-charts These charts provide an opportunity to make hypotheses about naming practice precise, and help to clarify interlanguage differences. For example, this kind of diagram has been used to identify the factors governing the use of T or V forms in Yiddish (S. Ervin-Tripp, 1972).



Farr's Law of Mean Familiarity

... as discovered by Lumer Farr, one of the senior lifemen in Stephen Potter's *One-upmanship* (1952), identifies a well-known inverse naming relationship in the following way:

The Guv'nor addresses:	
Co-director Michael Yates as	Mike
Assistant director Michael Yates as	Michael
Sectional manager Michael Yates as	Mr Yates
Sectional assistant Michael Yates as	Yates
Indispensable secretary Michael Yates as	Mr Yates
Apprentice Michael Yates as	Michael
Night-watchman Michael Yates as	Mike

SEXISM

The relationship between language and sex has attracted considerable attention in recent years, largely as a consequence of public concern over male and female equality. In many countries, there is now an awareness, which was lacking a generation ago, of the way in which language can reflect and help to maintain social attitudes towards men and women. The criticisms have been directed almost exclusively at the linguistic biases that constitute a male-orientated view of the world, fostering unfair sexual discrimination, and, it is argued, leading to a denigration of the role of women in society. English has received more discussion than any other language, largely because of the impact of early American feminism.

Several areas of grammar and vocabulary have been cited. In grammar, the issue that has attracted most attention is the lack of a sex-neutral, third-person singular pronoun in English, especially in its use after indefinite pronouns, e.g. *If anyone wants a copy, he can have one.* (In the plural, there is no problem, for *they* is available.) No natural-sounding option exists: *one* is considered very formal, and forms such as *he or she* are stylistically awkward. As a result, there have been many proposals for the introduction of a new English sex-neutral pronoun – including *tey, co, E, ne, thon, mon, heesh, ho, hesh, et, hir, jhe, na, per, xe, po,* and *person.* None of these proposals has attracted widespread support, but *co,* for example, has been used in some American communes, and *na* and *per* have been used by some novelists. Less radical alternatives include advice to restructure sentences to avoid the use of *he*-forms.

Many other examples of linguistic bias have been given. In the lexicon, particular attention has been paid to the use of 'male' items in sex-neutral contexts, such as *man* in generic phrases (*the man in the street, stone-age man,* etc.), and the potential for replacing it by genuinely neutral terms (*chairman* → *chairperson, salesman* → *sales assistant,* etc.). Another lexical field that is considered problematic is marital status, where bias is seen in such phrases as X's *widow* (but not usually *Y's widower),* the practice of changing the woman's surname at marriage, and the use of *Mrs* and *Miss* (hence the introduction of *Ms* as a neutral alternative). The extent of the bias is often remarked upon. In one computer analysis of child school books, male pronouns were four times as common as female pronouns. In another study, 220 terms were found in English for sexually promiscuous women, and only 22 for sexually promiscuous men. It is easy to see how sexual stereotypes would be reinforced by differences of this kind.

THE PROPER STUDY OF MANKIND IS MAN?

What has happened to sexist language, as a result of feminist criticism? So far, the effect has been far more

MAINTAINING SEXUAL STEREOTYPE LANGUAGE

This was the list of lecturers from the University of Reading's Department of Linguistic Science in 1983, as printed in the University calendar. Although gender is irrelevant to the job, the women in the Department were clearly identified by the use of a full first name, and/or by the use of *Mrs.* It is not possible to tell if the male members of staff are married.

Lecturers:

C. Biggs, *MA, Oxford; PhD Cambridge; Diploma in Linguistics, Cambridge*

R. W. P. Brasington, *MA, Oxford*

A. R. Butcher, *MA, Edinburgh; MPhil, London; Dr phil, Kiel*

F. Margaret Davison, *BA, Sussex; MA, Reading; Cert T Deaf, Manchester*

P. J. Fletcher, *BA, Oxford; MPhil, Reading; PhD, Alberta*

M. A. G. Garman, *BA, Oxford; PhD Edinburgh; Diploma in General Linguistics, Edinburgh*

G. A. Hughes, *BA, Montreal; Diploma in English as Second Language, Wales*

K. Johnson, *BA, Oxford; MA, Essex*

Carolyn A. Letts (Mrs Letts), *BA, Wales; MCST*

K. M. Petyt, *MA, Cambridge; MA, PhD, Reading; Diploma in Public and Social Administration, Oxford (Director of Extramural and Continuing Education)*

Marion E. Trim (Mrs Trim), *MSc, London; LCST*

Irene P. Warburton (Mrs Warburton), *BA, Athens; PhD, Indiana*

SEXIST LANGUAGE

People would bring their wives, mothers, and children.
Rise Up, O Men of God ...
Man, being a mammal, breastfeeds his young.
Mind that child – he may be deaf!
Man overboard!

These randomly selected cases of sexist language may provoke ridicule, anger, or indifference, but they would be unlikely to warrant a legal action to determine their meaning. However, there are other examples where a legal decision could hang on the sex-specific vs sex-neutral senses of *man*. In the USA, for example, there has been legal controversy over the application of the generic male pronoun in cases where it was disputed whether such phrases as 'a reasonable man' could legitimately be applied to women. And in a case heard in 1977, an appeal was made against a woman's murder conviction on the grounds that instructions to the jury were phrased using the generic male form; this, it was argued, could have biased the jury's response, giving them the impression that the objective standard to be applied was that applicable to an altercation between two men. Traditional safeguard phrases such as 'the masculine pronoun shall import the feminine' have turned out to be less than satisfactory in resolving such issues.

SEX-ROLE STEREOTYPING IN SCHOOLBOOKS

Sexual stereotyping has been especially noted in traditional children's reading books and textbooks. There were always more male characters than female, and they took part in a greater variety of roles and activities. In early reading books, it was always the boys who were daring, the girls who were caring. Pictures in science books would show experiments being conducted by boys, while girls looked on. There is now a widespread trend to avoid sex-role stereotypes in children's books, and to prepare children for a more egalitarian society.

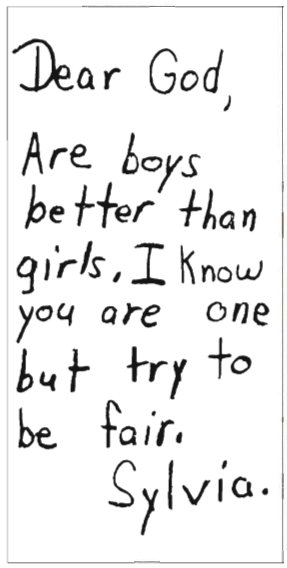
noticeable in writing than in speech. Several publishing companies have issued guidelines about ways of avoiding its use, and several writers and editors, in many important areas, now make a conscious effort to avoid unintentional biases – including such well-known bodies as the American Library Association, and writers such as Dr Benjamin Spock and the present author. Legal changes, such as the Sex Discrimination Act in Britain (1975), have caused job titles and much of the associated language to be altered. But is there any evidence of a significant change in practice throughout the language as a whole.

In 1984, an American study investigated the use of *man* and its compounds to refer to all humans, and the use of *he* and its inflected forms to refer to females as well as males, in a selection of publications taken at intervals between 1971 and 1979. The texts were samples of 75,000 running words from American women's magazines, science magazines, several newspapers, and both prepared and spontaneous remarks from the *Congressional Record;* a sample from *The Times Literary Supplement* was used, as a British comparison. The total sample was over half a million words.

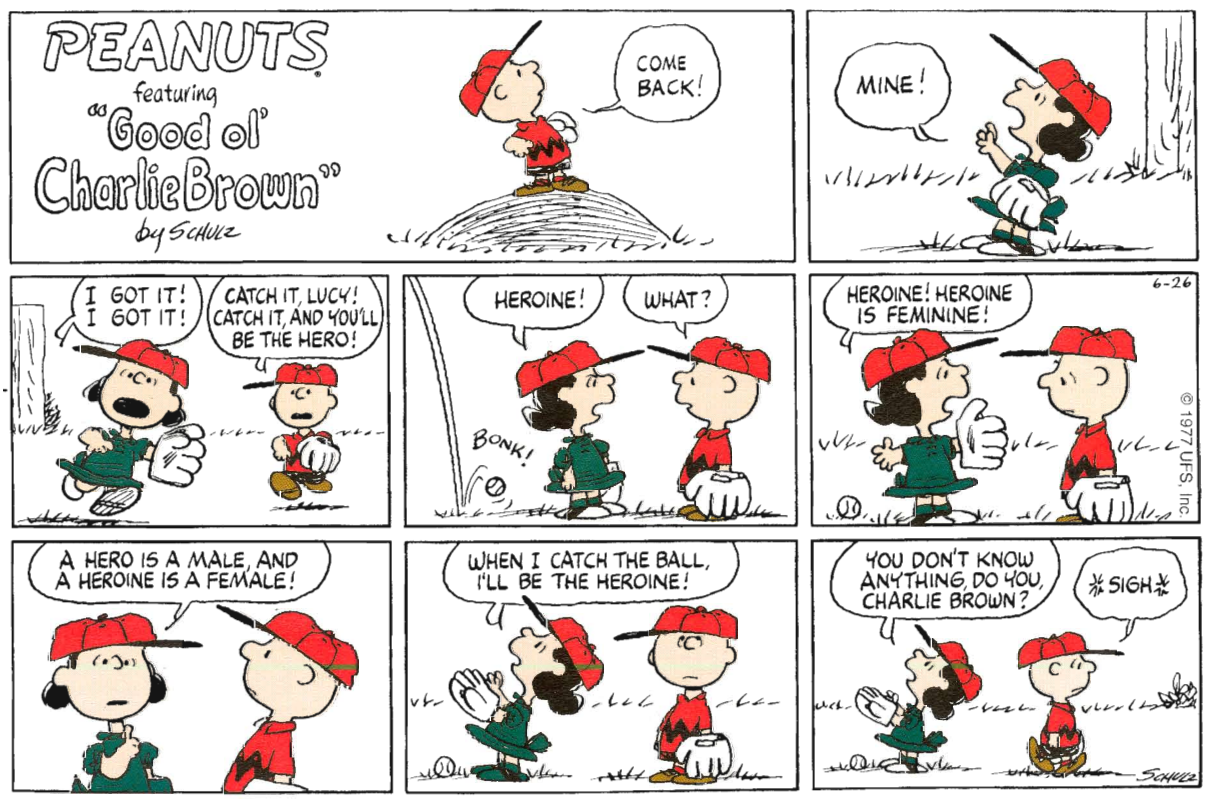
The results were dramatic. In the American corpus, the use of these forms fell from 12.3 per 5,000 words in 1971 to 4.3 per 5,000 in 1979. Women's magazines showed the steepest decline, followed by science magazines. By contrast, results for congressmen showed no decline at all, and results for congresswomen were mixed. There was no clear decline in the British publication, but rates were very low, and little can be deduced from such a small sample. (After R. L. Cooper, 1984.)

What took the place of these forms? There was no evidence that a straightforward replacement by such forms as *he or she* was taking place. Rather, it seems likely that people were using alternative linguistic devices to get round the problem, such as *they* along with a plural noun. (This is the solution I have found most congenial in the present work, in fact.)

There is thus clear evidence that the feminist movement had an observable impact in the 1970s on several important genres of written language – publications aimed at general audiences, not solely at women. Plainly, there has been a general raising of consciousness about the issue of linguistic sexism, at least as regards the written language. Whether this same consciousness would be found in everyday speech is unclear, as is the question of how long-term these linguistic effects will be. A great deal of social change has taken place in two decades, and this could be enough to make the associated linguistic changes permanent; but a decade or two is as nothing within the large time-scale of language change, and it remains to be seen whether the new trends in usage will continue, or whether there will be a reversal, with public opinion reacting against the extreme positions taken by some militant feminists.



Child's letter from *Children's Letters to God.*



14 · TYPOLOGY AND UNIVERSALS

The languages of the world present us with a vast array of structural similarities and differences. Why should this be so? One way of answering this question is to adopt a historical perspective, investigating the origins of language, and pointing to the importance of linguistic change – a perspective that is discussed in Part IX. An alternative approach is to make a detailed description of the similarities or differences, regardless of their historical antecedents, and proceed from there to generalize about the structure and function of human language.

There are two main ways of approaching this latter task. We might look for the structural features that all or most languages have in common; or we might focus our attention on the features that differentiate them. In the former case, we are searching for language *universals*; in the latter case, we are involving ourselves in language *typology*. In principle, the two approaches are complementary, but sometimes they are associated with different theoretical conceptions of the nature of linguistic enquiry.

SIMILARITY OR DIFFERENCE?

Since the end of the 18th century, the chief concern has been to explain the nature of linguistic diversity. This was the focus of comparative philology and dialectology, and it led to early attempts to set up genetic and structural typologies of languages (§50). The emphasis carried through into the 20th century when the new science of linguistics continually stressed the variety of languages in the world, partly in reaction against the traditions of 19th-century prescriptivism, where one language, Latin, had been commonly regarded as a standard of excellence (§1).

Since the 1950s, the focus on diversity has been replaced by a research paradigm, stemming from the work of the American linguist Noam Chomsky (1928–), in which the nature of linguistic universals holds a central place. Chomsky's generative theory of language proposes a single set of rules from which all the grammatical sentences in a language can be derived (p. 97). In order to define these rules in an accurate and economical way, a grammar has to rely on certain general principles – abstract constraints that govern the form it takes and the nature of the categories with which it operates. In this approach, these principles are conceived as universal properties of language – properties that are biologically necessary and thus innate (p. 236). The notion of universals is important, it is argued, not only because it deepens our understanding of language in its own right, but because it provides an essential first step in the task of understanding human intellectual capacity.

In Chomsky's view, therefore, the aim of linguistics is to go beyond the study of individual languages, to determine what the universal properties of language are, and to establish a 'universal grammar' that would account for the range of linguistic variation that is humanly possible. The question is simply: What are the limits on human language variability? Languages do not make use of all possible sounds, sound sequences, or word orders. Can we work out the reasons? It might be possible to draw a line between the patterns that are essential features of language, and those that no language ever makes use of (p. 97). Or perhaps there is a continuum between these extremes, with some features being found in most (but not all) languages, and some being found in very few. Questions of this kind constitute the current focus of many linguists' attention.

EXPRESSING COMPARISON

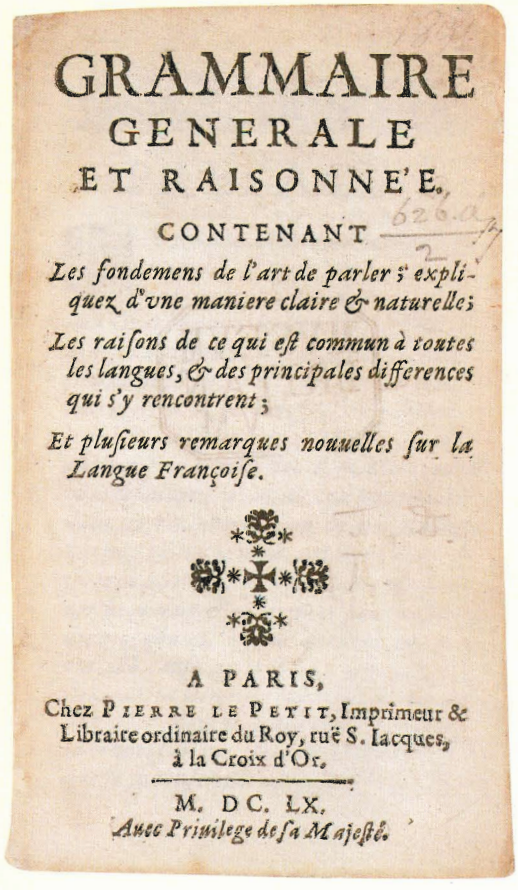
The English comparative construction, 'X is *bigger than* Y' involves three parts: the adjective (*big*), the markers of comparison (*-er* and *than*), and the standard of comparison (*Y*). This way of putting it is shared by many languages, including Berber, Greek, Hebrew, Malay, Maori, Songhai, Swahili, Thai, Welsh, and Zapotec.

However, the opposite order, in which the standard of comparison is expressed first, is also common. In Japanese, for example, it is 'Y *yori okii*' (literally 'Y than big'), and this way of putting it is shared by Basque, Burmese, Chibcha, Guaraní, Hindi, Kannada, and Turkish, among others. Finnish is a language which uses both constructions.

THE PORT-ROYAL GRAMMAR

Contemporary ideas about the nature of linguistic universals have several antecedents in the work of 17th-century thinkers. The *Grammaire générale et raisonnée* (1660) is widely recognized as the most influential treatise of this period. It is often referred to as the 'Port-Royal grammar', because it was written by scholars who belonged to the community of intellectuals and religious established between 1637 and 1660 in Port-Royal, Versailles.

Although published anonymously, the authorship of the grammar has been ascribed to Claude Lancelot (1615–95) and Antoine Arnauld (1612–94). Its subtitle, referring to 'that which is common to all languages, and their principal differences ...' provides a neat summary of the current preoccupation with universals and typology. However, the approach of modern linguistics is less concerned with how language relates to logic and reality, and more with its arbitrary properties.


GRAMMAIRE
GENERALE
ET RAISONNE'E,
CONTENANT
Les fondemens de l'art de parler; expliquez d'vne maniere claire & naturelle;
Les raiſons de ce qui eſt commun à toutes les langues, & des principales differences qui s'y rencontrent;
Et pluſieurs remarques nouuelles ſur la Langue Françoiſe.
A PARIS,
Chez PIERRE LE PETIT, Imprimeur & Libraire ordinaire du Roy, ruë S. Iacques, à la Croix d'Or.
M. DC. LX.
Auec Priuilege de ſa Majeſté.

BREADTH OR DEPTH?

The distinction between typological and universalist approaches to language study is doubtless ultimately an arbitrary one; and both have considerable insights to offer. But the two approaches, as currently practised, differ greatly in their procedures. Typologists typically study a wide range of languages as part of their enquiry, and tend to make generalizations that deal with the more observable aspects of structure, such as word order, word classes, and types of sound. In contrast with the empirical breadth of such studies, universalists rely on in-depth studies of single languages, especially in the field of grammar – English, in particular, is a common language of exemplification – and tend to make generalizations about the more abstract, underlying properties of language.

This focus on single languages might at first seem strange. If we are searching for universals, then surely we need to study many languages? Chomsky argues, however, that there is no paradox. Because English is a human language, it must therefore incorporate all universal properties of language, as well as those individual features that make it specifically 'English'. One way of finding out about these properties, therefore, is the detailed study of single languages. The more languages we introduce into our enquiry, the more difficult it can become to see the central features behind the welter of individual differences.

On the other hand, it can be argued that the detailed study of single languages is inevitably going to produce a distorted picture. There are features of English, for example, that are *not* commonly met with in other languages, such as the use of only one inflectional ending in the present tense (third-person, as in *she runs*), or the absence of a second-person singular / plural distinction (cf. French *tu / vous*). Without a typological perspective, some say, it is not possible to anticipate the extent to which our sense of priorities will be upset. If languages were relatively homogeneous entities, like samples of iron ore, this would not be a problem. But, typologists argue, languages are unpredictably irregular and idiosyncratic. Under these circumstances, a focus on breadth, rather than depth, is desirable.

RELATIVE OR ABSOLUTE?

The universalist ideal is to be able to make succinct and interesting statements that hold, without exception, for all languages. In practice, very few such statements can be made: the succinct ones often seem to state the obvious (e.g. all languages have vowels); and the interesting ones often seem to require considerable technical qualification. Most of the time, in fact, it is clear that 'absolute' (or exceptionless) universals do not exist. As a result, many linguists look instead for trends or tendencies across languages – 'relative' universals – which can be given statistical expression. For example, in over 99% of languages whose word order has been studied, grammatical subjects precede objects. And in a phonological study of over 300 languages (p. 167), less than 3% have no nasal consonant. Linguistic features that are statistically dominant in this way are often referred to as 'unmarked'; and a grammar that incorporates norms of this kind is known as a 'core grammar' (p. 97).

THREE TYPES OF UNIVERSALS

Substantive

Substantive universals comprise the set of categories that is needed in order to analyse a language, such as 'noun', 'question', 'first-person', 'antonym', and 'vowel'. Do all languages have nouns and vowels? The answer seems to be yes. But certain categories often thought of as universal turn out not to be so: not all languages have case endings, prepositions, or future tenses, for example, and there are several surprising limitations on the range of vowels and consonants that typically occur (§28). Analytical considerations must also be borne in mind. Do all languages have words? The answer depends on how the concept of 'word' is defined (p. 91).

Formal

Formal universals are a set of abstract conditions that govern the way in which a language analysis can be made – the factors that have to be written into a grammar, if it is to account successfully for the way sentences work in a language. For example, because all languages make statements and ask related questions (such as *The car is ready* vs *Is the car ready?*), some means has to be found to show the relationship between such pairs. Most grammars derive question structures from statement structures by some kind of transformation (in the above example, 'Move the verb to the beginning of the sentence'). If it is claimed that such transformations are necessary in order to carry out the analysis of these (and other kinds of) structures, as one version of Chomskyan theory does, then they would be proposed as formal universals. Other cases include the kinds of rules used in a grammar, or the different levels recognized by a theory (§13).

Implicational

Implicational universals always take the form 'If X, then Y', their intention being to find constant relationships between two or more properties of language. For example, three of the universals proposed in a list of 45 by the American linguist Joseph Greenberg (1915–) are as follows:

Universal 17. With overwhelmingly more-than-chance frequency, languages with dominant order VSO [=Verb–Subject–Object] have the adjective after the noun.

Universal 31. If either the subject or object noun agrees with the verb in gender, then the adjective always agrees with the noun in gender.

Universal 43. If a language has gender categories in the noun, it has gender categories in the pronoun.

As is suggested by the phrasing, implicational statements have a statistical basis, and for this reason are sometimes referred to as 'statistical' universals (though this is a somewhat different sense from that used in §15).

HOW MANY LANGUAGES?

It is impossible in principle to study all human languages, in order to find out about universals, for the simple reason that many languages are extinct, and there is no way of predicting what languages will emerge in the future. To be practical, typological or universal studies therefore need to be based on a sample of the 6,000 or so current languages of the world (§47). But how should a representative sample be achieved?

Several projects on language universals have had to address this basic question. The aim is to include as many different kinds of language as possible. Languages are selected from the main branches of every language family, insofar as these are known. They are not selected from the same local geographical area, in case they display a high degree of mutual influence. And the number of languages within each family has to be carefully considered. It would not be right to select an arbitrary five languages from each family – bearing in mind that Indo-Pacific, for example, has over 700 languages, whereas Dravidian has only about 25 (§52). The languages of New Guinea ought, statistically speaking, to constitute about 20% of any sample.

In practice, surveys have to be satisfied with what they can get. As few of the New Guinea languages have been studied in depth, for instance, it is currently impracticable to achieve the target of 20%. For such reasons, even the largest surveys work under considerable limitations. For example, in an American study of phonological universals (§28), the database was provided by a total of only 317 languages – about 5% of the whole. But the study nonetheless provided an enormous amount of valuable information (I. Maddieson, 1984).

BASIC GRAMMATICAL NOTIONS

The range of constructions that is studied by grammar is very large, and grammarians have often divided it into sub-fields. The oldest and most widely-used division is that between morphology and syntax.

MORPHOLOGY

This branch of grammar studies the structure of words. In the following list, all the words except the last can be divided into parts, each of which has some kind of independent meaning.

unhappiness	*un- -happi- -ness*
horses	*horse- -s*
talking	*talk- -ing*
yes	*yes*

Yes has no internal grammatical structure. We could analyse its constituent sounds, /j/, /e/, /s/, but none of these has a meaning in isolation. By contrast, *horse, talk,* and *happy* plainly have a meaning, as do the elements attached to them (the 'affixes'): *un-* carries a negative meaning; *-ness* expresses a state or quality; *-s* expresses plural; and *-ing* helps to convey a sense of duration. The smallest meaningful elements into which words can be analysed are known as *morphemes*; and the way morphemes operate in language provides the subject matter of *morphology.*

It is an easy matter to analyse the above words into morphemes, because a clear sequence of elements is involved. Even an unlikely word such as *antidisestablishmentarianism* would also be easy to analyse, for the same reason. In many languages (the so-called 'agglutinating' languages (p. 295)), it is quite normal to have long sequences of morphemes occur within a word, and these would be analysed in the same way. For example, in Eskimo the word *angyaghllangyugtuq* has the meaning 'he wants to acquire a big boat'. Speakers of English find such words very complex at first sight; but things become much clearer when we analyse them into their constituent morphemes:

angya-	'boat'
-ghlla-	an affix expressing augmentative meaning
-ng-	'acquire'
-yug-	an affix expressing desire
-tuq-	an affix expressing third person singular.

English has relatively few word structures of this type, but agglutinating and inflecting languages, such as Turkish and Latin, make widespread use of morphological variation. Many African languages, such as Swahili or Bilin, have verbs which can appear in well over 10,000 variant forms.

MORPHEME PROBLEMS

Not all words can be analysed into morphemes so easily. In English, for example, it is difficult to know how to analyse irregular nouns and verbs: *feet* is the plural of *foot*, but it is not obvious how to identify a plural morpheme in the word, analogous to the *-s* ending of *horses.* In the Turkish word *evinden* 'from his/her house', there is the opposite problem, as can be seen from the related forms:

ev	house
evi	his / her / its house
evden	from the house

It seems that the *-i* ending marks 'his / her / its', and the *-den* ending marks 'from' – in which case the combination of the two ought to produce *eviden.* But the form found in Turkish has an extra *n*, which does not seem to belong anywhere. Its use is automatic in this word (in much the same way as an extra *r* turns up in the plural of *child* in English – *child-r-en*). Effects of this kind complicate morphological analysis – and add to its fascination. Explanations can sometimes be found in other domains: it might be possible to explain the *n* in *evinden* on phonetic grounds (perhaps anticipating the following nasal sound), and the *r* in children is certainly a fossil of an older period of usage (Old English *childru*). To those with a linguistic bent, there is nothing more intriguing than the search for regularities in a mass of apparently irregular morphological data.

Another complication is that morphemes sometimes have several phonetic forms, depending on the context in which they occur. In English, for example, the past-tense morpheme (written as *-ed*), is pronounced in three different ways, depending on the nature of the sounds that precede it. If the preceding sound is /t/ or /d/, the ending is pronounced /id/, as in *spotted;* if the preceding sound is a voiceless consonant (p. 128), the ending is pronounced /t/, as in *walked*; and if the preceding sound is a voiced consonant or a vowel, the ending is pronounced /d/, as in *rolled.* Variant forms of a morpheme are known as *allomorphs.*

INFLECTIONAL AND DERIVATIONAL

Two main fields are traditionally recognized within morphology. *Inflectional morphology* studies the way in which words vary (or 'inflect') in order to express grammatical contrasts in sentences, such as singular/plural or past/present tense. In older grammar books, this branch of the subject was referred to as 'accidence'. *Boy* and *boys,* for example, are two forms of the 'same' word; the choice between them, singular vs plural, is a matter of grammar, and thus the business of inflectional morphology. *Derivational morphology,* however, studies the principles governing the construction of new words, without reference to the specific grammatical role a word might play in a sentence. In the formation of *drinkable* from *drink*, or *disinfect* from *infect*, for example, we see the formation of different words, with their own grammatical properties.

NEW WORDS OUT OF OLD

There are four normal processes of word formation in English:

- *prefixation* an affix is placed before the base of the word, e.g. *disobey*;
- *suffixation* an affix is placed after the base of the word, e.g. *kindness*;
- *conversion* a word changes its class without any change of form, e.g. *(the) carpet* (noun) becomes *(to) carpet* (verb);
- *compounding* two base forms are added together, e.g. *blackbird.*

There are also some less usual ways of making new words.

- *reduplication* a type of compound in which both elements are the same, or only slightly different, e.g. *goody-goody, wishy-washy, teeny-weeny*;
- *clippings* an informal shortening of a word, often to a single syllable, e.g. *ad, gents, flu, telly*;
- *acronyms* words formed from the initial letters of the words that make up a name, e.g. NATO, UNESCO, *radar* (= radio detection and ranging); a sub-type is an *alphabetism,* in which the different letters are pronounced, e.g. *VIP, DJ*;
- *blends* two words merge into each other, e.g. *brunch* (from 'breakfast' + 'lunch'), *telex* ('teleprinter' + 'exchange'.)

ABSO-BLOOMING-LUTELY

Morphemes can be classified into 'free' and 'bound' forms. Free morphemes can occur as separate words, e.g. *car, yes.* Bound morphemes cannot occur on their own, e.g. *anti-, -tion.* The main classes of bound morphemes are the prefixes and suffixes; but *infixes* are also possible – an affix which is inserted *within* a stem. The nearest we get to this in English is emphatic forms such as *abso-bloom-ing-lutely awful;* but in many languages, infixation is a normal morphological process. In Tagalog, for example, the form /um/ 'one who does' is infixed within the form /pi:lit/ 'effort' to produce /pumi:lit/, which means 'one who compelled'.

WORDS

Words sit uneasily at the boundary between morphology and syntax. In some languages – 'isolating' languages, such as Vietnamese (p. 295) – they are plainly low-level units, with little or no internal structure. In others – 'polysynthetic' languages, such as Eskimo – word-like units are highly complex forms, equivalent to whole sentences. The concept of 'word' thus ranges from such single sounds as English *a* to *palyamunurringkutjamunurtu* ('he/she definitely did not become bad') in the Western Desert language of Australia.

Words are usually the easiest units to identify, in the written language. In most writing systems, they are the entities that have spaces on either side. (A few systems use word dividers (e.g. Amharic), and some do not separate words at all (e.g. Sanskrit).) Because a literate society exposes its members to these units from early childhood, we all know where to put the spaces – apart from a small number of problems, mainly to do with hyphenation. Should we write *washing machine* or should it be *washing-machine? Well informed* or *well-informed? No one* or *no-one?*

It is more difficult to decide what words are in the stream of speech, especially in a language that has never been written down. But there are problems, even in languages like English or French. Certainly, it is possible to read a sentence aloud slowly, so that we can 'hear' the spaces between the words; but this is an artificial exercise. In natural speech, pauses do not occur between each word, as can be seen from any acoustic record of the way people talk. Even in very hesitant speech, pauses come at intervals – usually between major grammatical units, such as phrases or clauses (p. 95). So if there are no audible 'spaces', how do we know what the words are? Linguists have spent a great deal of time trying to devise satisfactory criteria – none of which is entirely successful.

There are no word spaces in the 4th century AD Greek *Codex Sinaiticus*. Word spaces were a creation of the Romans, and became widespread only in the Middle Ages.

FIVE TESTS OF WORD IDENTIFICATION

Potential pause
Say a sentence out loud, and ask someone to 'repeat it very slowly, with pauses'. The pause will tend to fall between words, and not within words. For example, *the / three / little / pigs / went / to / market.* But the criterion is not foolproof, for some people will break up words containing more than one syllable, e.g. *mar / ket.*

Indivisibility
Say a sentence out loud, and ask someone to 'add extra words' to it. The extra items will be added between the words and not within them. For example, *the pig went to market* might become *the big pig once went straight to the market,* but we would not have such forms as *pi-big-g* or *mar-the-ket.* However this criterion is not perfect either, in the light of such forms as *absobloominglutely.*

Minimal free forms
The American linguist Leonard Bloomfield (1887–1949) thought of words as 'minimal free forms' – that is, the smallest units of speech that can *meaningfully* stand on their own. This definition does handle the majority of words, but it cannot cope with several items which are treated as words in writing, but which never stand on their own in natural speech, such as English *the* and *of*, or French *je* ('I') and *de* ('of').

Phonetic boundaries
It is sometimes possible to tell from the sound of a word where it begins or ends. In Welsh, for example, long words generally have their stress on the penultimate syllable, e.g. (*cartref* 'home', *car'trefi* 'homes'. In Turkish, the vowels within a word harmonize in quality (p. 163), so that if there is a marked change in vowel quality in the stream of speech, a new word must have begun. But there are many exceptions to such rules.

Semantic units
In the sentence *Dog bites vicar,* there are plainly three units of meaning, and each unit corresponds to a word. But language is often not as neat as this. In *I switched on the light, the* has little clear 'meaning', and the single action of 'switching on' involves two words.

WORD CLASSES

Since the early days of grammatical study, words have been grouped into *word classes*, traditionally labelled the 'parts of speech'. In most grammars, eight such classes were recognized, illustrated here from English:

nouns	*boy, machine, beauty*
pronouns	*she, it, who*
adjectives	*happy, three, both*
verbs	*go, frighten, be*
prepositions	*in, under, with*
conjunctions	*and, because, if*
adverbs	*happily, soon, often*
interjections	*gosh, alas, coo*

In some classifications, participles (*looking, taken*) and articles (*a, the*) were separately listed.

Modern approaches classify words too, but the use of the label 'word class' rather than 'part of speech' represents a change in emphasis. Modern linguists are reluctant to use the notional definitions found in traditional grammar – such as a noun being the 'name of something'. The vagueness of these definitions has often been criticized: is *beauty* a 'thing'? is not the adjective *red* also a 'name' of a colour? In place of definitions based on meaning, there is now a focus on the structural features that signal the way in which groups of words behave in a language. In English, for example, the definite or indefinite article is one criterion that can be used to signal the presence of a following noun (*the car*); similarly, in Romanian, the article (*ul*) signals the presence of a preceding noun (*avionul* 'the plane').

Above all, the modern aim is to establish word classes that are coherent: all the words within a class should behave in the same way. For instance, *jump, walk,* and *cook* form a coherent class, because all the grammatical operations that apply to one of these words apply to the others also: they all take a third person singular form in the present tense (*he jumps / walks / cooks),* they all have a past tense ending in *-ed* (*jumped / walked / cooked),* and so on. Many other words display the same (or closely similar) behaviour, and this would lead us to establish the important class of 'verbs' in English. Similar reasoning would lead to an analogous class being set up in other languages, and ultimately to the hypothesis that this class is required for the analysis of all languages (as a 'substantive universal', §14).

CLASSIFYING NOUNS

Distinctions such as masculine / feminine and human / non-human are well known in setting up sub-classes of nouns, because of their widespread use in European languages. But many Indo-Pacific and African languages far exceed these in the number of noun classes they recognize. In Bantu languages, for example, we find such noun classes as human beings, growing things, body parts, liquids, inanimate objects, animals, abstract ideas, artefacts, and narrow objects.

However, these labels should be viewed with caution, as they are no more exact semantically than are the gender classes of European languages. In Swahili, for example, there are subclasses for human beings and insect / animal names, but the generic words 'insect' and 'animal' in fact formally belong to the 'human' class!

Gradience

Word classes should be coherent. But if we do not want to set up hundreds of classes, we have to let some irregular forms into each one. For example, for many speakers *house* is the only English noun ending in /s/, where the /s/ becomes /z/ when the plural ending is added (*houses*). Although in theory it is 'in a class of its own', in practice it is grouped with other nouns, with which it has a great deal in common.

Because of the irregularities in a language, word classes are thus not as neatly homogeneous as the theory implies. Each class has a core of words that behave identically, from a grammatical point of view. But at the 'edges' of a class are the more irregular words, some of which may behave like words from other classes. Some adjectives have a function similar to nouns (e.g. *the rich*); some nouns behave similarly to adjectives (e.g. *railway* is used adjectivally before *station*).

The movement from a central core of stable grammatical behaviour to a more irregular periphery has been called *gradience*. Adjectives display this phenomenon very clearly. Five main criteria are usually used to identify the central class of English adjectives:

(A) they occur after forms of *to be*, e.g. *he's sad*;
(B) they occur after articles and before nouns, e.g. *the big car*;
(C) they occur after *very*, e.g. *very nice*;
(D) they occur in the comparative or superlative form e.g. *sadder / saddest, more / most impressive*; and
(E) they occur before *-ly* to form adverbs, e.g. *quickly*.

We can now use these criteria to test how much like an adjective a word is. In the matrix below, candidate words are listed on the left, and the five criteria are along the top. If a word meets a criterion, it is given a +; *sad*, for example, is clearly an adjective (*he's sad, the sad girl, very sad, sadder / saddest, sadly*). If a word fails the criterion, it is given a – (as in the case of *want*, which is nothing like an adjective: **he's want, *the want girl, *very want, *wanter / wantest, *wantly*).

	A	B	C	D	E
happy	+	+	+	+	+
old	+	+	+	+	–
top	+	+	+	–	–
two	+	+	–	–	–
asleep	+	–	–	–	–
want	–	–	–	–	–

The pattern in the diagram is of course wholly artificial because it depends on the way in which the criteria are placed in sequence; but it does help to show the gradual nature of the changes as one moves away from the central class, represented by *happy*. Some adjectives, it seems, are more adjective-like than others.

WHAT PART OF SPEECH IS *ROUND*?

You cannot tell what class a word belongs to simply by looking at it. Everything depends on how the word 'behaves' in a sentence. *Round* is a good illustration of this principle in action, for it can belong to any of five word classes, depending on the grammatical context.

Adjective
Mary bought a round table.

Preposition
The car went round the corner.

Verb
The yacht will round the buoy soon.

Adverb
We walked round to the shop

Noun
It's your round. I'll have a whiskey.

A DUSTBIN CLASS?

Several of the traditional parts of speech lacked the coherence required of a well-defined word class – notably, the adverb. Some have likened this class to a dustbin, into which grammarians would place any word whose grammatical status was unclear. Certainly, the following words have very little structurally in common, yet all have been labelled 'adverb' in traditional grammars:

tomorrow	*very*	*no*
however	*quickly*	*when*
not	*just*	*the*

The, an adverb? In such contexts as *The more the merrier.*

NOUN TENSES?

Some languages formally mark the expression of time relations on word classes other than the verb. In Japanese, adjectives can be marked in this way, e.g. *shiroi* 'white', *shirokatta* 'was white', *shirokute* 'being white', etc. In Potowatomi, the same ending that expresses past time on verbs can be used on nouns

/nkašatəs/	I am happy
/nkəšatsəpən/	I was once happy
/nosˑ/	my father
/nospˑən/	my dead father
/nčiman/	my canoe
/nčimanpən/	my former canoe (lost, stolen)

(After C. F. Hockett, 1958, p. 238.)

FIVE MOODS

A range of attitudes can be expressed by the mood system of the verb. In Fox, one mood expresses the meaning 'God forbid that this should happen!'; another, 'What if it did happen! What do I care!' In Menomini, there is a five-term mood system:

/piˑw/	he comes / is coming / came
/piˑwen/	he is said to be coming / it is said that he came
/piˑʔ/	is he coming / did he come?
/piasah/	so he *is* coming after all!
/piapah/	but he was going to come! (and now it turns out he is not)

(After C. F. Hockett, 1958, p. 237.)

DUAL AND TRIAL NUMBER

Four numbers are found in the language spoken on Aneityum Island (Melanesia): singular, dual, trial, plural. The forms are shown for 1st and 2nd person: /ñ/ is a palatal nasal; /j/ is a palatal affricate or stop; excl./incl. = exclusive/inclusive of speaker:

/añak/	I
/akaja/	we (incl.)
/ajama/	we (excl.)
/akajau/	we two (incl.)
/ajamrau/	we two (excl.)
/akataj/	we three (incl)
/ajamtaj/	we three (excl.)
/aek,aak/	you
/ajourau/	you two
/ajoutaj/	you three
/ajowa/	you (pl.)

A FOURTH PERSON

A fourth-person contrast is made in the Algonquian languages, referring to non-identical animate third persons in a particular context. In Cree, if we speak of a man, and then (secondarily) of another man, the forms are different: /ˈna:peːw/ vs /ˈna:pewa/. This fourth person form is usually referred to as the 'obviative'.
(After L. Bloomfield, 1933, p. 257.)

FIFTEEN CASES

Nominative (subject), *genitive* (of), *accusative* (object), *inessive* (in), *elative* (out of), *illative* (into), *adessive* (on), *ablative* (from), *allative* (to), *essive* (as), *partitive* (part of), *translative* (change to), *abessive* (without), *instructive* (by), and *comitative* (with).

The Finnish case system seems fearsome to those brought up on the six-term system of Latin. But the less familiar cases are really quite like prepositions – except that the forms are attached to the end of the noun as suffixes, instead of being separate words placed before, as in English.

GRAMMATICAL CATEGORIES

In many languages, the forms of a word vary, in order to express such contrasts as number, gender, and tense. These categories are among the most familiar of all grammatical concepts, but their analysis can lead to surprises. In particular, it emerges that there is no neat one-to-one correspondence between the grammatical alterations in a word's form and the meanings thereby conveyed. Plural nouns do not always refer to 'more than one'; a first-person pronoun does not always refer to the person who is talking; and masculine nouns are not always male.

Category	Typical formal contrasts	Typical meanings conveyed	Examples	But note...
aspect (verbs)	perfect(ive), imperfect(ive)	completeness, habituality, continuousness, duration, progressiveness	Russian *ya pročital* (pf.) vs *ya čital* (impf.), roughly 'I read' vs 'I used to read / was reading'; English *she sings* (as a job) vs *she's singing* (now).	Adverbs can change the meaning, as when *always* changes the 'in progress' meaning of *John is driving from London* to a habitual (and often irritated) meaning: *John's always driving from London.*
case (nouns, pronouns, adjectives)	nominative, vocative, accusative, genitive, partitive	actor, possession, naming, location, motion towards	English gen. *boy's, girls';* Latin nom. *puella* 'girl', gen. *puellae* 'of the girl'; Serbo-Croat *grad* 'town', loc. *gradu* 'at a town'.	Cases may have several functions. The English genitive is sometimes called the 'possessive', but it can express other meanings than possession, e.g. *the man's release, a week's leave, a summer's day.*
gender (nouns, verbs, adjectives)	masculine, feminine, neuter, animate, inanimate	male, female, sexless, living	Spanish masc. *el muchacho* 'boy', fem. *la muchacha* 'girl'; German masc. *der Mann* 'the man', fem. *die Dame* 'the lady', neut. *das Ende* 'the end'; Russian past tense singular masc. *čital,* fem. *čitala,* neut. *čitalo* 'read'.	There is no necessary correlation between grammatical gender and sex. In German, 'spoon' is masculine (*der Löffel*); 'fork' is feminine (*die Gabel*); 'knife' is neuter (*das Messer*). French 'love' *amour* is masculine in the singular, but often feminine in the plural.
mood (verbs)	indicative, subjunctive, optative	factuality, possibility, uncertainty, likelihood	Latin *requiescit* 'he / she / it rests' vs *requiescat* 'may he / she rest'; English *God save the Queen, if I were you.*	Although a major section in traditional grammars, many European languages no longer make much use of the subjunctive. It is often restricted to formulaic phrases or very formal situations.
number (nouns, verbs, pronouns)	singular, dual, trial, plural	one, two, more than one, more than two, more than three	Swedish *bil* 'car', *bilar* 'cars'; Dutch *ik roep* 'I call', *wij roepen* 'we call'; Samoan /ʔoe/ 'you' (sing.), /ʔoulua/ 'you two', /ʔoutou/ 'you' (pl.).	Nouns plural in form may refer to singular entities (e.g. *binoculars, pants*), and some nouns functioning as singulars refer to several events (e.g. *athletics, news*). The two crops known as *wheat* and *oats* look very similar; but in English one is singular and the other is plural.
person (pronouns, verbs)	first person, second person, third person, fourth person	speaker, addressee, third party, fourth party	Welsh *mi* 'I', *ni* 'we'; Menomini /nenah/ 'I' /kenah/ 'thou', /wenah/ 'he'; Latin *amo* 'I love', *amas* 'you love' (sing.), *amat* 'he / she / it loves'.	First person can refer to addressee (Doctor (to patient): *How are we today?*) or to a third party (Secretary – to friend, about the boss: *We're not in a good mood today*). Third person can refer to self (Wife: *How's my husband?* Husband: *He's hungry*).
tense (verbs)	present, past, future	present time, past time, future time	Italian *io parlo* 'I speak', *io ho parlato* 'I have spoken', *io parlavo* 'I was speaking'; Gaelic *chuala mi* 'I heard', *cluinneadh mi* 'I'll hear'.	Tense and time do not always correspond. Present tense–past time: *Minister dies* (headline). Present tense–future time: *I'm leaving tomorrow.*
voice (verbs)	active, passive middle, causative	who did action what was acted upon, what caused action	Classical Greek active *didàsko* 'I teach', middle *dídàskomai* 'I get myself taught'; Portuguese active *cortou* 'cut', passive *foi cortada* 'was cut'; Tigrinya active *qätäle* 'he killed', causative *ʔaqtäle* 'he caused to kill'.	There are several active verbs in English which have no passive (e.g. *She has a car* will not transform into **A car is had by her*), and several passives which have no active (e.g. *He was said to be angry* will not transform into **Someone said him to be angry*).

SYNTAX

Syntax is the way in which words are arranged to show relationships of meaning within (and sometimes between) sentences. The term comes from *syntaxis,* the Greek word for 'arrangement'. Most syntactic studies have focused on sentence structure, for this is where the most important grammatical relationships are expressed.

THE SENTENCE

Traditionally, grammars define a sentence in such terms as 'the complete expression of a single thought'. Modern studies avoid this emphasis, because of the difficulties involved in saying what 'thoughts' are. *An egg* can express a thought, but it would not be considered a complete sentence. *I shut the door, as it was cold* is one sentence, but it could easily be analysed as two thoughts.

Some traditional grammars give a logical definition to the sentence. The most common approach proposes that a sentence has a 'subject' (= the topic) and a 'predicate' (= what is being said about the topic). This approach works quite well for some sentences, such as *The book is on the table,* where we can argue that *the book* is what the sentence is 'about'. But in many sentences it is not so easy to make this distinction. *It's raining* is a sentence, but what is the topic? And in *Michael asked Mary for a pen,* it is difficult to decide which of Michael, Mary, or the pen is the topic – or whether we have *three* topics! Also, some modern grammars treat subjects and topics in completely different ways.

In some written languages, it is possible to arrive at a working definition of 'sentence' by referring to the punctuation one is taught to use in school. Thus, an English sentence for many people 'begins with a capital letter and ends with a full stop' (or some other mark of 'final' punctuation). The problem is that many languages (e.g. in Asia) do not make use of such features; and even in those that do, punctuation is not always a clear guide. It may be omitted (in notices and legal documents, for example); and it proves difficult to prescribe rules governing its use other than 'good practice'. People therefore often disagree about the best way to punctuate a text. In some manuals of style, it is recommended that one should not end a sentence before a coordinating conjunction (*and, or, but*). But there are often cases where an author might feel it necessary – for reasons of emphasis, perhaps – to do the opposite.

It is even more difficult to identify sentences in speech, where the units of rhythm and intonation often do not coincide with the places where full stops would occur in writing. In informal speech, in particular, constructions can lack the careful organization we associate with the written language (p. 52). It is not that conversation lacks grammar: it is simply that the grammar is of a rather different kind, with sentences being particularly difficult to demarcate. In the following extract, it is not easy to decide whether a sentence ends at the points marked by pauses (–), or whether this is all one, loosely constructed sentence:

when they fed the pigs/ they all had to stand well back/ – and they were allowed to take the buckets/ – but they weren't allowed to get near the pigs/ you see/ – so they weren't happy ...

Linguistic approaches

Despite all the difficulties, we continue to employ the notion of 'sentence', and modern syntacticians try to make sense of it. But they do not search for a satisfactory definition of 'sentence' at the outset – an enterprise that is unlikely to succeed, with over 200 such definitions on record to date. Rather, they aim to analyse the linguistic constructions that occur, recognizing the most independent of them as sentences. Thus, because the following constructions can stand on their own as utterances, and be assigned a syntactic structure, they would be recognized as sentences:

she asked for a book/
come in/
the horse ran away because the train was noisy/

The following combination of units, however, could not be called a sentence:

will the car be here at 3 o'clock/ it's raining/.

The syntax of the first unit and that of the second do not combine to produce a regular pattern. It would be just as possible to have:

it's raining / will the car be here at 3 o'clock /

or either unit without the other. Within each unit, however, several kinds of rules of syntactic order and selection are apparent. We may not say:

*will be here at 3 o'clock the car/
*will be here the car/
*car at 3 o'clock/.

Each unit in the sequence, then, is a sentence; but the combination does not produce a 'larger' syntactic unit.

A sentence is thus the largest unit to which syntactic rules apply – 'an independent linguistic form, not included by virtue of any grammatical construction in any larger linguistic form' (L. Bloomfield, 1933, p. 170). But this approach has its exceptions, too. In particular, we have to allow for cases where sentences are permitted to omit part of their structure and thus be dependent on a previous sentence (*elliptical* sentences), as in:

A: Where are you going?
B: To town.

Several other types of exception would be recognized in a complete grammatical description.

MINOR SENTENCE TYPES

A language contains many sentence-like units which do not conform to the regular patters of formation. Here is a selection from English:

Yes
Gosh!
Least said, soonest mended.
How come you're early?
Oh to be free!
All aboard!
Down with racism!
No entry.
Taxi!
Good evening.
Happy birthday!
Checkmate.

A sign like this has a regular syntactic structure, but it does not use normal sentence punctuation.

ASPECTS OF SENTENCE SYNTAX

Hierarchy

Hilary couldn't open the windows.

One of the first things to do in analysing a sentence is to look for groupings within it – sets of words (or morphemes, p. 90) that hang together. In this example, we might make an initial division as follows:

Hilary / couldn't open / the windows.

Units such as *couldn't open* and *the windows* are called *phrases.* The first of these would be called a *verb phrase,* because its central word (or 'head') is a verb, *open;* the second would be called a *noun phrase,* because its head is a noun, *windows.* Other types of phrase also exist – adjective phrases, for example, such as *very nice.*

Phrases may in turn be divided into their constituent *words* (p. 91):

couldn't + open *the + windows*

And words may be divided into their constituent *morphemes,* if there are any:

could + n't *window + s*

This conception of sentence structure as a hierarchy of levels, or ranks, may be extended 'upwards'. The sentence can be made larger by linking several units of the same type:

Hilary opened the windows, but David couldn't open the doors.

Here, too, we have a sentence, but now we have to recognize two major units within it – each of which has a structure closely resembling that of an independent sentence. These units are traditionally referred to as *clauses.* In the above example, the clauses have been 'coordinated' through the use of the conjunction *but.* An indefinite number of clauses can be linked within the same sentence.

A five-rank hierarchy is a widely used model of syntactic investigation:

sentences are analysed into	↓	sentences are used to build
clauses are analysed into	↓	clauses are used to build
phrases are analysed into	↓	phrases are used to build
words are analysed into	↓	words are used to build
morphemes	↓	morphemes

Morphemes are the 'lower' limit of grammatical enquiry, for they have no grammatical structure. Similarly, sentences form the 'upper' limit of grammatical study, because they do not usually form a part of any larger grammatical unit.

CLAUSES

The various units that make up the structure of a clause are usually given functional labels, such as *Subject* (S), *Verb* (V), *Complement* (C), *Object* (O), and *Adverbial* (A). A number of clause types can be identified in this way, such as:

S + V	The dog + is running.
S + V + O	The man + saw + a cow.
S + V + C	The car + is + ready.
S + V + A	A picture + lay + on the ground.
S + V + O + O	I + gave + John + a book.
S + V + O + C	He + called + John + a fool.
S + V + O + A	Mary + saw + John + yesterday.

Several approaches to grammatical analysis make use of elements of this kind, though there is considerable variation in definition and terminology. Languages also vary greatly in the way in which these elements are identified. In English, for example, word order is the main factor, with only occasional use being made of morphology (e.g. *he* (subject) *saw* (verb) *him* (object)). In Latin, word-endings provide the main clues to element function, word order being irrelevant (e.g. *puer puellam vidit* 'the boy saw the girl'). In Japanese, basic grammatical relations are marked by special particles: *ga* (subject), *o* (direct object), *ni* (indirect object), and *no* (genitive). For example,

kodomo ga	*tomodachi no*	*inu ni*	*mizu o*	*yaru*
the child	friend's	to dog	water	gives

'The child gives water to his / her friend's dog.'

PHRASES

Most phrases can be seen as expansions of a central element (the *head),* and these are often referred to as 'endocentric' phrases:

cars
the cars
the big cars
all the big cars
all the big cars in the garage

Phrases which cannot be analysed in this way are then called 'exocentric': *inside / the cars.*

The internal structure of an endocentric phrase is commonly described in a three-part manner:

all the big	*cars*	*in the garage*
PREMODIFICATION	HEAD	POSTMODIFICATION

COORDINATION VS SUBORDINATION

Coordination is one of two main ways of making sentences more complex; the other is known as *subordination,* or 'embedding'. The essential difference is that in the former the clauses that are linked are of equal grammatical status, whereas in the latter, one clause functions as part of another (the 'main' clause). Compare:

Coordinate clause:
The boy left on Monday *and the girl left on Tuesday.*

Subordinate clause:
The boy left on Monday *when John rang.*

The phrase *on Monday* is part of the clause, giving the time when the action took place. Similarly, the unit *when John rang* is also part of the clause, for the same reason. But *when John rang* is additionally a clause in its own right.

CONCORD

Grammatical links between words are often signalled by concord or 'agreement'. A form of one word requires a corresponding form of another, as when in English a singular noun 'agrees with' a singular verb in the present tense: *the man walks* vs *the men walk.*

The purpose of concord varies greatly between languages. In Latin, it is an essential means of signalling which words go together. In the absence of fixed word-order patterns, sentences would otherwise be uninterpretable. For example, in *parvum puerum magna puella vidit* 'the tall girl saw the small boy', we know that the boy is small and the girl is tall only through the agreement of the endings, *-um* vs *-a.*

On the other hand, concord plays much less of a role in modern French, in cases such as *le petit garçon et la grande fille* 'the little boy and the big girl'. Because the position of adjectives is fixed (before the noun, in these cases), it would not pose any problems of intelligibility if there were no difference between the masculine and feminine forms:

le petit garçon
**la petit fille*
**le petite garçon*
la petite fille

If French allowed free word order, as in Latin, so that one could say **le garçon et la fille petit grande,* then concord would be needed to show which adjective should go with which noun – but this does not happen. The gender system is thus of limited usefulness, though it still has a role to play in certain syntactic contexts, such as cross-reference (*J'ai vu* ***un*** *livre et une plume.* ***Il*** *était* ***nouveau.*** 'I saw a book and a pen. It [i.e. the book] was new.').

IMMEDIATE CONSTITUENT DIAGRAMS

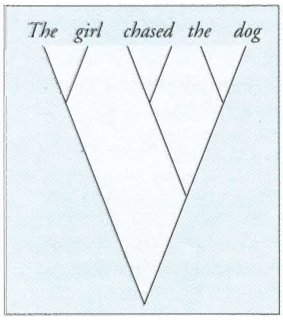
One of the most widely used techniques for displaying sentence structure is the use of *immediate constituent* (IC) analysis. This approach works through the different levels of structure within a sentence in a series of steps. At each level, a construction is divided into its major constituents, and the process continues until no further divisions can be made. For example, to make an IC analysis of the sentence *The girl chased the dog,* we carry out the following steps:

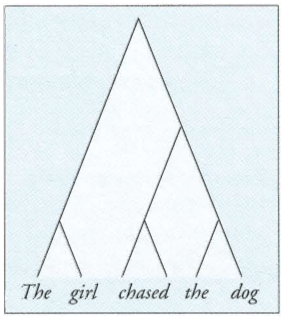
1. Identify the two major constituents, *the girl* and *chased the dog.*
2. Divide the next-biggest constituent into two, *viz. chased the dog* into *chased* and *the dog.*
3. Continue dividing constituents into two until we can go no further, *viz. the girl* and *the dog* into *the + girl, the + dog,* and *chased* into *chase* + the *-ed* ending.

The order of segmentation can be summarized using lines or brackets. If the first cut is symbolized by a single vertical line, the second cut by two lines, and so on, the sentence would look like this:

the /// girl / chase /// -ed // the /// dog

However, a much clearer way of representing constituent structure is through the use of 'tree diagrams':



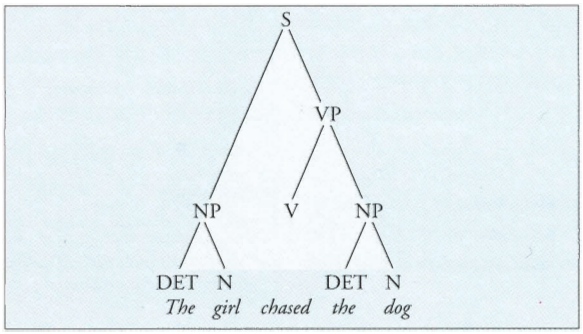


The second kind of tree diagram is in fact the normal convention in modern linguistics.

Such representations of structure are very helpful, as far as they go. But not all sentences are as easy to analyse in IC terms as this one. It is sometimes not clear where the cuts should be made (e.g. whether to divide *the three old men* into *the + three old men* or *the three old + men,* or *the three + old men*). More important, the process of segmenting individual sentences does not take us very far in understanding the grammar of a language. IC analyses do not inform us about the identity of the sentence elements they disclose, nor do they provide a means of showing how sentences relate to each other grammatically (as with statements and questions, actives and passives). To develop a deeper understanding of grammatical structure, alternative approaches must be used.

PHRASE STRUCTURE

A good way of putting more information into an analysis would be to name, or *label,* the elements that emerge each time a sentence is segmented. It would be possible to use functional labels such as 'subject' and 'predicate', but the approach that is most widely practised has developed its own terminology and abbreviations, so these will be used here. Taking the above sentence (S), the first division produces a 'noun phrase' (NP) *the girl* and a 'verb phrase' (VP) *chased the dog.* (This is a broader sense of 'verb phrase' than that used on p. 95, as it includes both the verb and the noun phrase that follows.) The second division recognizes a 'verb' (V) *chased* and another noun phrase *the dog.* The next divisions would produce combinations of 'determiner' (DET) and 'noun' (N) *the + girl, the + dog.* This is the 'phrase structure' of the sentence, and it can be displayed as a tree diagram:



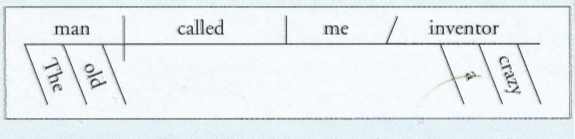
This kind of representation of the phrase structure of a sentence is known as a 'phrase marker' (or 'P-marker'). Phrase structures are also sometimes represented as labelled sets of brackets, but these are more difficult to read:

$[_S[_{NP}[_{DET}\text{the}][_N\text{girl}]][_{VP}[_V\text{chased}][_{NP}[_{DET}\text{the}][_N\text{dog}]]]]$

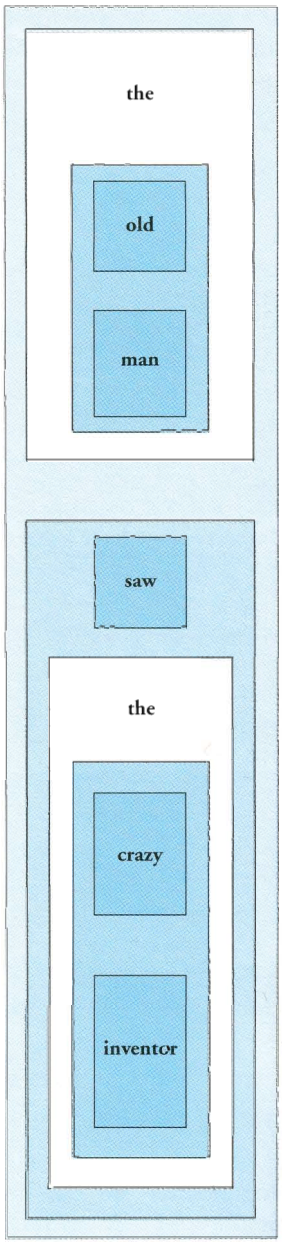
DIAGRAMMING

A frequent practice in American schools is the use of a system of vertical and slanting lines to represent the various relationships in a sentence. The representations are often called 'Reed & Kellogg' diagrams, after the authors of a 19th-century English textbook. A long vertical line marks the boundary between subject and predicate; a short vertical line divides verb and direct object; and a short slanting line marks off a complement. Other items are drawn in beneath the main parts of the sentence.

The old man called me a crazy inventor.



The approach shows the relationships between words clearly, but it cannot handle variations in word order: both *I turned off the light* and *I turned the light off* would be diagrammed in the same way.



A little-used 'Chinese box' representation of sentence structure

RULES

Analyses of single sentences are illuminating, as far as they go, but grammarians are concerned to move beyond this point, to see whether their analyses work for other sentences in the language. To what other sentences might the above sequence of steps, and the resulting P-marker, also apply? In Noam Chomsky's approach, first outlined in *Syntactic Structures* (1957), the jump from single-sentence analysis is made by devising a set of rules that would 'generate' tree structures such as the above. The procedure can be illustrated using the following rules (but several details from the original approach are omitted for clarity):

S → NP + VP
VP → V + NP
NP → DET + N
V → *chased*
DET → *the*
N → *girl, dog*

The first rule states that a sentence can consist of a noun phrase and a following verb phrase; the second, that a verb phrase can consist of a verb plus a following noun phrase; the third, that a noun phrase can consist of a determiner plus a noun. Each abstract category is then related to the appropriate words, thus enabling the sentence to be generated. Grammars that generate phrase structures in this way have come to be called 'phrase structure grammars' (PSGs).

If we follow these rules through, it can be seen that there is already a significant increase in the 'power' of this grammar over the single-sentence analysis used previously. If we choose *the girl* for the first NP, and *the dog* for the second, we generate *the girl chased the dog*; but if the choices are made the other way round, we generate the sentence *the dog chased the girl.* By the simple device of adding a few more words to the rules, suddenly a vast number of sentences can be generated:

V → *chased, saw, liked ...*
DET → *the, a*
N → *girl, man, horse ...*

the girl chased the horse
the man saw the girl
the horse saw the man etc.

However, if *went* were introduced into the rules, as a possible V, ungrammatical sentences would come to be generated, such as **the girl went the man.* In working out a generative grammar, therefore, a means has to be found to block the generation of this type of sentence, at the same time permitting such sentences as *the man went* to be generated. The history of generative syntax since 1957 is the study of the most efficient ways of writing rules, so as to ensure that a grammar will generate all the grammatical sentences of a language and none of the ungrammatical ones.

Transformations

This tiny fragment of a generative grammar from the 1950s suffices only to illustrate the general conception underlying the approach. 'Real' grammars of this kind contain many rules of considerable complexity and of different types. One special type of rule that was proposed in the first formulations became known as a *transformational* rule. These rules enabled the grammar to show the relationship between sentences that had the same meaning but were of different grammatical form. The link between active and passive sentences, for example, could be shown – such as *the horse chased the man* (active) and *the man was chased by the horse* (passive). The kind of formulation needed to show this is:

$$NP_1 + V + NP_2 \rightarrow NP_2 + Aux + V\textit{en} + by + NP_1$$

which is an economical way of summarizing all the changes you would have to introduce, in order to turn the first sentence into the second. If this formula were to be translated into English, four separate operations would be recognized:

(i) The first noun phrase in the active sentence (NP_1) is placed at the end of the passive sentence.
(ii) The second noun phrase in the active sentence (NP_2) is placed at the beginning of the passive sentence.
(iii) The verb (V) is changed from past tense to past participle (V*en*), and an auxiliary verb (Aux) is inserted before it.
(iv) A particle *by* is inserted between the verb and the final noun phrase.

This rule will generate all regular active-passive sentences.

In subsequent development of generative grammar, many kinds of transformational rules came to be used, and the status of such rules in a grammar has proved to be controversial (§65). Recent generative grammars look very different from the model proposed in *Syntactic Structures.* But the fundamental conception of sentence organization as a single process of syntactic derivation remains influential, and it distinguishes this approach from those accounts of syntax that represent grammatical relations using a hierarchy of separate ranks (p. 95).

RULES AND 'RULES'

The 'rules' of a generative grammar are not to be identified with the prescriptive 'rules' that formed part of traditional grammar (p. 3). A prescriptive grammatical rule is a statement – such as 'You should never end a sentence with a preposition' – that tells us whether we are right or wrong to use a particular construction. Generative rules have no such implication of social correctness. They are objective descriptions of the grammatical patterns that occur.

GENERATIVE NOTATION

A major feature of generative grammar is the way special notations have been devised to enable rules to be expressed in an economical way. In particular, different types of brackets, such as ( ), [ ], and { } are given different meanings. Round brackets, for example, enclose a grammatical element that is *optional* in a sentence; that is, the sentence would be grammatical even if the element were left out. The rule

NP → DET (ADJ) N

means that a noun phrase can consist of *either* a determiner, adjective, and noun *or* simply a determiner and noun (*the old man* or *the man*). A grammar could, of course, list the two possibilities separately, as

NP → DET + N

NP → DET + ADJ + N

but collapsing them into a single rule, through the use of the ( ) convention, saves a great deal of space, and represents something we all 'know' about the structure of the noun phrase.

PRINCIPLES AND PARAMETERS

Government and binding theory is an approach to generative grammar which developed in the 1980s. It takes its name from the way it focuses on the conditions which formally relate (or 'bind') certain elements of a sentence, and on the structural contexts within which these binding relationships apply ('govern').

The approach holds that the same *principles* of syntax operate in all languages, though they can differ slightly (along certain *parameters*) between languages (§14). For example it is a syntactic principle that in a noun phrase there is a chief element (the *head*), which will be the noun (*the new* ***President***), and that other nouns may accompany it (*the President of America*). But whether the accompanying nouns occur before or after the head varies between languages: they occur after it in English, but before it in Japanese (*Amerika no Daitoryo*).

WORD ORDER

The term 'word order' is somewhat ambiguous, for it can refer both to the order of words in a phrase, and to the order of multi-word units within a sentence. Given the sentence

The cat sat on the mat

both the following involve word-order problems – but they are of very different kinds:

**cat the*	*sat*	*mat the on*
**sat*	*the cat*	*on the mat*

In linguistic description, word-order studies usually refer to the second type of problem – that is, the sequence in which grammatical elements such as Subject, Verb, and Object occur in sentences. A great deal of attention has been paid to the way in which languages vary the order of these elements, as part of typological studies (§14). Word order, it is hoped, will be a more satisfactory way of classifying languages than the older morphological method (which recognized such types as isolating and inflecting, p. 295), into which many languages do not fit neatly.

In comparing word orders across languages, it is important to appreciate that what is being compared is the 'basic' or 'favourite' pattern found in each language. For example, in English, we will encounter such sequences as:

SVO *the boy saw the man*
OVS *Jones I invited – not Smith*
VSO *govern thou my song* (Milton)
OSV *strange fits of passion have I known* (Wordsworth)
SOV *pensive poets painful vigils keep* (Pope)

However, only the first of these is the natural, usual, 'unmarked' order in English; the others all convey special effects of an emphatic or poetic kind. The same principle must apply in studying word order in all languages, but it is often not so easy to establish which is the normal word-order pattern and which is the pattern that conveys the special effect. The mere fact of talking to a foreigner, for instance, might motivate a native speaker to change from one order to another, and it often requires great ingenuity on the part of the linguist to determine whether such stylistic changes are taking place.

Typology

Apart from cases of free word order (e.g. Latin, Quechua, Navajo, Fore), there are six logical possibilities: SVO, SOV, VSO, VOS, OSV, OVS. Of these, over 75% of the world's languages use SVO (as in English, French, Hausa, Vietnamese) or SOV (as in Japanese, Amharic, Tibetan, Korean). A further 10–15% use VSO (e.g. Welsh, Tongan, Squamish). Examples of VOS are Malagasy, Tzotzil, and Houailou.

Until recently, Object-initial languages were conspicuous by their absence, and it was thought that perhaps these did not exist. But a group of OVS languages have now been found, all in the Amazon basin, mainly belonging to the Carib family, e.g. Hixkaryana, Apalai, Bacairi, Makusi. A few other languages (e.g. Jamamadi, Apurina) seem to be OSV. But there is some variability in the data that have been collected so far, with both OVS and OSV being used by some languages.

Word-order generalizations often need careful qualification. Latin, for example, is said to have a free word order, but in fact SOV is a very common pattern in that language. Modern Hebrew is SVO, but Classical Hebrew seemed to favour VSO. German prefers SVO in main clauses, but SOV in subordinate clauses. In Tagalog, the V usually comes first, but there is great variation in what follows, with both OS and SO being widely used. In Japanese, SOV is favoured, but OSV is also very common.

LISU

This Lolo-Burmese language seems to have free word order, yet it has no morphological cases to mark Subject and Object. A sentence Noun–Verb–Noun might therefore mean either 'N1 did V to N2' or 'N2 did V to N1'. In theory, such a language ought to be unintelligible! But in fact the speakers survive, by relying on context, the use of alternative grammatical constructions, and a modicum of common sense.

OSV IN SPACE

Sick have I become.
Strong am I with the Force.
Your father he is.
When nine hundred years old you reach, look as good you will not.

The rarity of OSV constructions and languages perhaps explains the impact of this strange speech style, used by the Jedi Master, Yoda, in the film *The Empire Strikes Back* (1983).

DEEP AND SURFACE STRUCTURE

In the standard approach to generative grammar, sentences are analysed in terms of two levels of organization, known as *deep structure* and *surface structure*. At the 'deep' (or 'underlying') level, a sentence structure is represented in an abstract way, displaying all the factors that govern how it should be interpreted. At the 'surface' level, there is a more concrete representation, giving the string of morphemes that closely corresponds to what we would hear if the sentence were spoken.

This distinction was used to explain sentence ambiguities, by arguing that in such cases a single surface structure correlates with more than one deep structure. An early Chomskyan example was *Flying planes can be dangerous*, which can be related to two underlying sentences: *Planes which fly can be dangerous* and *To fly planes can be dangerous*.

The distinction was also used to relate sentences that have different surface forms but the same underlying structure, as in the case of active and passive sentences. *Cats chase mice* and *Mice are chased by cats* were said to have different surface structures, but the same deep structure.

The interpretation and status of the two notions has altered greatly in generative theory over the years (§65), but the basic insight is one that has achieved widespread recognition in linguistics.

Honorific grammar

Several languages make use of a special set of grammatical contrasts, in which different levels of politeness or respect are expressed, according to the mutual status of the participants (§10). An 'honorific' system, as it is often called, is well developed in several oriental languages, such as Korean, Javanese, Tibetan, and Japanese; and although its use is changing, especially among younger generations of speakers, it still plays an important role in the marking of social relationships.

Japanese honorific expression shares with many other languages certain characteristics of formal speech. Local dialect forms are avoided; loan words are often used (Chinese loans, in the case of Japanese); sentences are longer and involve more circumlocution and negative expression (cf. English 'I wonder whether you mightn't ...'). What differentiates Japanese from European languages is the way in which pronouns, verbs, adjectives, and many types of grammatical construction change their form depending on their honorific status. A large number of special forms are permitted, which are classified into 'respect words' (*sonkeigo*), 'condescending words' (*kenzjo-go*), and 'polite words' (*teinei-go*).

Honorific markers in the morphological system include: (a) a specific honorific prefix, *o-* or *go-*; (b) the complete replacement of a word, e.g. *iu* 'say' becomes *ossharu*; and (c) a complex system of titular forms (where English would say 'Mr, Mrs, Miss'), all suffixes attached to the name:

-sama	very polite
-san	neutral
-chan	diminutive
-kun	for men only
sensei	traditionally used to a person who was 'born earlier', but now used to someone whose capabilities are respected, especially a teacher or politician

A wide range of pronoun forms is used. Among the first-person forms, we find:

watakushi	very formal male; less formal female
watashi	formal male; neutral female
atakushi	rare male; snobbish female
atashi	chiefly female, colloquial
washi	dialectal, chiefly male, older generation
boku	exclusively male, proscribed in talking to superiors (but cf. p. 21)
ore	colloquial male

Among the second-person forms, we find:

anata	standard, polite, not used to superiors
anta	informal
sochira	polite, very formal
kimi	chiefly men to men of equal or lower status
omae	informal, colloquial, somewhat pejorative
kisama and *temē*	derogatory, very impolite

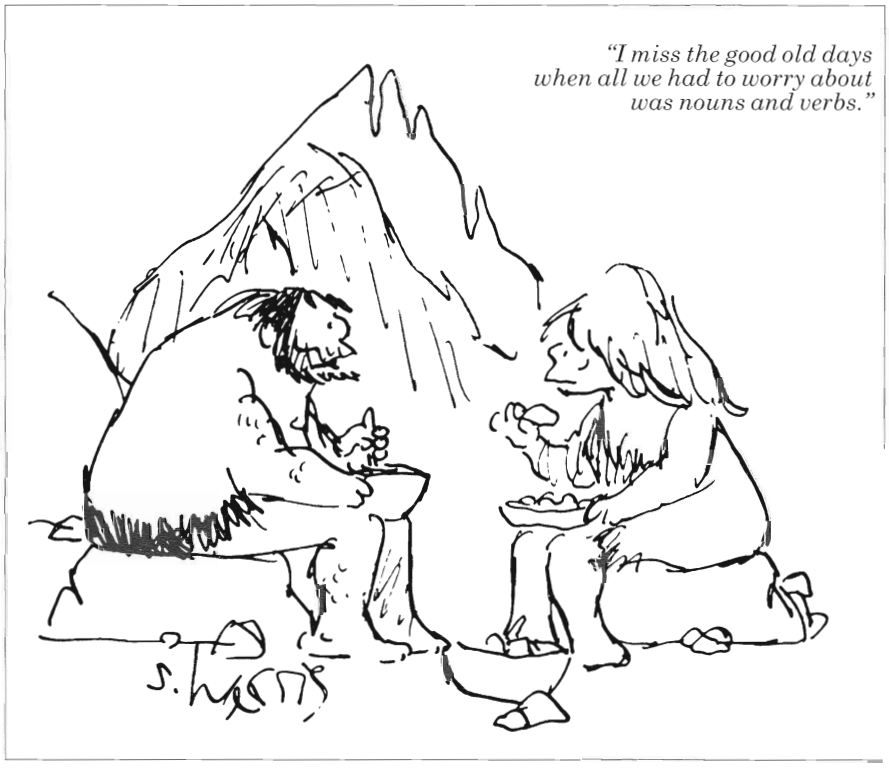
(After S. I. Harada, 1976.)

MISCELLANY

Even a brief survey of grammatical issues leaves one somewhat in awe at the extraordinary variety of patterns that exist in the languages of the world. Repeatedly the lesson is brought home that there is nothing sacrosanct or superior about the grammar of any one language – a lesson that is particularly apposite for English users, whose language holds a special position in modern world society (§59). The following structural differences illustrate this important principle still further.

- English counts in tens and units, as reflected in our number-names: *41* = 'four tens one'. Welsh counts in a mixture of tens and twenties: *20 = ugain, 30 = deg ar hugain* 'ten on twenty', *40 = deugain* 'two twenties', *50 = deg a deugain*, 'ten and forty'. French also makes some use of twenty: *91 = quatre-vingt-onze* 'four-twenty-eleven'. Old Hawaiian made use of forty as a counting unit: 50 was 'forty and ten'; 968 was expressed as 'two four-hundreds and four forties and eight'. Some number systems involve counting backwards: English sees 199 as '100 plus 99'; Yoruba sees it as '200 less 1'. Several languages have no number system: Andamanese makes do with two number-words, one and one-plus. Khoisan languages express one, two, and occasionally three, but rarely more.
- English has a single pair of demonstratives, *this* and *that*, which basically refer to 'near' vs 'further away'. To make other semantic distinctions, we have to use a circumlocution, e.g. 'that one over there'. Japanese has a three-way system: *kono* = near the speaker, *sono* = near the hearer, *ano* = distant from both (in time or place). The Australian language Alyawarra has a four-term system: *nhinha* = 'this', *yanha* = 'that (near)', *nhaka* = 'that (far)', and *awutha* = 'the one mentioned before'. Eskimo has around 30 separate demonstrative forms, expressing such notions as 'that in there', 'that high up there', 'that unseen'.
- English has a single pair of response words that can be used to reply to all questions (other than those beginning with a question-word, such as *why*): *yes* and *no*. In Welsh, there is an indefinite number of response forms, the choice depending on the grammatical form of the question. For example, a question beginning *A oes ...?* (Is there ...?) is replied to by *oes* (yes) or *nag oes* (no); *Ydy Gwen yn mynd?* (Is Gwen going? → *Ydy / Nag ydy, Ydych chi'n mynd?* (Are you going?) → *Ydw / Nag ydw, Allwch ...?* (Can you ...?) → *Galla / Na alla*. The principle underlying this proliferation of forms is straightforward, however. In most cases, the reply simply repeats the verb form, allowing for changes in pronouns. It is as if in English there was a system:

Are you going? Yes-I-go.
Is he there? No-he-isn't.
Did I see ? Yes-you-saw.



17 · SEMANTICS

Semantics is the study of meaning in language. The term did not come to be widely used until the 20th century, but the subject it represents is very old, reaching back to the writings of Plato and Aristotle, and attracting the special interest of philosophers, logicians, and (these days) linguists (§65). The linguistic approach aims to study the properties of meaning in a systematic and objective way, with reference to as wide a range of utterances and languages as possible. It is thus broader than the approach taken by many logicians and philosophers, who have tended to concentrate on a restricted range of sentences (typically, statements, or 'propositions') within a single language. But logical analysis nonetheless exercises a major influence on contemporary linguistic semantics (p. 107).

Any scientific approach to semantics has to be clearly distinguished from a pejorative sense of the term that has developed in popular use, when people talk about the way language can be manipulated in order to mislead the public. A newspaper headline might read 'Unemployment reduced to semantics' – referring to a new way of counting the unemployed which makes it appear that there are fewer of them. Or someone might say in an argument, 'That's just semantics', implying that the point is purely a verbal quibble, bearing no relationship to anything in the real world. This kind of nuance is absent when we talk about semantics from the objective viewpoint of linguistic research.

THE MEANINGS OF MEANING

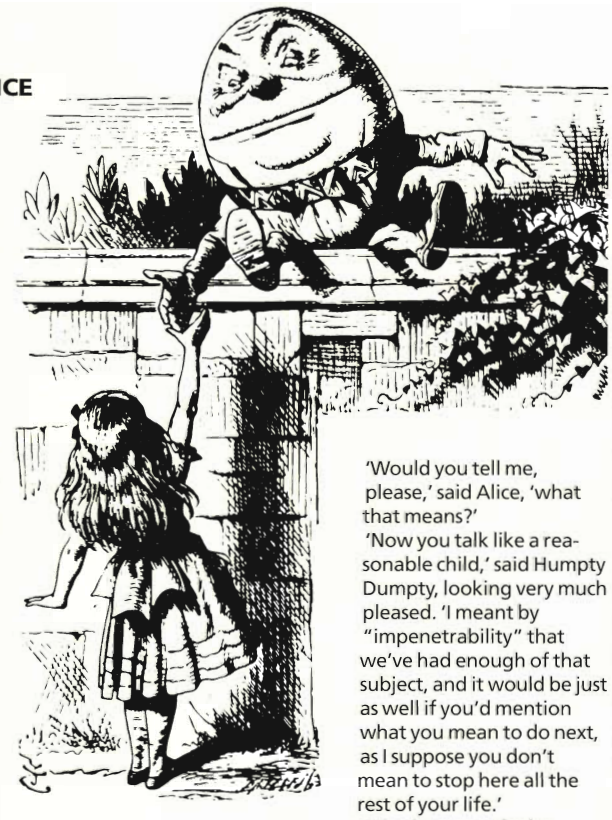
In an important early book on the subject, C. K. Ogden & I. A. Richards's *The Meaning of Meaning* (1923), 16 different meanings of the words 'mean / meaning' were distinguished. Here are some of them:

John means to write. 'intends'
A green light means go. 'indicates'
Health means everything. 'has importance'
His look was full of meaning. 'special import'
What is the meaning of life? 'point, purpose'
What does 'capitalist' mean to you? 'convey'
What does 'cornea' mean? 'refer to in the world'

It is the last kind of use that comes closest to the focus of linguistic semantics; but even this is a special kind of enquiry. The question asks for a definition, which is a somewhat unusual form of reply, found more in dictionaries than in everyday speech, that involves the 'translation' of the difficult word into 'easier' words. The study of the properties of definitions is an important part of semantics, but it is only a part. Of greater importance is the study of the way in which words and sentences convey meaning in everyday situations of speech and writing.

SEMANTICS AND ALICE

One of the favourite quotations of semanticists is from Lewis Carroll's *Through the Looking Glass* (1872, Chapter 6), in which Humpty Dumpty turns our conventional understanding of meaning on its head, and thus makes us see more clearly what it has to be about. If everyone were to use words in an idiosyncratic way, as Humpty suggests, the result would be communication anarchy. Only in certain fields – such as literature (§12) – do we tolerate personal deviations from the semantic norms of the language.



'There's glory for you!'

'I don't know what you mean by "glory,"' Alice said.

Humpty Dumpty smiled contemptuously. 'Of course you don't – till I tell you. I meant "there's a nice knock-down argument for you!"'

'But "glory" doesn't mean "a nice knock-down argument,"' Alice objected.

'When *I* use a word,' Humpty Dumpty said, in rather a scornful tone, 'it means just what I choose it to mean – neither more nor less.'

'The question is,' said Alice, 'whether you *can* make a word mean so many different things.'

'The question is,' said Humpty Dumpty, 'which is to be master – that's all.'

Alice was too much puzzled to say anything: so after a minute Humpty Dumpty began again. 'They've a temper, some of them – particularly verbs, they're the proudest – adjectives you can do anything with, but not verbs – however, *I* can manage the whole lot of them! Impenetrability! That's what *I* say!'

'Would you tell me, please,' said Alice, 'what that means?'

'Now you talk like a reasonable child,' said Humpty Dumpty, looking very much pleased. 'I meant by "impenetrability" that we've had enough of that subject, and it would be just as well if you'd mention what you mean to do next, as I suppose you don't mean to stop here all the rest of your life.'

'That's a great deal to make one word mean,' Alice said in a thoughtful tone.

'When I make a word do a lot of work like that,' said Humpty Dumpty, 'I always pay it extra.'

'Oh!' said Alice. She was too much puzzled to make any other remark.

'Ah, you should see 'em come round me of a Saturday night,' Humpty Dumpty went on, wagging his head gravely from side to side, 'for to get their wages, you know.'

THREE CONCEPTIONS OF MEANING

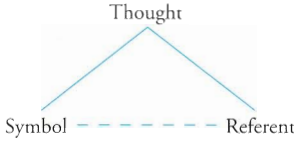
Words → things

A popular view is that words 'name' or 'refer to' things – a view that can be found in the pages of Plato's *Cratylus.* Proper names like *London, Bill Brown,* and *Daddy* illustrate this conception, as do several other words and phrases – the labels attached to objects for sale in a shop, or those found on a paint colour chart. But there are large numbers of words where it is not possible to

see what 'thing' the word refers to: verbs such as *ask* or *find*; adjectives such as *difficult* or *popular*; nouns such as *consistency* or *tradition*. In fact, the majority of words seem unable to be related to things, in any clear way.

Words → concepts → things

This view denies a direct link between words and things, arguing that the relationship can be made only through the use of our minds. For every word, there is an associated concept. One of the best-known formulations of this position is the 'semiotic triangle' of Ogden and Richards (1923, p. 99):



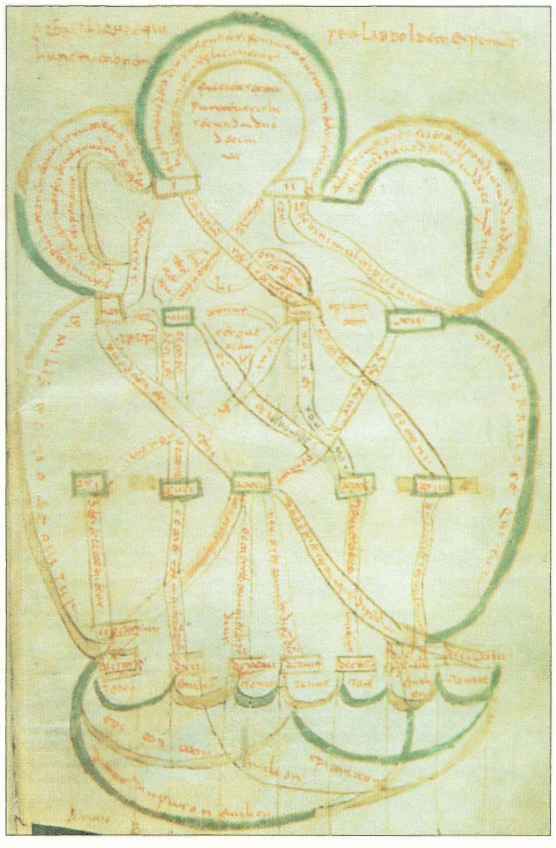
The main criticism of this approach is the insuperable difficulty of identifying 'concepts'. The 'concept' underlying a word such as *tradition* is no easier to define than the 'thing' referred to by *tradition*. Some words do have meanings that are relatively easy to conceptualize, but we certainly do not have neat visual images corresponding to every word we say. Nor is there any guarantee that a concept which might come to mind when I use the word *table* is going to be the same as the one you, the reader, might bring to mind.

Stimuli → words → responses

Leonard Bloomfield (1887–1949) expounded a behaviourist view of meaning in his book *Language* (1933): meaning is something that can be deduced solely from a study of the situation in which speech is used – the stimulus (S) that led someone to speak (r), and the response (R) that resulted from this speech (s). He draws this as follows:

S ——————→ r.....s ——————→ R

In Bloomfield's example, Jill is hungry, sees an apple (S), and asks Jack to get it for her (r); this linguistic stimulus (s) leads to Jack getting the apple (R). Bloomfield argues that you can tell what the meaning of r ... s must be just be observing the events that accompanied it. However, in very many situations it is difficult to demonstrate what the relevant features of the stimulus / response are – a real problem when events are not clearly visible in physical terms (as in the expression of feelings). And it proves even more difficult to handle cases where people do not act in the 'predicted' way (if Jack did not fetch the apple, perhaps because of a quarrel with Jill at Monte Carlo two years before).



A design by Isidore of Seville (c. AD 555–636) The design attempts to show a link between a word's shape and its meaning. Isidore believed that the basic meaning of a word could be found if it could be traced back to its primitive shape. The discussion is found in the ninth book of his *Originum sive etymologiarum libri XX*, which is largely about questions of semantic history and the origins of language.

NATURAL OR CONVENTIONAL?

The Greek philosophers were the first to debate the nature of meaning, from which two main views emerged. The *naturalist* view, deriving largely from Plato (427–347 BC), maintained that there was an intrinsic connection between sound and sense. The *conventionalist* view, largely Aristotelian, held that this connection was purely arbitrary (§65).

In their extreme forms, both views are untenable. If the naturalist view were valid, we would be able to tell the meaning of words just by hearing them. Only onomatopoeic words (§30), such as *bow wow* and *splash*, come close to this, and even they change greatly from language to language. But naturalistic thinking is still widely encountered, especially in the concern many people have over the use of certain words (to do with death or sex, for example, p. 61), or in the readiness with which they make judgments about the appropriateness of words. 'Look at them, sir,' says Aldous Huxley's character Old Rowley, pointing to swine wallowing in the mud, 'Rightly is they called "pigs".' (*Crome Yellow*, 1921).

The conventionalist position is nearer the truth, as it emphasizes the arbitrary relationship between words and things – a principle accepted by modern semanticists. There is nothing in the form of the word *pig* that bears any direct relationship to the 'thing'. But it is equally untenable to think of language, as the conventionalists did, solely as the result of an *agreement* between people to use a word in a certain way. Such a procedure would presuppose the prior existence of language, to formulate the agreement in the first place. Diodorus of Megara (4th century BC) nonetheless supported the conventionalist position to the extent of calling his slaves by the names of Greek particles!

MODERN SEMANTICS

In the past, semantic debate has been largely concerned with discovering what 'meaning' is, as a concept in its own right. The enquiries have undoubtedly increased our understanding of the nature of the problem, but an accepted definition of 'meaning' is as far away today as it was in Plato's time. Why should this be so?

It is now widely held that 'meaning' is not some kind of 'entity' separate from language – any more than measures such as 'height' or 'length' have some kind of independent existence. To say that objects 'have height' means only that they are so many units high; it does not mean that there is an abstract property of 'height' that exists independently of objects. In the same way, it is argued, to say that words 'have meaning' means only that they are used in a certain way in a sentence. We can examine the meaning of individual words and sentences – but there is no 'meaning' beyond that.

In modern linguistics, then, meaning is studied by making detailed analyses of the way words and sentences are used in specific contexts. It is an approach shared by several philosophers and psychologists (p. 418). Ludwig Wittgenstein (1889–1951), in particular, stressed its importance in his dictum: 'the meaning of a word is its use in the language'.

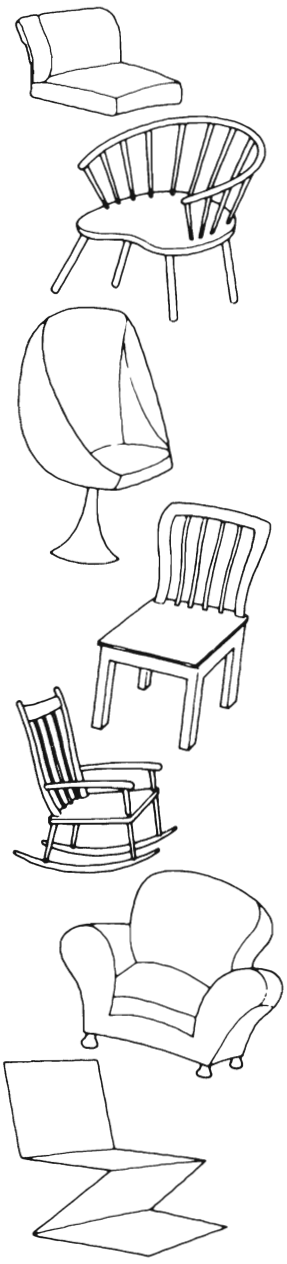
SENSE vs REFERENCE

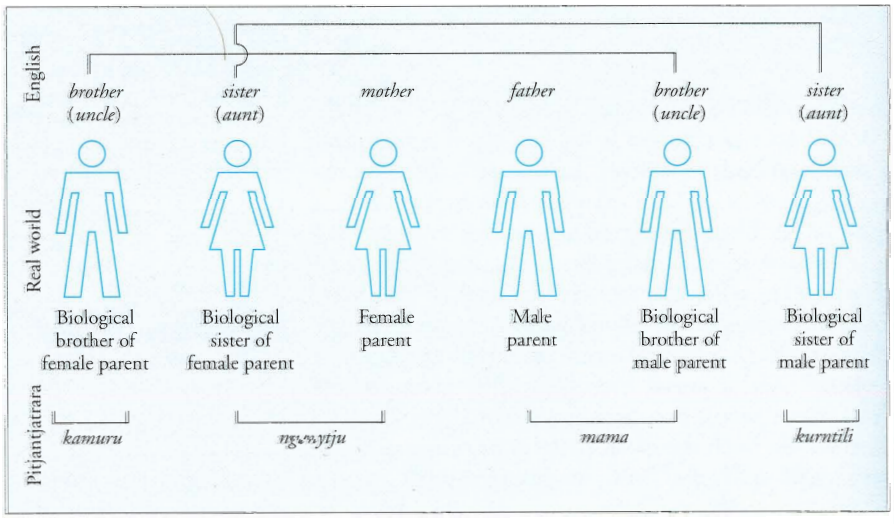
Semantics is not directly concerned with the study of the external world, or its conceptualization. The world of non-linguistic experience is the province of physicists, geographers, psychologists, and others. Nor, as we have seen (p. 101), is semantics easily able to cope with the study of how language *refers* to this external world – the notion of 'reference'. Rather, the primary focus of the modern subject is on the way people relate words to each other within the framework of their language – on their 'sense', rather than their reference.

The distinction between sense and reference is a critical one, because it allows us to study the many cases where we happily use words, even though they do not naturally correspond to the way things are in the world. This may be difficult to see if we restrict our study to a single language, but when we look at how different languages 'parcel out' the world, the distinction is forced upon us. For example, in the 'real' world, mothers and fathers have brothers and sisters. In English, there are no single words expressing the notions 'mother's brother', 'father's brother', 'mother's sister', or 'father's sister', and we have to use a circumlocution to make the distinction. In the Australian language Pitjanjatjara, however, we have a different situation: *ngunytju* = 'mother's sister', *kamuru* = 'mother's brother', *kurntili* = 'father's sister', and *mama* = 'father's brother'. There is also a complication (to English ways of thinking): *mama* also means 'father', and *ngunytju* also means 'mother'. What is plain, though, is that the same biological relationships are given quite different linguistic treatment between the two languages. Family photographs would look the same, but the words would have different senses (see below).

But even within a single language, we need to distinguish sense from reference, to explain the way language makes divisions where there are none in reality. The neat scientific classifications of fauna and flora, where each name has its place in a system of terms, are not typical of language. In everyday life, we use such words as *hill* and *mountain, cup* and *glass,* or *stream* and *river,* where the real-world notions are quite indeterminate. When does a stream become a river, or a hill a mountain? And would all agree about which of the pictures (right) count as a *chair*?

There is also the problem of how we explain what a word's meaning is. Let us imagine someone who had encountered the word *chair* and did not know what it meant. One procedure would be to explain its reference: we could take the person to a chair and point to it. But this would be of limited help, for how would the person know from that experience which *other* objects in the world should also be called chairs? The wrong deduction might also be made, that what we were pointing at was the quality 'wooden', or the concept of 'furniture' – the kind of error children make when they learn vocabulary (§42). A better procedure would be to explain the sense of the word, using a rough definition such as a 'seat with four legs and a back'. Such a definition would enable the person to look out for other objects with similar properties, and thus use the word appropriately. The definition could then be sharpened, as related words were met (e.g. *armchair, stool*). But this whole process of vocabulary learning continues without any direct reference to the objects in the real world: there is total reliance on the use of words to explain the sense of other words – a process that reaches its logical conclusion in a dictionary (§18).





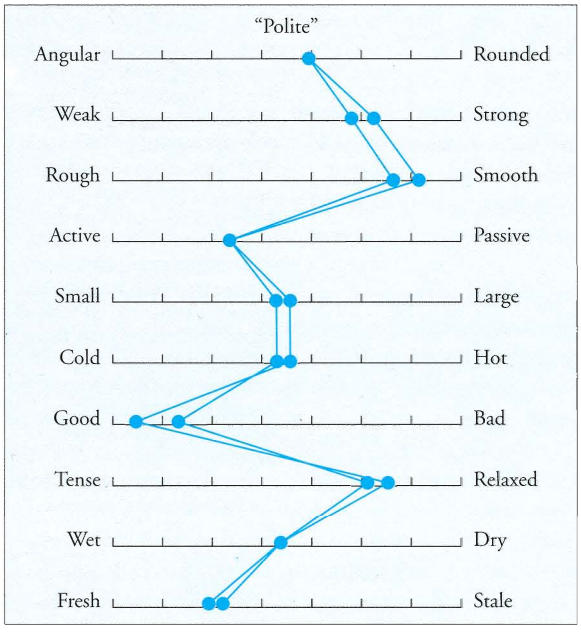
SEMANTIC SPACE

Psychologists also share the concern to establish the semantic properties of individual words, and several approaches have been proposed to plot differences and quantify the psychological 'distance' between words.

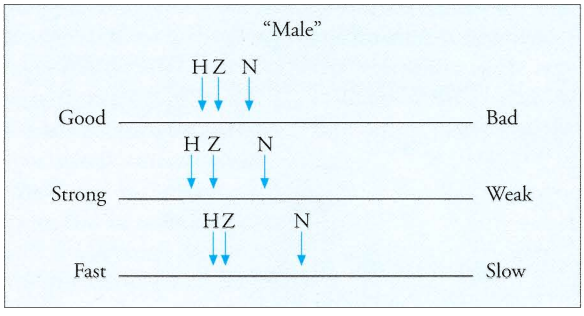
A pioneering work in this field was C. E. Osgood, G. Suci, & P. Tannenbaum, *The Measurement of Meaning* (1957), which was a study of 'affective' meaning – the emotional reactions attached to a word. Each word was subjected to a test that they called a 'semantic differential' – the name reflecting the view that it was possible to analyse meaning into a range of different dimensions. Osgood likened his procedure to a game of Twenty Questions, in which each question (e.g. 'Is it good or bad? fast or slow? small or large?') would aim to locate a concept in semantic space. The questions were presented as seven-point scales, with the opposed adjectives at each end, such as

good ——————— bad

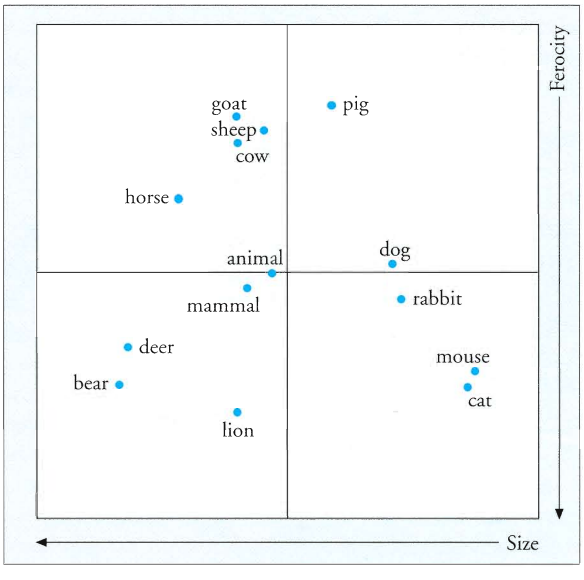
and subjects were asked to rate words in terms of where they would fall on these scales. If they felt that *car* was 'good', for example, they would place a mark towards the 'good' end of the first scale; if 'bad', towards the other end. The seven positions allowed for variations in degree of feeling. Ten of the scales are illustrated below, giving the average responses from the two groups of 20 subjects to the word *polite* (after C. E. Osgood, 1952):



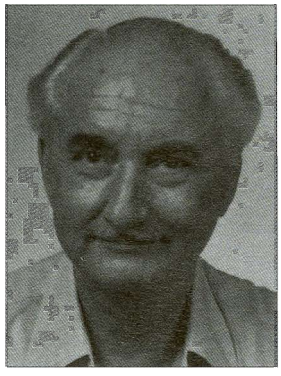
The method was also used to make comparisons between cultural groups. For example, *noise* is a highly affective concept for the Japanese, who tended to react to it using the extremes of the polar scales; it is not so for Americans or Kannada-speaking Indians. The word *male* varies in its connotations between Hopi (H), Zuñi (Z), and Navaho (N) Indians, the first two groups being fairly close together (after H. Maclay & E. E. Ware, 1961).



The semantic differential procedure is a limited one. It does not provide information about the basic meaning of a word but only about the emotions the word generates. It tells us, for example, that *mother* might be 'very good', 'slightly strong', etc., but it does not tell us that the word means 'adult female parent'. To display this kind of information, other ways of working with semantic space are required. We can illustrate this using the results of a technique in which people judge the similarities between words. In the diagram, mammal names are located in a space where the horizontal dimension represents size and the vertical dimension represents ferocity (after L. J. Rips *et al.*, 1973). Larger animals are on the left; more ferocious animals are towards the bottom. The more similar any two animals are thought to be, the closer they are placed in the space. (There is no necessary correspondence with zoological reality, as can be seen from the closeness between cats and mice.)



This is a very simple analysis, which it would be more difficult to make for words where the relevant dimensions of meaning are less clear-cut (items of furniture, for example). But the general approach is illuminating, with considerable research potential.



Charles E. Osgood (1916–)

SEMANTIC STRUCTURE

One of the most productive approaches to the semantic analysis of vocabulary has come from the application of structuralist ideas (§65). From this viewpoint, language is a network of systematic relationships between units. In phonology, for example, the relationships exist between sounds – or phonemes (§28). What are the equivalent semantic units, and how are they related?

Lexemes

So far in this section, we have used the term 'word' to discuss semantic units, and this is the traditional use. People readily talk about the 'meaning of words'. However, if we wish to enquire precisely into semantic matters, this term will not do, and an alternative must be found. There are three main reasons.

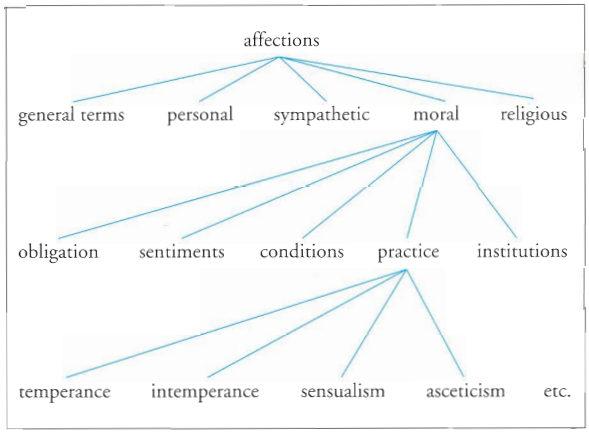
1. The term *word* is used in ways that obscure the study of meaning. The forms *walk, walks, walking,* and *walked* could all be called 'different words'; yet from a semantic point of view, they are all variants of the same underlying unit, 'walk'. If the variants are referred to as 'words', though, what should the underlying unit be called? It would not be particularly clear to say that 'these four words are different forms of the same word'.
2. The term *word* is useless for the study of idioms, which are also units of meaning. A much-used example is *kick the bucket* (= 'die'). Here we have a single unit of meaning, which happens to consist of three words. Again, it would hardly be clear to talk of this unit as a 'word', if we then go on to say that this word consists of three words.
3. The term *word* has in any case been appropriated for use elsewhere in linguistic study – in the field of grammar, where it does sterling service at the junction between syntax and morphology (p. 90).

For such reasons, most linguists prefer to talk about the basic units of semantic analysis with fresh terminology, and both *lexeme* and *lexical item* are in common use. We may now avoid the lack of clarity referred to above, and say that the 'lexeme' WALK occurs in several variant forms – the 'words' *walk, walks,* etc. Similarly, we can say that the 'lexeme' KICK THE BUCKET contains three 'words'; and so on. It is lexemes that are usually listed as headwords in a dictionary. Accordingly, we shall put this term to use in the remaining parts of this section.

SEMANTIC FIELDS

One way of imposing some order on vocabulary is to organize it into 'fields' of meaning. Within each field, the lexemes interrelate, and define each other in specific ways. For example, the various lexemes for 'parts of the body' (*head, neck, shoulders,* etc.) form a semantic field, as do the different lexemes for 'vehicles', 'fruit', 'tools', or 'colour'. It has been argued that the whole of a language's vocabulary is structured into fields; but there is in fact a great deal of variation as we move from one part of the language to another. There would be little difficulty gathering together all the English lexemes for 'body parts', for example; but it would be very difficult to do the same job for 'noise' or 'ornaments'.

There have been many philosophical and linguistic attempts to classify the concepts or words in a language – notably, those associated with the 17th-century quest for a universal language (§58). In recent times, the most influential and popular work has been the *Thesaurus* of Peter Mark Roget (1779–1869), first published in 1852. Roget divided the vocabulary into six main areas: *abstract relations, space, matter, intellect, volition,* and *affections.* Each area was given a detailed and exhaustive sub-classification, producing 1,000 semantic categories in all. One path through the thesaurus is illustrated below:



Groups of words are then listed under each of these headings and classified into the main parts of speech. For example, in the 1962 edition of the work (p. 625), we find the following items listed as a section within *temperance* (numbers refer to other thesaurus sections; keywords are in italics):

abstainer, total a., teetotaller 948n. *sober person*; prohibitionist, pussyfoot; vegetarian, fruitarian, Pythagorean; Encratite; dieter, banter, faster; enemy of excess, Spartan 945n. *ascetic.*

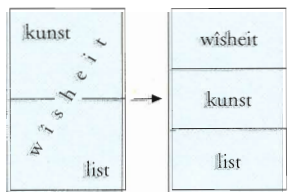
Thesauri of this kind have now been produced for several languages, and prove to be a useful adjunct to many practical linguistic activities, such as professional writing, translating, and setting or solving crosswords. For the semanticist, however, their value is limited, as they contain no information about the sense relationships between individual lexemes, and items that come from different regional, social, or professional varieties (§§8–11) are juxtaposed without comment. To study the structure of a semantic field, more precise means of plotting the sense relations between lexemes need to be used.

Roget (1779–1869) by William Brockedon, 1835

SEMANTIC CHANGE

The linguistic approach to semantic fields was first propounded by German scholars in the 1930s. In one of the earliest studies (J. Trier, 1934), the approach showed how the structure of a semantic field can change over time. Middle High German terms for 'knowledge' changed greatly between 1200 and 1300. In 1200, a German had no separate lexeme for the quality of cleverness. The language contained *kunst* ('courtly skills') and *list* ('non-courtly skills'), and there was also *wîsheit* for any form of knowledge, whether courtly or not, mundane or divine.

A hundred years later, everything was different. *Wîsheit* had developed the restricted meaning of 'religious experience'; *kunst* was beginning to take on the meaning of 'art / skill', and *wizzen* (modern *Wissen*) had more the meaning of 'knowledge'. *List* had left the field entirely, as it had begun to develop pejorative connotations (cf. its sense of 'cunning' or 'trick' in Modern German). The whole of this change can be summarized in the form of two diagrams:



For a similar use of diagrams in the comparison of modern languages, see p. 106.

SENSE RELATIONSHIPS

How are the lexemes of a language organized? To think of them as a list, such as we might find in a dictionary, is highly misleading. There is no semantic reality in alphabetical order; on the contrary, alphabetical order destroys semantic structure, keeping apart lexemes that should belong together (such as *aunt* and *uncle,* or *big* and *little*). Rather, we need to develop an alternative conception, based on our intuitions that groups of lexemes are related in sense.

Accounts of semantic structure recognize several kinds of sense relations between lexemes. Some result from the way lexemes occur in sequences (*syntagmatic* relations); others from the way in which lexemes can substitute for each other (*paradigmatic* relations) (§65). For example, in the sentence *It was a very auspicious* —, English speakers 'know' that the omitted word will be one of a very small set (e.g. *occasion, event*) – unless, of course, a literary or humorous point is being made (*It was a very auspicious kilt*). This would be a syntagmatic semantic relationship. By contrast, the relationship between the following two sentences is a paradigmatic one: *Is that a new radio? No, it's an old radio.* The substitution of *old* for *new* results in a change of meaning that we recognize as an 'opposite'.

Several types of paradigmatic relationship have been recognized, some of which form a familiar part of language syllabuses in school. These include:

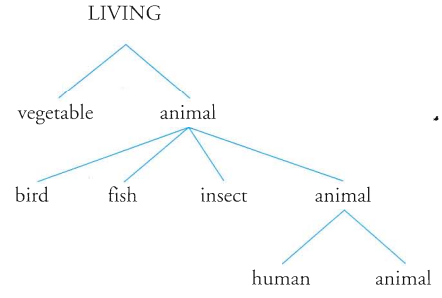
• *Synonymy* This is the relationship of 'sameness' of meaning, e.g. *kingly / royal / regal, pavement / sidewalk, youth / youngster.* The search for synonyms is a long-standing pedagogical exercise, but it is as well to remember that lexemes rarely (if ever) have *exactly* the same meaning. There are usually stylistic, regional, emotional, or other differences to consider. And context must be taken into account. Two lexemes might be synonymous in one sentence but different in another: *range* and *selection* are synonyms in *What a nice — of furnishings,* but not in *There's the mountain —.*

• *Hyponymy* This less familiar relationship refers to the notion of 'inclusion', whereby we can say that 'an X is a kind of Y'. For example, *rose* is a hyponym of *flower, car* of *vehicle.* Several lexemes will be 'co-hyponyms' of the same superordinate term: *rose, pansy, tulip* ... Once again, it must be stressed that this is a linguistic, and not a real-world classification. Languages differ in their superordinate terms, and in the hyponyms they accept under one such term. For instance in classical Greek the lexemes for 'carpenter', 'doctor', 'flautist', and other occupations are all hyponyms of *demiourgos*; but there is no equivalent superordinate term in English. We simply do not have a single 'occupational' term that would allow us to say 'A carpenter / doctor / flautist, etc is a kind of —'. Likewise, *potato* is a hyponym of *vegetable* in English, but *Kartoffel* is not included among *Gemüse* in German (after J. Lyons, 1963).

THE 'ANIMAL' KINGDOM

Animal is a strange lexeme in English, because it can be used at three levels in a hierarchy of inclusion:

1. in a classification of living things, it contrasts with *vegetable,* to include birds, fishes, and insects;
2. it contrasts with *bird, fish,* and *insect* to include humans and beasts;
3. it contrasts with *human.*



• *Antonymy* This is the relationship of 'oppositeness of meaning'. Antonyms are often thought of in the same breath as synonyms, but they are in fact very different. There may be no true synonyms, but there are several kinds of antonyms. Some of the most important types are:

– *gradable* antonyms, such as *big / small, good / bad,* which permit the expression of degrees (*very big, quite small,* etc.);
– *nongradable* antonyms (also called *complementary* terms), which do not permit degrees of contrast, such as *single / married, male / female;* it is not possible to talk of *very male, quite married,* etc., except in jest; and
– *converse* terms: two-way contrasts that are interdependent, such as *buy / sell* or *parent / child;* one member presupposes the other.

• *Incompatibility* Under this heading are grouped sets of lexemes that are mutually exclusive members of the same superordinate category. For example, *red, green,* etc. are incompatible lexemes within the category *colour*: it would not be possible to say 'I am thinking of a single colour, and it is green and red.' On the other hand, *red* is not incompatible with such lexemes as *round* or *dirty* (something can be at once 'red and round'). Terms for fruit, flowers, weekdays, and musical instruments illustrate other incompatible sets. Once again, we must be prepared for some unexpected usages – as in English, where *black, white,* and *grey* are not always included within the category of colour (as with *black-and-white* films and TV sets), and where *red* can be excluded from this category (as with snooker, where one may proceed to play the 'coloured' balls only after all the red balls have been potted).

THE COMPANY LEXEMES KEEP

'You shall know a word by the company it keeps', said the British linguist J. R. Firth (1890–1960) in 1957, referring to the syntagmatic tendency of lexemes to work together ('collocate') in predictable ways. *Blond* collocates with *hair, flock* with *sheep, neigh* with *horse.* Some collocations are totally predictable, such as *spick* with *span,* or *addled* with *brains* or *eggs.* Others are much less so: *letter* collocates with a wide range of lexemes, such as *alphabet* and *spelling,* and (in another sense) *box, post* and *write.* Yet other lexemes are so widely used that they have no predictable collocates at all, such as *have* and *get.*

Collocation should not be confused with 'association of ideas'. The way lexemes work together may have nothing to do with 'ideas'. We say in English *green with jealousy* (not *blue, red,* etc.), though there is nothing literally 'green' about 'jealousy'. *Coffee* can be *white,* though the colour is brown. Both lads and lasses may be well rounded enough to be called *buxom,* but this lexeme is used only with the latter.

Collocations differ greatly between languages, and provide a major difficulty in mastering foreign languages. In English, we 'face' problems and 'interpret' dreams; but in modern Hebrew, we have to 'stand in front of' problems and 'solve' dreams. In Japanese the verb for 'drink' collocates with water and soup, but also with tablets and smoking.

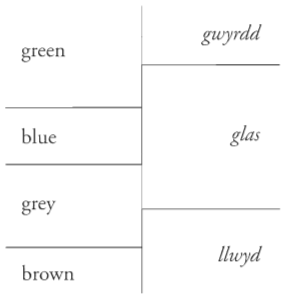
The more fixed a collocation is, the more we think of it as an 'idiom' – a pattern to be learned as a whole, and not as the 'sum of its parts'. Thus we find French *broyer du noir* (lit. 'grind' + 'black'), meaning to 'have the blues' or 'be browned off' – a nice instance of the arbitrary use of colour terms.

Collocations are quite different from the idiosyncratic links between ideas that can be verbally expressed. On a psychiatrist's couch, we may 'free associate', responding to *farm* with *Easter,* or *jam* with *mother.* This is not collocation, which is a link between lexemes made by *all* who speak a language.

COLOUR LEXEMES

The range of colours is a continuous band, lacking any clear physical boundaries. The semantic field of colour has therefore attracted particular attention because it demonstrates very clearly the different patterns of lexical use in a language. English has 11 basic colour lexemes: *white, black, red, green, yellow, blue, brown, purple, pink, orange,* and *grey.* In contrast:

- There were no generic lexemes for 'brown' or 'grey' in Latin; modern Romance forms (such as French *brun, gris*) have been borrowed from Germanic. Navaho has a single lexeme for both.
- Navaho also makes no lexical distinction between 'blue' and 'green'. On the other hand, it has two terms for 'black', distinguishing the black of darkness from the black of such objects as coal.
- Russian makes a distinction between two kinds of 'blue', *sinij* vs *goluboj*, where English has to use circumlocutions: 'dark blue' vs 'sky blue'. Hungarian has two terms for 'red'.
- Japanese *ao* can mean 'green', 'blue', or 'pale', depending on context (e.g. vegetables, sea, clouds).
- In Hanunóo, there are just four basic colour terms, 'black', 'white', 'red', and 'green'.
- Some New Guinea Highland languages have terms only for 'black' and 'white' – perhaps better translated as 'dark' vs 'light'.
- In some languages the situation is more difficult to express in words, and a field diagram is clearer. Literary Welsh, for example, divides the green–brown part of the spectrum quite differently from English:



Modern Welsh is similar to English, but even so, *glas* is used for the colour of growing things (though it otherwise is equivalent to *blue*).

Colour universals?

The differences between the colour terms of various languages are striking, and might lead us to conclude that each language has worked out a unique system in a totally arbitrary way. A 1969 study by B. Berlin & P. Kay, however, argued the opposite. After studying the colour systems of 98 languages, they concluded that there is a universal inventory of only 11 basic colour categories, and all languages use either these 11 or fewer. 'Basic' was interpreted to mean that the terms used only a single morpheme (excluding *light brown*, etc.), were in common use (excluding *indigo*), applied to many objects (excluding *blond*), and were not contained within another colour (excluding *scarlet*). They also claimed (p. 25) that these basic terms were ordered, as follows:

[white, black] <[red]< [green, yellow] <[blue]<[brown]< [purple, pink, orange, grey]

If a language has a term to the right of the sign <, it will also have all the terms to the left.

These claims are not without controversy. Obtaining reliable data from native speakers about such matters is a problem, especially as their judgments might have been coloured by their exposure to other languages. Some languages, also, seem to have 12 basic terms (e.g. Russian). But the research has demonstrated some impressive similarities across a wide range of languages.

POLYSEMY OR HOMONYMY?

- *Polysemy* refers to cases where a lexeme has more than one meaning: for example, *chip* can mean a piece of wood, food, or electronic circuit. People see no problem in saying that 'the word *chip* has several different meanings in English'.
- *Homonymy* refers to cases where two (or more) different lexemes have the same shape: for example, *bank* is both a building and an area of ground. Again, people see no problem in saying that 'these are two different words in English'.

This second reaction would also be given to those cases where lexemes were only 'half' identical in shape:

– *homophones*, which have the same pronunciation, but different spelling (e.g. *threw* vs *through*);
– *homographs*, which have the same spelling, but different pronunciation (e.g. *wind* – air movement vs bend).

The distinction seems clear enough, and dictionaries treat cases of multiple meaning either as polysemy or as homonymy. But in fact it is not always easy to decide which we are dealing with, and dictionaries sometimes differ in their decisions. Are *table* (furniture) and *table* (arrangement of data) two different words, or the same word with two meanings? Dictionaries usually go for the latter solution, on grounds of a shared etymology. On the other hand, *pupil* (in school) and *pupil* (of the eye) are usually listed as different words – though in fact they have the same historical origin. French *voler* 'fly' and *voler* 'steal' are similar: they are now thought of as different words, but both derive from Latin *volare*. There is often a conflict between historical criteria and present-day intuition, in sorting out cases of polysemy and homonymy.

KINSHIP CONTRASTS

Another semantic field which has been much studied is that of kinship. Here too there are interesting differences between languages:

- Hungarian had no terms for 'brother' or 'sister' until the 19th century, though it did have separate terms for 'elder' and 'younger' brothers and sisters.
- Malay has a generic term for both 'sibling' and 'cousin'.
- There is no single term for 'grandfather' or 'grandmother' in Swedish: *farfar* = 'father's father', *morfar* = 'mother's father', *farmor* = 'father's mother', *mormor* = 'mother's mother'.
- In Njamal (Australia), some terms express generation distance, e.g. a man can use *maili* both for 'father's father' and 'daughter's son's wife's sister' – both are two generations away.
- Latin distinguished 'father's brother' (*patruus*), 'father's sister' (*matertera*), 'mother's brother' (*avunculus*), and 'mother's sister' (*amita*), but modern Romance languages have reduced these to two (e.g. French *oncle* and *tante*, derived from the maternal terms).

DEIXIS

Every language has a set of lexemes which can be interpreted only with reference to the speaker's position in space or time. These are known as *deictic* forms (from the Greek word for 'pointing'), and the conditions governing their use have attracted especial attention in recent semantics. They fall into three main types.

- *Personal deixis* The use of pronouns, such as *I* and *you*, which identify who is taking part in the discourse.
- *Spatial deixis* Forms that distinguish the position of the speaker in relation to other people or objects, such as *this / that, here / there* (p. 99), *bring / take, come / go. Come,* for example, implies direction towards the speaker – *Come here!* (but not **Go here!*).
- *Temporal deixis* Forms that distinguish time with reference to the speaker, such as *now, yesterday, then,* and the various kinds of tense marker.

SEMANTIC COMPONENTS

A further way to study lexical meaning is by analysing lexemes into a series of semantic features, or components. *Man,* for example, could be analysed as ADULT, HUMAN, and MALE. The approach was originally devised by anthropologists as a means of comparing vocabulary from different cultures, and it has been developed by semanticists as a general framework for the analysis of meaning.

Whole systems of relationships can be established, using a small set of components. For example, the components ADULT/NON-ADULT and MALE/FEMALE can be used for the following:

man (ADULT, MALE), woman (ADULT, FEMALE)
boy (NON-ADULT, MALE), girl (NON-ADULT, FEMALE).

Many animals display a similar pattern (though lacking a male / female non-adult distinction):

MALE	FEMALE	NON-ADULT
bull	cow	calf
ram	ewe	lamb
boar	sow	piglet

In componential analysis, contrasts are usually presented in terms of + or −, and often drawn in a matrix. Thus, we could use +MALE and −MALE (or, of course +FEMALE and −FEMALE) to summarize the above possibilities:

	bull	ram	boar	cow	ewe	sow	calf	lamb	piglet
MALE	+	+	+	-	-	-	+-	+-	+-
FEMALE	-	-	-	+	+	+	+-	+-	+-

The analyses become more interesting, as the lexemes become more complex. Here, for instance, is a possible matrix for some human motion verbs.

	NATURAL	HURRIED	FORWARD	ONE FOOT ALWAYS ON GROUND
walk	+	−	+	+
march	−	+	+	+
run	−	+	+	−
limp	−	−	+	+

It is easy, using a system of this kind, to see what lexical gaps there are in a language. For example, this matrix suggests there is no single English lexeme expressing the notion of 'human using legs to move backwards'. On the other hand, it is not always so easy to decide which are the relevant components of a lexeme and whether they can be applied in a binary (+/−) way. Would *swim* be +HURRIED or −HURRIED in this matrix? Or, in other fields, would *soup* be +EAT or −EAT, and *porridge* +LIQUID or −LIQUID?

SENTENCE MEANING

The study of meaning takes us by degrees through the whole of a language, and it proves difficult to draw a neat line around the semantic component of any linguistic framework (§13). Much of the focus of traditional semantics has been on vocabulary, but contemporary semantics is increasingly concerned with the analysis of sentence meaning – or, at least, of those aspects of sentence meaning that cannot be predicted from the 'sum' of the individual lexemes.

- *Prosodic meaning* The way a sentence is said, using the prosody of the language (§29), can radically alter the meaning. Any marked change in emphasis, for example, can lead to a sentence being interpreted in a fresh light. Each of these sentences carries a different implication, as the stress (indicated by capitals) moves:

John's bought a red CAR (not a red bicycle).
John's bought a RED car (not a green one).
JOHN's bought a red car (not Michael).

The prosody informs us of what information in the sentence can be taken for granted (is 'given') and what is of special significance (is 'new').

- *Grammatical meaning* The categories that are established by grammatical analysis can also be analysed from a semantic point of view. A sentence such as *John read a book yesterday* consists of Subject + Verb + Object + Adverbial (p. 95); but it can also be analysed as an 'actor' performing an 'action' on a 'goal' at a certain 'time'. There is a great deal to be said about the 'semantic roles' played by syntactic elements – an area of study that falls uneasily between semantics and grammar.
- *Pragmatic meaning* The function performed by the sentence in a discourse needs to be considered. The meaning of the sentence *There's some chalk on the floor* seems plain enough; but in some situations it would be interpreted as a statement of fact ('Have you seen any chalk?') and in others as a veiled command (as when a teacher might point out the chalk to a child in class). The pragmatic study of sentence function is reviewed in §21, but it overlaps greatly with the field of semantics—especially the 'semantics of misunderstanding'.
- *Social meaning* The choice of a sentence may directly affect the social relationships between the participants. We may convey such impressions as politeness, rudeness, competence, or distance, and this will affect our status and role within a community. 'What do you mean by talking to me like that?' is a question that raises larger issues than the meaning of the individual lexemes and sentences that have been used.
- *Propositional meaning* Perhaps the most important trend in modern semantics is the investigation of sentence meaning using ideas derived from philosophy and logic. In this kind of approach, a careful distinction is drawn between sentences (grammatical units, p. 94) and propositions. A proposition is the unit of meaning that identifies the subject matter of a statement; it describes some state of affairs, and takes the form of a declarative sentence, e.g. *Mary loves Michael.* In such theories as 'truth-conditional semantics', sentences are analysed in terms of the underlying propositions they express, and these propositions are then tested to see whether they would be true or false, in relation to the real world. The theories are controversial, and require not a little expertise in formal logic to be understood. But they may in due course provide a level of general explanation for semantic observations that the subject has hitherto lacked.

GRAMMAR OR SEMANTICS

The uncertain boundary between semantics and grammar is a classic problem in linguistic theory. It can be illustrated by the many sentences that are used in a habitual manner, and are thus semi-idiomatic in type, falling midway between the 'straightforward' idioms such as *raining cats and dogs* and clear cases of sentences which follow the normal rules of grammar, such as *The man kicked the ball.*

In one study, a large number of habitually used expressions were collected, based on the lexeme *think.* They included:

Come to think of it ...
What do you think?
I thought better of it.
Think nothing of it.
Think it over.
It doesn't bear thinking about.
I thought you knew.
I think so.
What I think is ...
I was just thinking aloud.
Who'd have thought it?
Who do you think you are?
(After A. Pawley & F. H. Syder, 1983, pp. 213–14.)

It is argued that people have memorized expressions of this kind, as part of the process of building up fluent connected speech (the phenomenon is less obvious in the written language). On the other hand, these 'lexicalized sentence stems', as they were called, are plainly not as 'fixed' in their structure as conventional idioms, and their meaning can be predicted quite accurately from their constituent lexemes (unlike, say *raining cats and dogs*). The result is an area of usage that lies midway between the domain of grammar, which focuses on productive sentence types, and that of the lexicon, which focuses on the properties of particular lexical items.

18 · DICTIONARIES

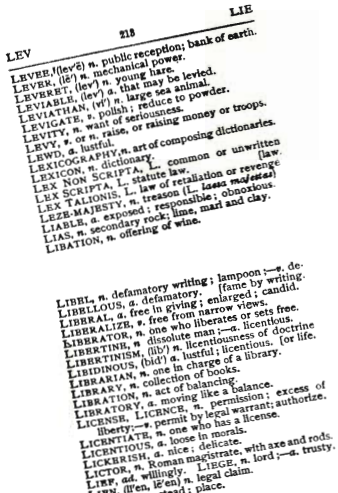
A dictionary is a reference book that lists the words of one or more languages, usually in alphabetical order, along with information about their spelling, pronunciation, grammatical status, meaning, history, and use. The process of compiling dictionaries is known as *lexicography*, and the people who carry out this task are *lexicographers* – 'harmless drudges', as Dr Johnson defined them.

In literate societies, most homes have a dictionary, but there is enormous variation in the way this is used. Some people constantly use them as a serious educational tool, aiming to improve their own or their children's 'word power'. Others use them only for fun – as the arbiter in a game of Scrabble, for instance (p. 64). Others do not use them at all and do not replace them when they fall badly out of date. The continued use of 10- or 20-year-old dictionaries is by no means uncommon.

For a book that is viewed with a level of respect normally accorded only to the Bible, it is remarkable how casually dictionary-users treat their dictionaries. When people are asked what factors govern their choice of dictionary, most cite linguistically irrelevant matters, such as price, pictorial content, and size – not in terms of number of entries, but whether it would fit on a shelf, or in a pocket. Many people expect a dictionary to contain encyclopedic information about historical events, people, and places. Most admit they have never bothered to read the Preface to their dictionary – the place where the layout and conventions of the book are systematically explained. As a consequence they are unable to say what the various abbreviations and symbols mean, or why they are there. The general conclusion is inescapable: most people who would check out every tiny feature of their new car before buying it are unaware of the power that lies under the bonnet of their dictionary.

THE RANGE OF DICTIONARIES

Dictionaries come in all shapes and sizes, from the massive unabridged works, such as the 2,662-page Merriam-Webster *Third New International Dictionary of the English Language* (1961), to the tiniest of pocket-size works, such as the 386-page Nimmo's *Thumb English Dictionary*, each page of which is less than 9 cm high. Coverage (the number of headwords the work contains) and treatment (the kind of information provided under each headword) thus vary enormously. The simplest way of showing this is to compare the amount of detail given for the same word in dictionaries of different sizes. Here are the entries for *insular* in a large, medium-sized, and small dictionary:


LEV 218 LIE

LEVER, (lēv) n. public reception; bank of earth.
LEVER, (lĕ') n. mechanical power.
LEVERET, (lev') n. young hare.
LEVIABLE, (ev') a. that may be levied.
LEVIATHAN, n. large sea animal.
LEVIGATE, v. polish; reduce to powder.
LEVITY, n. want of seriousness.
LEVY, v. or n. raise, or raising money or troops.
LEWD, a. lustful.
LEXICOGRAPHY, n. art of composing dictionaries.
LEXICON, n. dictionary.
LEX NON SCRIPTA, L. common or unwritten law.
LEX SCRIPTA, L. statute law.
LEX TALIONIS, L. law of retaliation or revenge
LEZE-MAJESTY, n. treason (L. laesa majestas)
LIABLE, a. exposed; responsible; obnoxious.
LIAS, n. secondary rock; lime, marl and clay.
LIBATION, n. offering of wine.

LIBEL, n. defamatory writing; lampoon;—v. de-
LIBELLOUS, a. defamatory. [fame by writing.
LIBERAL, a. free in giving; enlarged; candid.
LIBERALIZE, v. free from narrow views.
LIBERATOR, n. one who liberates or sets free.
LIBERTINE, n. dissolute man;—a. licentious.
LIBERTINISM, (lib') n. licentiousness of doctrine
LIBIDINOUS, (bid') a. lustful; licentious. [or life.
LIBRARIAN, n. one in charge of a library.
LIBRARY, n. collection of books.
LIBRATION, n. act of balancing.
LIBRATORY, a. moving like a balance.
LICENSE, LICENCE, n. permission; excess of
liberty;—v. permit by legal warrant; authorize.
LICENTIATE, n. one who has a license.
LICENTIOUS, a. loose in morals.
LICKERISH, a. nice; delicate.
LICTOR, n. Roman magistrate, with axe and rods.
LIEF, ad. willingly. LIEGE, n. lord;—a. trusty.
LIEN, (lē'en, lē'en) n. legal claim.
LIEU, (lū) n. stead; place.

A page from the 'thumb' dictionary (75% of real size)

The Random House Dictionary of the English Language (c. 260,000 headwords, 1987)

(in′sə lər, ins′yə-), *adj.* **1.** of or pertaining to an island or islands: *insular possessions.* **2.** dwelling or situated on an island. **3.** forming an island: *insular rocks.* **4.** detached; standing alone; isolated. **5.** of, pertaining to, or characteristic of islanders. **6.** narrow-minded or illiberal; provincial: *insular attitudes towards foreigners.* **7.** *Pathol.* occurring in or characterized by one or more isolated spots, patches, or the like. **8.** *Anat.* pertaining to an island of cells or tissue, as the islets of Langerhans. –*n.* **9.** an inhabitant of an island; islander. [1605–15; <LL *insulār* (*is*). See INSULA, -AR¹] –**in′su . lar . ism,** *n.* –**in′su . lar′i . ty,** *n.* –**in′su . lar . ly,** *adv.*

Longman Dictionary of the English Language (c. 90,000 headwords, 1984)

/ˈinsyoolə/ *adj* **1a** of or being an island **b** living or situated on an island <*~residents*> **2** *of a plant or animal* having a restricted or isolated natural range or habitat **3a** of island people <*surviving customs*> **b** that results (as if) from lack of contact with other peoples or cultures; narrow-minded, illiberal **4** *anatomy* of an island of cells or tissue [LL *insularis,* fr L *insula* island] – **insularism** *n,* **insularly** *adv,* **insularity** *n*

Penguin English Dictionary (c. 40,000 headwords, 2nd edn, 1969)

[*in*sewler] *adj* of or like an island; of or like inhabitants of an island; cut off from general currents of thought; narrow-minded, smugly intolerant.

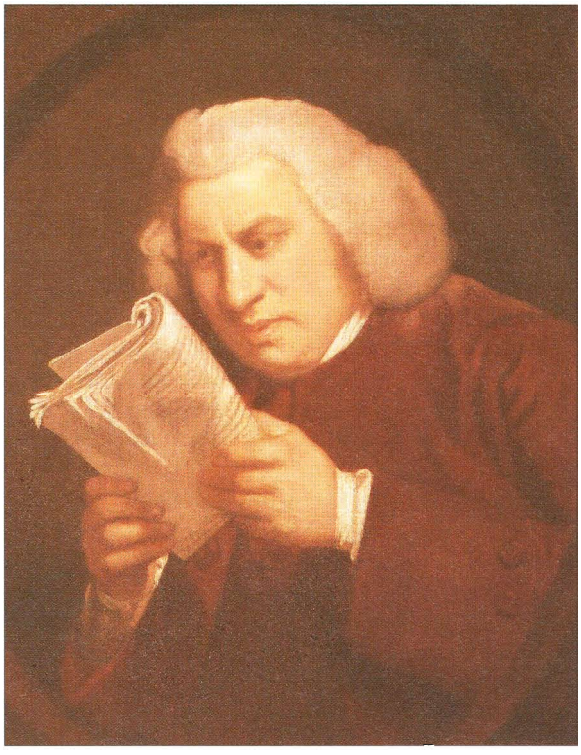
For further comparison, Nimmo's tiny book says simply: '*a.* surrounded by water'.

Apart from variations of format – the use of bold face, numbered senses, etc. – there are major differences in the range and depth of information provided. It is worth spending five minutes making a point-by-point comparison, to see exactly how much information is lost as the dictionaries become smaller. And the moral is plain: for serious study of a language's word-hoard, only the largest dictionaries will suffice.

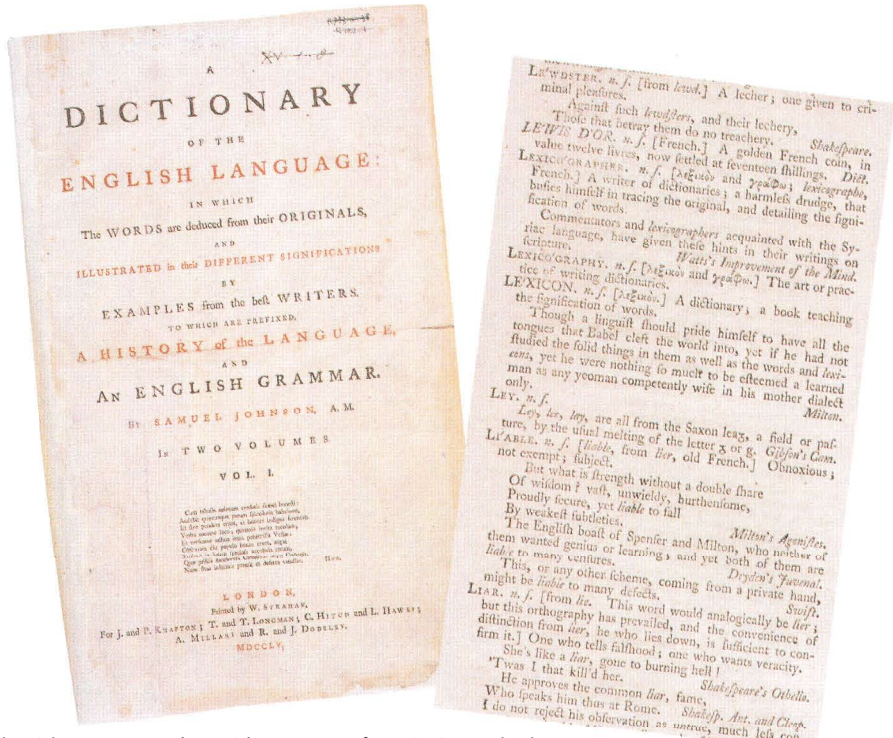
HOW BIG IS A DICTIONARY?

Dictionaries usually claim to contain 'X,000' words. But this grand total can mean several different things. It might refer just to the number of *headwords* in the dictionary – that is, the bold-face items that occur at the beginning of each entry. Or it might include in addition all the subsidiary bold-face items that occur within an entry: under *quick*, for example, there will be *~ly* and *~ness*. Different word classes might be counted separately (e.g. *play* noun vs *play* verb), as might idioms, and irregular grammatical forms (e.g. *go, went*). Depending on what you decide to count, you can end up with two very different totals for the same dictionary. Claims about size should therefore be viewed with caution.

The best way to evaluate the coverage of a dictionary is to compare the words and senses it includes with another dictionary of about the same size. It is notable how even the largest dictionaries present great differences in their coverage – the variation being particularly noticeable in the way they treat world regional vocabulary (how many Australian, South African, or West Indian forms does an English dictionary include, for example?), local dialect words, abbreviations, slang and sub-standard forms, new coinages and borrowings. The use of illustrations and the inclusion of encyclopedic information (names of people, places, historical events, etc.) is also a major source of difference, especially between British and American dictionaries. It has been estimated that the lack of correspondence in large English dictionaries can be as great as 50% – indicating that a truly comprehensive dictionary of the language has yet to be compiled.



Dr Samuel Johnson (1709–84) by Joshua Reynolds, 1775.



A
DICTIONARY
OF THE
ENGLISH LANGUAGE:
IN WHICH
The WORDS are deduced from their ORIGINALS,
AND
ILLUSTRATED in their DIFFERENT SIGNIFICATIONS
BY
EXAMPLES from the beſt WRITERS.
TO WHICH ARE PREFIXED,
A HISTORY of the LANGUAGE,
AND
AN ENGLISH GRAMMAR.
By SAMUEL JOHNSON, A.M.
IN TWO VOLUMES.
VOL. I.
LONDON,
Printed by W. STRAHAN;
For J. and P. KNAPTON; T. and T. LONGMAN; C. HITCH and L. HAWES;
A. MILLAR and R. and J. DODSLEY.
MDCCLV.

LE'WDSTER, *n. ſ.* [from *lewd.*] A lecher; one given to criminal pleaſures.
Againſt ſuch *lewdſters*, and their lechery,
Thoſe that betray them do no treachery. *Shakeſpeare.*
LEWIS D'OR, *n. ſ.* [French.] A golden French coin, in value twelve livres, now ſettled at ſeventeen ſhillings. *Dict.*
LEXICOGRAPHER, *n. ſ.* [λεξικὸν and γράφω; *lexicographe*, French.] A writer of dictionaries; a harmleſs drudge, that buſies himſelf in tracing the original, and detailing the ſignification of words.
Commentators and *lexicographers* acquainted with the Syriac language, have given theſe hints in their writings on ſcripture. *Watts's Improvement of the Mind.*
LEXICOGRAPHY, *n. ſ.* [λεξικὸν and γράφω.] The art or practice of writing dictionaries.
LEXICON, *n. ſ.* [λεξικὸν.] A dictionary; a book teaching the ſignification of words.
Though a linguiſt ſhould pride himſelf to have all the tongues that Babel cleft the world into, yet if he had not ſtudied the ſolid things in them as well as the words and *lexicon*, yet he were nothing to muckt to be eſteemed a learned man as any yeoman competently wiſe in his mother dialect only. *Milton.*
LEY, *n. ſ.*
Ley, *ley*, *ley*, are all from the Saxon *leag*, a field or paſture, by the uſual melting of the letter *g* or *g*. *Gibſon's Cam.*
LI'ABLE, *n. ſ.* [*liable*, from *lier*, old French.] Obnoxious; not exempt; ſubject.
But what is ſtrength without a double ſhare
Of wiſdom? vaſt, unwieldy, burthenſome,
Proudly ſecure, yet *liable* to fall
By weakeſt ſubtleties. *Milton's Agoniſtes.*
The Engliſh boaſt of Spenſer and Milton, who neither of them wanted genius or learning; and yet both of them are *liable* to many cenſures. *Dryden's Juvenal.*
This, or any other ſcheme, coming from a private hand, might be *liable* to many defects. *Swift.*
LIAR, *n. ſ.* [from *lie*. This word would analogically be *lier*; but this orthography has prevailed, and the convenience of diſtinction from *lier*, he who lies down, is ſufficient to confirm it.] One who tells falſhood; one who wants veracity.
She's like a *liar*, gone to burning hell!
'Twas I that kill'd her. *Shakeſpeare's Othello.*
He approves the common *liar*, fame,
Who ſpeaks him thus at Rome. *Shakeſp. Ant. and Cleop.*
I do not reject his obſervation as untrue, much leſs con-

The title page, together with an extract from Dr Samuel Johnson's *Dictionary of the English Language*

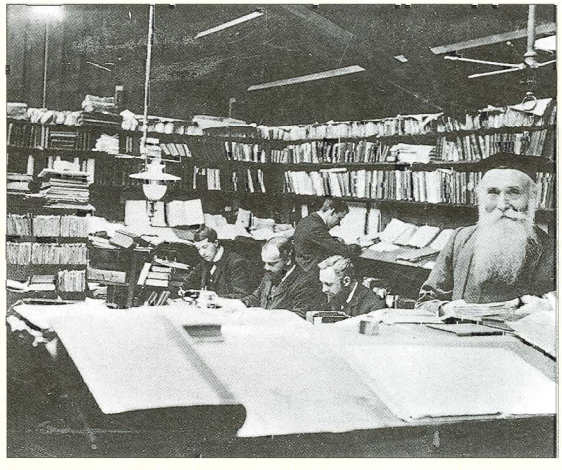
JOHNSON'S DICTIONARY

This great dictionary was published in two volumes in 1755, and its influence on subsequent lexicography was unequalled. Noah Webster went so far as to compare Johnson's contribution to Newton's in mathematics. The dictionary had four main features that set it off from previous English works:

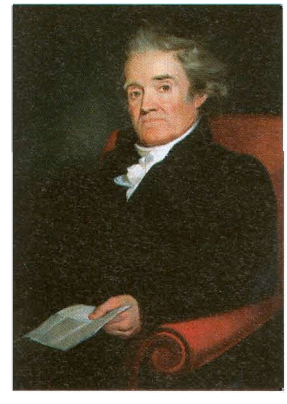
- It aimed to be a scholarly record of the whole of a language – a marked contrast with the haphazard dictionaries of 'hard words' previously compiled.
- It was based on a corpus of examples of usage, largely from the period 1560 to 1660; certain types of words were excluded (e.g. proper names – a decision which resulted in a major difference between British and American dictionaries, for while the former followed Johnson, the latter did not).
- It introduced a literary dimension, departing from the previous concentration on technical language. Half of all Johnson's quotations come from Shakespeare, Dryden, Milton, Addison, Bacon, Pope, and the Bible. The dictionary is very much the language of the 'best' authors.
- Dictionaries become more authoritative – and authoritarian – as a consequence. They became increasingly used in a normative way, as guides to *good* usage – a bias which 20th-century dictionaries are only just beginning to correct.

However, Johnson's was not the first monolingual English dictionary. The credit for this must go to Robert Cawdrey's *A Table Alphabeticall ...*, a 2,500-entry work printed in 1604.

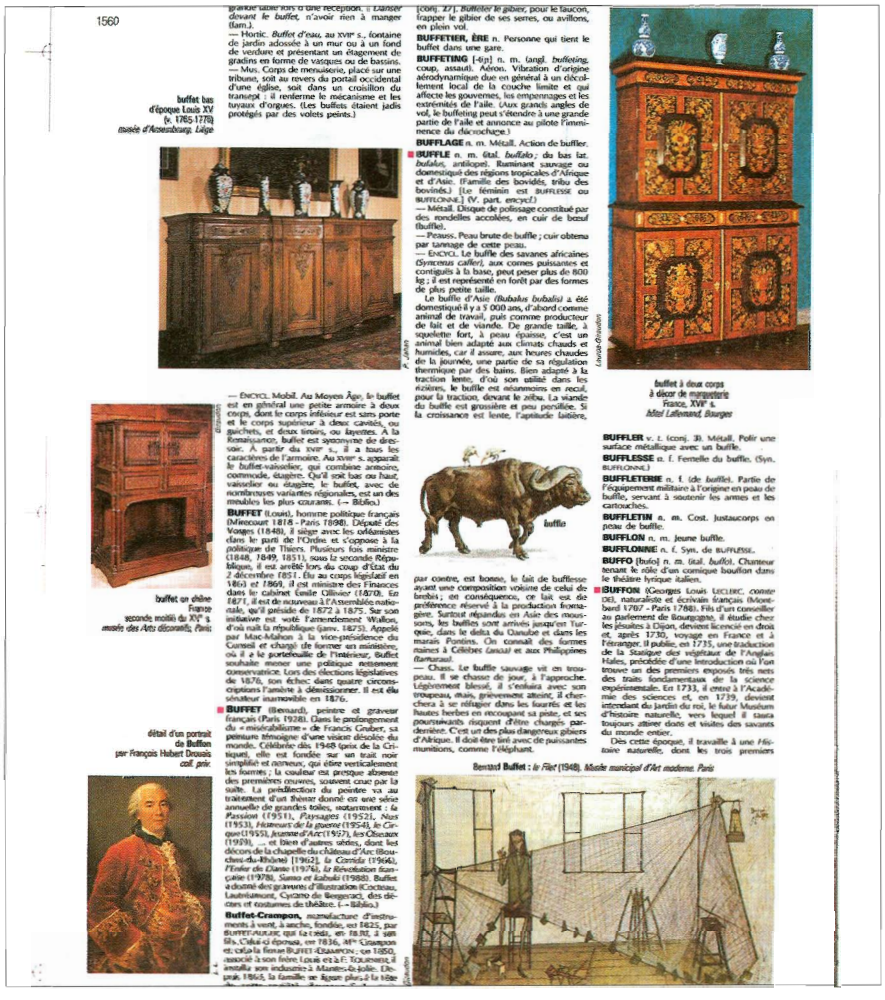
From Johnson's Preface
My purpose was to admit no testimony of living authors, that I might not be misled by partiality, and that none of my contemporaries might have reason to complain ... I have studiously endeavoured to collect examples and authorities from the writers before the restoration, whose words I regard as *the wells of English undefiled,* as the pure sources of genuine diction.



James Murray (1837–1915) The first editor of the *Oxford English Dictionary* is seen at work in his Scriptorium. Murray, the son of a village tailor from Hawick, Scotland, was largely self-educated. He left school at 14, and was a teacher and bank clerk before becoming a lexicographer. His Scriptorium, where most of the editing took place, was built in the back garden of his house in Oxford. Murray planned the whole of the New English Dictionary, and he edited more than half of its first edition himself. The first instalment was published in 1884: A–ANT, 352 pages, price 12s. 6d. (or 62½p today. It took 44 years to complete the dictionary, in 125 instalments – four times longer than had been expected. The complete work, totalling 424,825 entries, was then published as the *Oxford English Dictionary (OED)* from 1933 onwards, with four supplements issued from 1972. An integrated edition appeared in 1989, and was also released on compact disk; a third edition is planned for 2005.



Noah Webster (1758–1843)
Webster's works on spelling, grammar, and lexicon constituted the first major account of American English, and gave the variety a clear identity and status. His *An American Dictionary of the English Language*, published in two volumes in 1828, consisted of around 70,000 entries. Particular attention was paid to the inclusion of scientific terms, and to etymological background. The latest revision is the *Third New International Dictionary* (Merriam, 1961), containing over 450,000 entries, which took 757 editor-years to complete. Supplements appeared in 1976, 1983, and 1986.



1560

buffet bas d'époque Louis XV (v. 1765-1770) musée d'Ansembourg, Liège

buffet en chêne France seconde moitié du XVIe s. musée des Arts décoratifs, Paris

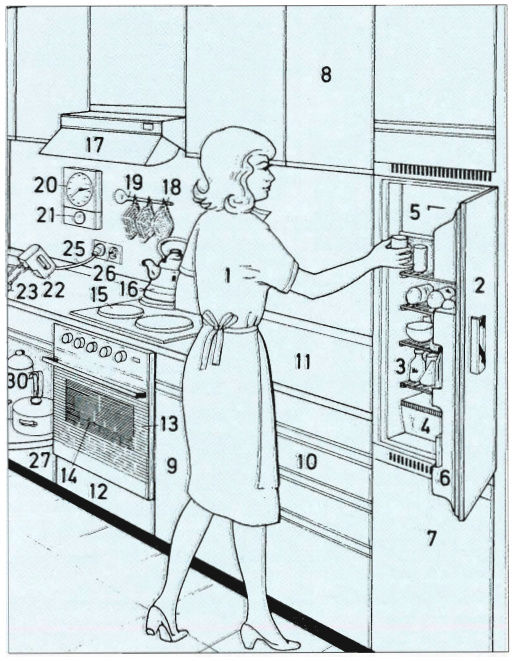
détail d'un portrait de Buffon par François Hubert Drouais coll. priv.

buffet à deux corps à décor de marqueterie France, XVIIe s. Hôtel Lallemand, Bourges

buffle

Bernard Buffet : le Filet (1948). Musée municipal d'Art moderne, Paris

Above: a page from the *Duden pictorial Encyclopedia in Five Languages* (1958 edition)



16 der Wasserkessel (Flötenkessel)
– *kettle (whistling kettle)*
17 der Wrasenabzug (Dunstabzug)
– *cooker hood*
18 der Topflappen
– *pot holder*
19 der Topflappenhalter
– *pot holder rack*

16 der Wasserkessel (Flötenkessel)
– *el hervidor (el hervidor de silbato* m; Amér.: *la pava)*
17 der Wrasenabzug (Dunstabzug)
– *la campana*
18 der Topflappen
– *la manopla (el agarrador)*
19 der Topflappenhalter
– *el cuelgamanoplas*

16 der Wasserkessel (Flötenkessel)
– *la bouilloire, la bouilloire à sifflet* m
17 der Wrasenabzug (Dunstabzug)
– *la hotte*
18 der Topflappen
– *la manique, le protège-main*
19 der Topflappenhalter
– *l'accroche-manique m, le crochet à protège-main*

16 der Wasserkessel (Flötenkessel)
– *il bollitore dell'acqua*
17 der Wrasenabzug (Dunstabzug)
– *la cappa aspirante*
18 der Topflappen
– *la presina da cucina*
19 der Topflappenhalter
– *il portapresine (l'appendipresine* m)

THE HISTORY OF LEXICOGRAPHY

The earliest dictionaries had a practical aim. They were often bilingual or polyglot word-lists aimed at the traveller and the missionary, or glossaries written to help people understand words which were dialectal, technical, or rare. From as early as the 5th century BC, the Greeks were compiling *glossai,* explaining difficult words in such authors as Homer. The first vocabulary lists in English were similar: these were 8th-century Anglo-Saxon glosses, in which English words were written between the Latin lines. Later, these glosses were collected together as lists. But random collections of words or glosses are not dictionaries: to count as a dictionary, the words need to be organized in a systematic way – such as through the use of the alphabetical principle.

The history of lexicography goes back over 2,000 years, to ancient China, Greece, and Rome. But there has never been a time when some kind of lexicographical work was not in progress. Some periods were particularly fruitful.

- Arabic dictionaries flourished from around the 8th century onwards.

Left: a page from the *Grand Larousse Universel* (1982)
It shows the mixture of lexical, literary, and encyclopedic information which is the hallmark of this tradition.

Below: a page from a German–Latin school book, *Teutsch lateinisches Wörter-Büchlein*
Printed around 1722, it contains about 6,000 words, each of which is illustrated.

- There was a flurry of activity in several languages following the invention of printing.
- The Accademia della Crusca produced its dictionary in 1612 (the first to be compiled by a team of people), and prompted several other national dictionary projects.
- Polyglot dictionaries were particularly numerous in the 17th century, with the development of trade and missionary activities around the world.
- The 18th century saw a fresh direction in lexicography, following the discoveries of the comparative philologists (§50), and the first major historical dictionaries began to be compiled.
- The 19th century saw many large-scale dictionary projects, produced by teams of compilers, and several specialized dictionaries (such as of dialect or technical words). Different kinds of dictionaries began to be produced, notably the Larousse series (from 1856), with its distinctive pictorial and encyclopedic character.
- The 20th century has seen the development of lexicography as a scholarly subject, largely under the influence of linguistics, and promoted especially by the growth of academic societies, such as the Dictionary Society of North America (1975), and the European Association for Lexicography (EURALEX, 1983)

TWENTY QUESTIONS TO ASK WHEN YOU BUY A DICTIONARY

1 Is the paper of good, hard-wearing quality?
2 Will the binding allow it to be opened flat?
3 Are (especially long) entries clearly laid out?
4 Does it have the words you most want to look up? (Keep a note of some words which have caused you problems, and use them as a quick check.)
5 Does it have good international coverage?
6 Does it contain encyclopedic information?
7 Does it have illustrations of difficult concepts?
8 Are the definitions clearly distinguished, and organized on a sensible principle?
9 Are the definitions easy to understand, and helpful (e.g. avoiding vicious circularity, as when X is defined as Y, and Y is then defined as X)?
10 Does it give citations (examples of usage), and are they real or artificial?
11 Does it give guidance about usage?
12 Does it use a good set of stylistic labels (e.g. formal, slang, medical, archaic)?
13 Does it give etymological information?
14 Does it give guidance about capitalization, spelling variation, and where syllable boundaries go (i.e. where to hyphenate)?
15 Does it give pronunciation variants, and is the phonetic transcription easy to follow?
16 Does it contain idioms, phrases, proverbs, etc.?
17 Does it contain lists of synonyms and antonyms?
18 Does it give useful cross-references to other words of related meaning?
19 Does it give information about word class, inflectional endings, and other relevant features of grammar?
20 Are there useful appendices (e.g. abbreviations, measures)?

THE PRESENT ... THE FUTURE ...

Since the 1970s, the flow of dictionaries has been unabated, as publishers try to meet the needs of an increasingly language-conscious age. In English, for example, new editions and supplements to the well-known dictionaries have appeared, and several publishers have launched new general series. Reader's Digest produced its *Great Illustrated Dictionary* in 1984, the first full-colour English dictionary, in the encyclopedic tradition of Larousse. Prominent also have been the dictionaries for special purposes (foreign language teaching, linguistics, medicine, chemistry, etc.). For the first time, spoken vocabulary has begun to find its way into dictionaries (though by no means all are yet willing to include the more colloquial words and uses).

But this outpouring is as nothing compared with the flood yet to come. The 1980s will one day be seen as a watershed in lexicography – the decade in which computer applications began to alter radically the methods and the potential of lexicography. Gone are the days of painstaking manual transcription and sorting on paper slips: the future is on disk, in the form of vast lexical databases, continuously updated, that can generate a dictionary of a given size and scope in a fraction of the time it used to take. Special programs are already available enabling people to ask the dictionary special questions (such as: 'find all words that entered the language in 1964' or 'find all words ending in *-esse*'). Access to large machine-dictionaries will become routine in offices and homes. One day, we shall not look up a word in a dictionary on a shelf but ask our home computer for the information we need. That day is not far off.

WORDS THAT HAVE NEVER BEEN USED – OUTSIDE THE DICTIONARY!

Early dictionaries were often little more than lists of 'hard' words. After all, it was reasoned, why should a dictionary include the words everyone already knew? The result was the inclusion of many very rare words – and some whose usage has never been officially recorded, outside of the dictionaries in which they appear! Examples are *commemorable* and *liquescency*.

But for non-words, the prize must go to *Dord*, first used in a 1930s dictionary, and found subsequently in several others. It seems there was a file in the compiler's office which contained the entry 'D or d' as abbreviations for the word 'density'. The information was accidentally put into the dictionary as *Dord*, meaning 'density', and a new word was born.

SOME IMPORTANT EVENTS BJ (BEFORE JOHNSON)

5th c. BC Protagoras of Abdera compiled a glossary of unfamiliar words in Homer.
3rd c. BC The poet Philetas of Cos compiled a glossary of unusual poetic, technical, and dialect words.
2nd c. BC Aristophanes of Byzantium compiled a dictionary of current and obsolete words.
1st c. BC Marcus Verrius Flaccus compiled the first Latin lexicon, *Libri de significatu verborum.*
1st c. Valerius Harpocration compiled a lexicon of the Attic orators.
2nd c. First systematic Chinese dictionary, *Hsuo Wên,* compiled by Hsü Shên.
5th c. Hesychius of Alexandria compiled a large lexicon of Classical Greek.
6th c. Compilation of a Sanskrit dictionary by the Hindu grammarian, Amarasiṃha.
8th c. The first general Arabic dictionary, *Kitāb al-'ayn,* compiled by Al-Khalīl Ibn Aḥmad.
10th–11th c. Compilation of a Byzantine encyclopedic dictionary, the *Suda.*
11th c. First Chinese–Japanese encyclopedic dictionary, by Minamoto no Shitagô.
12th c. Compilation of the Greek *Etymologicon magnum,* author unknown.
13th c. Johannes Balbus Januensis compiled the encyclopedic dictionary, the *Catholicon,* one of the most influential dictionaries of the middle ages, and the first to be printed (in 1460).
1477 The earliest printed bilingual dictionary: the *Vocabolista italiano-tedesco* (Venice).
1499 Probably the first dictionary to be printed in England: the Latin–English *Promptuorium parvulorum* (London, Richard Pynson).
1511 The first printed Dutch dictionary: Noël de Berlaimont's *Vocabulaire.*
1539 Compilation of Robert Estienne's *Dictionnaire françois-latin.*
1596 The first published Russian dictionary: Laurentii Zizanii's *Leksys ... synonima sloveno-rosskaia.*
1606 Publication of Jean Nicot's *Thresor de la langue francoyse,* the first systematic French dictionary.
1611 Publication of the first major Spanish dictionary, *Tesoro de la lengua castellana o española* of Covarrubias y Horozco.
(After R. L. Collison, 1982.)

20 · DISCOURSE AND TEXT

The traditional concern of linguistic analysis has been the construction of sentences (§16); but in recent years there has been an increasing interest in analysing the way sentences work in sequence to produce coherent stretches of language.

Two main approaches have developed. *Discourse analysis* focuses on the structure of naturally occurring spoken language, as found in such 'discourses' as conversations, interviews, commentaries, and speeches. *Text analysis* focuses on the structure of written language, as found in such 'texts' as essays, notices, road signs, and chapters. But this distinction is not clear-cut, and there have been many other uses of these labels. In particular, both 'discourse' and 'text' can be used in a much broader sense to include *all* language units with a definable communicative function, whether spoken or written. Some scholars talk about 'spoken and written discourse'; others about 'spoken and written text'. In Europe, the term *text linguistics* is often used for the study of the linguistic principles governing the structure of all forms of text.

The search for larger linguistic units and structures has been pursued by scholars from many disciplines. Linguists investigate the features of language that bind sentences when they are used in sequence. Ethnographers and sociologists study the structure of social interaction, especially as manifested in the way people enter into dialogue. Anthropologists analyse the structure of myths and folk-tales. Psychologists carry out experiments on the mental processes underlying comprehension. And further contributions have come from those concerned with artificial intelligence, rhetoric, philosophy, and style (§12).

These approaches have a common concern: they stress the need to see language as a dynamic, social, interactive phenomenon – whether between speaker and listener, or writer and reader. It is argued that meaning is conveyed not by single sentences but by more complex exchanges, in which the participants' beliefs and expectations, the knowledge they share about each other and about the world, and the situation in which they interact, play a crucial part.

CONVERSATION

Of the many types of communicative act, most study has been devoted to conversation, seen as the most fundamental and pervasive means of conducting human affairs (p. 52). These very characteristics, however, complicate any investigation. Because people interact linguistically in such a wide range of social situations, on such a variety of topics, and with such an unpredictable set of participants, it has proved very difficult to determine the extent to which conversational behaviour is systematic, and to generalize about it.

There is now no doubt that such a system exists. Conversation turns out, upon analysis, to be a highly structured activity, in which people tacitly operate with a set of basic conventions. A comparison has even been drawn with games such as chess: conversations, it seems, can be thought of as having an opening, a middle, and an end game. The participants make their moves and often seem to follow certain rules as the dialogue proceeds. But the analogy ends there. A successful conversation is not a game: it is no more than a mutually satisfying linguistic exchange. Few rules are ever stated explicitly (some exceptions are 'Don't interrupt!', and 'Look at me when I talk to you'). Furthermore, apart from in certain types of argument and debate, there are no winners.

Conversational success

For a conversation to be successful, in most social contexts, the participants need to feel they are contributing something to it and are getting something out of it. For this to happen, certain conditions must apply. Everyone must have an opportunity to speak: no one should be monopolizing or constantly interrupting. The participants need to make their roles clear, especially if there are several possibilities (e.g. 'Speaking as a mother / linguist / Catholic ...'). They need to have a sense of when to speak or stay silent; when to proffer information or hold it back; when to stay aloof or become involved. They need to develop a mutual tolerance, to allow for speaker unclarity and listener inattention: perfect expression and comprehension are rare, and the success of a dialogue largely depends on people recognizing their communicative weaknesses, through the use of rephrasing (e.g. 'Let me put that another way') and clarification (e.g. 'Are you with me?').

There is a great deal of ritual in conversation, especially at the beginning and end, and when topics change. For example, people cannot simply leave a conversation at any random point, unless they wish to be considered socially inept or ill-mannered. They have to choose their point of departure (such as the moment when a topic changes) or construct a special reason for leaving. Routines for concluding a conversation are particularly complex, and cooperation is crucial if it is not to end abruptly, or in an embarrassed silence. The parties may prepare for their departure a

CONVERSATION ANALYSIS

In recent years, the phrase 'conversation analysis' has come to be used as the name of a particular method of studying conversational structure, based on the techniques of the American sociological movement of the 1970s known as *ethnomethodology*.

The emphasis in previous sociological research had been deductive and quantitative, focusing on general questions of social structure. The new name was chosen to reflect a fresh direction of study, which would focus on the techniques (or 'methods') used by people themselves (oddly referred to as 'ethnic'), when they are actually engaged in social – and thus linguistic — interaction. The central concern was to determine how individuals experience, make sense of, and report their interactions.

In conversation analysis, the data thus consist of tape recordings of natural conversation, and their associated transcriptions. These are then systematically analysed to determine what properties govern the way in which a conversation proceeds. The approach emphasizes the need for empirical, inductive work, and in this it is sometimes contrasted with 'discourse analysis', which has often been more concerned with formal methods of analysis (such as the nature of the rules governing the structure of texts).

long way in advance, such as by looking at their watches or giving a verbal early warning. A widespread convention is for visitors to say they must leave some time before they actually intend to depart, and for the hosts to ignore the remark. The second mention then permits both parties to act.

The topic of the conversation is also an important variable. In general it should be one with which everyone feels at ease: 'safe' topics between strangers in English situations usually include the weather, pets, children, and the local context (e.g. while waiting in a room or queue); 'unsafe' topics include religious and political beliefs and problems of health. There are some arbitrary divisions: asking what someone does for a living is generally safe: asking how much they earn is not. Cultural variations can cause problems: commenting about the cost of the furniture or the taste of a meal may be acceptable in one society but not in another.

It is difficult to generalize about what is normal, polite, or antisocial in conversational practice, as there is so much cultural variation. Silence, for example, varies in status. It is an embarrassment in English conversations, unless there are special reasons (such as in moments of grief). However, in some cultures (e.g. Lapps, Danes, the Western Apache) it is quite normal for participants to become silent. Often, who speaks, and how much is spoken, depends on the social status of the participants – for example, those of lower rank may be expected to stay silent if their seniors wish to speak (p. 38). Even the basic convention of 'one person speaks at a time' may be broken. In Antigua, for example, the phenomenon of several people speaking at once during a whole conversation is a perfectly normal occurrence.

CONVERSATIONAL MAXIMS

The success of a conversation depends not only on what speakers say but on their whole approach to the interaction. People adopt a 'cooperative principle' when they communicate: they try to get along with each other by following certain conversational 'maxims' that underlie the efficient use of language. Four basic maxims have been proposed (after H. P. Grice, 1975):

- The *maxim of quality* states that speakers' contributions to a conversation ought to be true. They should not say what they believe to be false, nor should they say anything for which they lack adequate evidence.
- The *maxim of quantity* states that the contribution should be as informative as is required for the purposes of the conversation. One should say neither too little nor too much.
- The *maxim of relevance* states that contributions should clearly relate to the purpose of the exchange.
- The *maxim of manner* states that the contribution should be perspicuous – in particular, that it should be orderly and brief, avoiding obscurity and ambiguity.

Other maxims have also been proposed, such as 'Be polite', 'Behave consistently'. The principle of relevance has recently attracted most attention, as it has been proposed as a fundamental explanatory principle for a theory of human communication (D. Sperber & D. Wilson, 1986).

Listeners will normally assume that speakers are following these criteria. Speakers may of course break (or 'flout') these maxims – for example, they may lie, be sarcastic, try to be different, or clever – but conversation proceeds on the assumption that they are not doing so. Listeners may then draw inferences from what speakers *have* said (the literal meaning of the utterance) concerning what they have *not* said (the implications, or 'implicatures' of the utterance). For example,

A: I need a drink. B: Try The Bell.

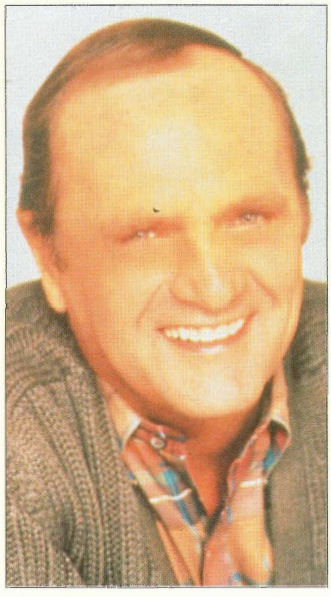
If B is adhering to the cooperative principle, several implicatures arise out of this dialogue: for example, The Bell must be a place that sells drinks; it must be open (as far as B knows); it must be nearby. If B is not being cooperative (e.g. if he knows that The Bell is closed, or is the name of a greengrocer's), he is flouting the maxims of quality and relevance.

Deliberate flouting of this kind is uncommon, of course, and only occurs in such special cases as sarcasm, joking, or deliberate unpleasantness. More likely is the inadvertent flouting of conversational maxims – as would happen if B genuinely did not know that The Bell was closed, and accidentally sent A on a wild goose chase. In everyday conversation, misunderstandings often take place as speakers make assumptions about what their listeners know, or need to know, that turn out to be wrong. At such points, the conversation can break down and may need to be 'repaired', with the participants questioning, clarifying, and cross-checking. The repairs are quickly made in the following extract, through the use of such pointers as 'I told you' and 'sorry'.

A: Got the time? B: No, I told you, I lost my watch. A: Oh, sorry, I forgot.

But it is quite common for participants not to realize that there has been a breakdown, and to continue conversing at cross purposes.

Bob Newhart
Newhart's comedy routines often rely on the audience's awareness of discourse conventions. His 'driving instructor' sketch, for example, gives us only half of the conversation, from the instructor's viewpoint, leaving the responses of the learner driver to our imagination. Joyce Grenfell's 'teaching young children' sketches were based on the same principle.



GEORGE – DON'T DO THAT

This extract from one of Joyce Grenfell's nursery school monologues shows how the reader can survive using just one side of a dialogue. The task is made easier here by the fact that it is a standard teaching technique to reinforce what a young child has just said by repeating or expanding it (as do parents: see p. 233).

Now then, let's all put on our Thinking Caps, shall we, and think what flowers we are going to choose to be.
Lavinia? – What flower are you?
A bluebell. Good.
Peggy?
A red rose. That's nice.
Neville?
A *wild* rose. Well done, Neville!
Sidney? – Sidney, pay attention, dear, and don't pummel Rosemary – what flower are you going to choose to be?
A *horse* isn't a flower, Sidney.
(From J. Grenfell, 1977, p. 30.)

Joyce Grenfell (1910–79)

CONVERSATIONAL TURNS

Probably the most widely recognized conversational convention is that people take turns to speak. But how do people know when it is their turn? Some rules must be present, otherwise conversations would be continually breaking down into a disorganized jumble of interruptions and simultaneous talk. In many formal situations, such as committee meetings and debates, there are often explicit markers showing that a speaker is about to yield the floor, and indicating who should speak next ('I think Mr Smith will know the answer to that question'). This can happen in informal situations too ('What do you think, Mary?'), but there the turn-taking cues are usually more subtle.

People do not simply stop talking when they are ready to yield the floor. They usually signal some way in advance that they are about to conclude. The clues may be semantic ('So anyway, ...', 'Last but not least ...'); but more commonly the speech itself can be modified to show that a turn is about to end – typically, by lowering its pitch, loudness, or speed. Body movements and patterns of eye contact are especially important. While speaking, we look at and away from our listener in about equal proportions; but as we approach the end of a turn, we look at the listener more steadily. Similarly, when talking to a group of people, we often look more steadily at a particular person, to indicate that in our view this should be the next speaker.

Listeners are not passive in all of this. Here too there are several ways of signalling that someone wants to talk next. Most obviously, the first person in a group actually to start speaking, after the completion of a turn, will usually be allowed to hold the floor. More subtly, we can signal that we want to speak next by an observable increase in body tension – by leaning forward, or producing an audible intake of breath. Less subtly, we can simply interrupt – a strategy which may be tolerated, if the purpose is to clarify what the speaker is saying, but which more usually leads to social sanctions.

EXCHANGES

Because conversational discourse varies so much in length and complexity, analysis generally begins by breaking an interaction down into the smallest possible units, then examining the way these units are used in sequences. The units have been called 'exchanges' or 'interchanges', and in their minimal form consist simply of an initiating utterance (I) followed by a response utterance (R), as in:

I: What's the time?
R: Two o'clock.

Two-part exchanges (sometimes called 'adjacency pairs') are commonplace, being used in such contexts as questioning / answering, informing / acknowledging, and complaining / excusing. Three-part exchanges are also important, where the response is followed by an element of feedback (F). Such reactions are especially found in teaching situations:

TEACHER: Where were the arrows kept? (I)
PUPIL: In a special kind of box. (R)
TEACHER: Yes, that's right, in a box. (F)

What is of particular interest is to work out the constraints that apply to sequences of this kind. The teacher–feedback sequence would be inappropriate in many everyday situations:

A: Did you have a good journey?
B: Apart from a jam at Northampton.
A: *Yes, that's right, a jam at Northampton.

Unacceptable sequences are easy to invent:

A: Where do you keep the jam?
B: *It's raining again.

On the other hand, with ingenuity it is often possible to imagine situations where such a sequence could occur (e.g. if B were staring out of the window at the time). And discourse analysts are always on the lookout for unexpected, but perfectly acceptable, sequences in context, such as:

A: Goodbye.
B: Hello.

(used, for example, as A is leaving an office, passing B on the way in). Many jokes, too, break discourse rules as the source of their effect:

A: Yes, I can.
B: Can you see into the future?

MISUNDERSTANDINGS

An important aim of discourse analysis is to find out why conversations are not always successful. Misunderstanding and mutual recrimination is unfortunately fairly common. Participants often operate with different rules and expectations about the way in which the conversation should proceed – something that is particularly evident when people of different cultural backgrounds interact. But even within a culture, different 'rules of interpretation' may exist.

It has been suggested, for example, that there are different rules governing the way in which men and women participate in a conversation (pp. 21, 120). A common source of misunderstanding is the way both parties use head nods and *mhm* noises while the other is speaking – something that women do much more frequently than men. Some analysts have suggested that the two sexes mean different things by this behaviour. When a woman does it, she is simply indicating that she is listening, and encouraging the speaker to continue, but the male interprets it to mean that she is agreeing with everything he is saying. By contrast, when a man does it, he is signalling that he does not necessarily agree, whereas the woman interprets it to mean that he is not always listening. Such interpretations are plausible, it is argued, because they explain two of the most widely reported reactions from participants in cross-sex conversations – the male reaction of 'It's impossible to say what a woman really thinks', and the female reaction of 'You never listen to a word I say.' (After D. N. Maltz & R. A. Borker, 1982.)

CONVERSATION MANOEUVRES

Conversational turn-taking is often marked by clear signals of direction

Openings
Guess what ...
Sorry to trouble you ...
Lovely day!
Got a match?
Can I help you?
Good morning.
Excuse me ...
Did you hear the one about ...
Can you spare a minute?
Halt! Who goes there?
But not: *How much do you earn?

Ongoing checks
By the speaker:
Do you see?
Can you guess what he said?
Are you with me?
Do I make myself clear?
Don't you think?
Let me put it another way ...
Don't get me wrong ...
What I'm trying to say is ...
By the listener:
You mean ...
Have I got you right?
Mhm.
I don't get you.
Let's get that straight ...

Changing topic
Introducing a new topic:
That reminds me ...
Incidentally ...
That's a good question.
By the way ...
Speaking of John ...
Where was I?
Concluding a topic:
So it goes.
That's life.
Makes you think, doesn't it.
Let's wait and see.

Ending
Sorry, but I have to go now.
Nice talking to you.
Well, must get back to work.
Gosh, is that the time?
I mustn't keep you.
Gotta run. (*especially US*)

TEXTUAL STRUCTURE

To call a sequence of sentences a 'text' is to imply that the sentences display some kind of mutual dependence; they are not occurring at random. Sometimes the internal structure of a text is immediately apparent, as in the headings of a restaurant menu; sometimes it has to be carefully demonstrated, as in the network of relationships that enter into a literary work. In all cases, the task of textual analysis is to identify the linguistic features that cause the sentence sequence to 'cohere' – something that happens whenever the interpretation of one feature is dependent upon another elsewhere in the sequence. The ties that bind a text together are often referred to under the heading of *cohesion* (after M. A. K. Halliday & R. Hasan, 1976). Several types of cohesive factor have been recognized:

• *Conjunctive relations* What is about to be said is explicitly related to what has been said before, through such notions as contrast, result, and time:

I left early. *However,* Jean stayed till the end.

Lastly, there's the question of cost.

• *Coreference* Features that cannot be semantically interpreted without referring to some other feature in the text. Two types of relationship are recognized: *anaphoric* relations look backwards for their interpretation, and *cataphoric* relations look forwards:

Several people approached. They seemed angry.

Listen to this: *John's getting married.*

• *Substitution* One feature replaces a previous expression:

I've got a pencil. Do you have *one*?

Will we get there on time? I think *so.*

• *Ellipsis* A piece of structure is omitted, and can be recovered only from the preceding discourse:

Where *did you see the car*? ∧ In the street.

• *Repeated forms* An expression is repeated in whole or in part:

Canon Brown arrived. Canon Brown was cross.

• *Lexical relationships* One lexical item enters into a structural relationship with another (p. 105):

The *flowers* were lovely. She liked the *tulips* best.

• *Comparison* A compared expression is presupposed in the previous discourse:

That house was *bad.* This one's far *worse.*

Cohesive links go a long way towards explaining how the sentences of a text hang together, but they do not tell the whole story. It is possible to invent a sentence sequence that is highly cohesive but nonetheless incoherent (after N. E. Enkvist, 1978, p. 110):

> A week has seven *days.* Every *day* I feed my *cat. Cats* have four legs. *The cat* is on the *mat. Mat* has three letters.

A text plainly has to be *coherent* as well as cohesive, in that the concepts and relationships expressed should be relevant to each other, thus enabling us to make plausible inferences about the underlying meaning.

TWO WAYS OF DEMONSTRATING COHESION

Paragraphs are often highly cohesive entities. The cohesive ties can stand out very clearly if the sentences are shuffled into a random order. It may even be possible to reconstitute the original sequence solely by considering the nature of these ties, as in the following case:

1. However, nobody had seen one for months.
2. He thought he saw a shape in the bushes.
3. Mary had told him about the foxes.
4. John looked out of the window.
5. Could it be a fox?

(The original sequence was 4,2,5,3,1.)

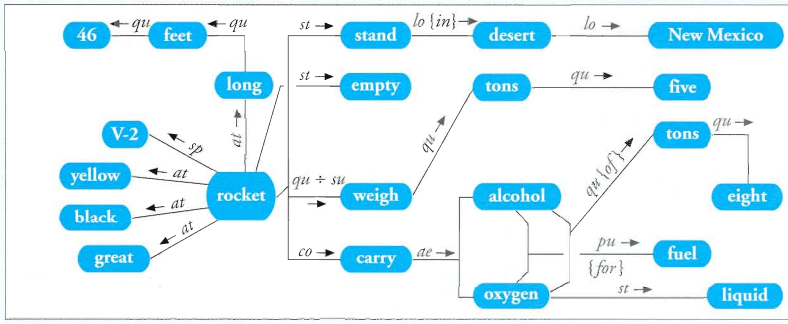
We can use graphological devices to indicate the patterns of cohesion within a text. Here is the closing paragraph of James Joyce's short story 'A Painful Case'. The sequence of pronouns, the anaphoric definite articles, and the repeated phrases are the main cohesive features between the clauses and sentences. Several of course refer back to previous parts of the story, thus making this paragraph, out of context, impossible to understand.

> He turned back the way he had come, the **rhythm** of the engine pounding in his ears. He began to doubt the reality of what memory told him. He halted under a tree and allowed **the rhythm** to die away. He could not feel her near him in the DARKNESS nor her *voice* touch his ear. He waited for some minutes *listening.* He could *hear* ***NOTHING***: the NIGHT was ***perfectly silent.*** He *listened* again: ***perfectly silent.*** He felt that he was ***ALONE.***

MACROSTRUCTURES

Not all textual analysis starts with small units and works from the 'bottom up' (p. 71); some approaches aim to make very general statements about the macrostructure of a text. In psycholinguistics, for example, attempts have been made to analyse narratives into schematic outlines that represent the elements in a story that readers remember. These schemata have been called 'story-grammars' (though this is an unusually broad sense of the term 'grammar', cf §16).

In one such approach (after P. W. Thorndyke, 1977), simple narratives are analysed into four components: *setting, theme, plot,* and *resolution.* The setting has three components: the *characters,* a *location,* and a *time.* The theme consists of an *event* and a *goal.* The plot consists of various *episodes,* each with its own *goal* and *outcome.* Using distinctions of this kind, simple stories are analysed into these components, to see whether the same kinds of structure can be found in each (p. 79). Certain similarities do quickly emerge; but when complex narratives are studied, it proves difficult to devise more detailed categories that are capable of generalization, and analysis becomes increasingly arbitrary.



Conceptual structure One way of representing the conceptual structure of a text (after R. de Beaugrande & W. Dressler, 1981, p. 100). This 'transition network' summarizes the following paragraph:

A great black and yellow V-2 rocket 46 feet long stood in a New Mexico desert. Empty, it weighed five tons. For fuel it carried eight tons of alcohol and liquid oxygen.

The abbreviations identify the types of semantic links which relate the concepts (following the direction of the arrows):

ae	affected entity
at	attribute of
co	containment of
lo	location of
pu	purpose of
qu	quantity of
sp	specification of
st	state of
su	substance of

21 · PRAGMATICS

Pragmatics studies the factors that govern our choice of language in social interaction and the effects of our choice on others. In theory, we can say anything we like. In practice, we follow a large number of social rules (most of them unconsciously) that constrain the way we speak. There is no law that says we must not tell jokes during a funeral, but it is generally 'not done'. Less obviously, there are norms of formality and politeness that we have intuitively assimilated, and that we follow when talking to people who are older, of the opposite sex, and so on. Writing and signing behaviour are constrained in similar ways.

Pragmatic factors always influence our selection of sounds, grammatical constructions, and vocabulary from the resources of the language. Some of the constraints are taught to us at a very early age – in British English, for example, the importance of saying *please* and *thank you*, or (in some families) of not referring to an adult female in her presence as *she* (p. 248). In many languages, pragmatic distinctions of formality, politeness, and intimacy are spread throughout the grammatical, lexical, and phonological systems, ultimately reflecting matters of social class, status, and role (§10, p. 99). A well-studied example is the pronoun system, which frequently presents distinctions that convey pragmatic force – such as the choice between *tu* and *vous* in French.

Languages differ greatly in these respects. Politeness expressions, for instance, may vary in frequency and meaning. Many European languages do not use their word for *please* as frequently as English does; and the function and force of *thank you* may also alter (e.g. following the question 'Would you like some more cake?', English *thank you* means 'yes', whereas French *merci* would mean 'no'). Conventions of greeting, leave-taking, and dining also differ greatly from language to language. In some countries it is polite to remark to a host that we are enjoying the food; in others it is polite to stay silent. On one occasion, at a dinner in an Arabic community, the present author made the mistake of remarking on the excellence of the food before him. The host immediately apologized, and arranged for what was there to be replaced!

Pragmatic errors break no rules of phonology, syntax, or semantics. The elements of *How's tricks, your majesty?* will all be found in English language textbooks and dictionaries, but for most of us the sequence is not permissible from a pragmatic viewpoint. Pragmatics has therefore to be seen as separate from the 'levels' of language represented in linguistic models of analysis (§13). It is not a 'part' of language structure, but its domain is so closely bound up with structural matters that it cannot be ignored in this section of the encyclopedia.

THE IDENTITY OF PRAGMATICS

Pragmatics is not at present a coherent field of study. A large number of factors govern our choice of language in social interaction, and it is not yet clear what they all are, how they are best interrelated, and how best to distinguish them from other recognized areas of linguistic enquiry. There are several main areas of overlap.

Semantics (§17) Pragmatics and semantics both take into account such notions as the intentions of the speaker, the effects of an utterance on listeners, the implications that follow from expressing something in a certain way, and the knowledge, beliefs, and presuppositions about the world upon which speakers and listeners rely when they interact.

Stylistics (§12) *and sociolinguistics* (§§10, 63) These fields overlap with pragmatics in their study of the social relationships which exist between participants, and of the way extralinguistic setting, activity, and subject-matter can constrain the choice of linguistic features and varieties.

Psycholinguistics (§§7, 38) Pragmatics and psycholinguistics both investigate the psychological states and abilities of the participants that will have a major effect upon their performance – such factors as attention, memory, and personality.

Discourse analysis (§20) Both discourse analysis and pragmatics are centrally concerned with the analysis of conversation, and share several of the philosophical and linguistic notions that have been developed to handle this topic (such as the way information is distributed within a sentence, deictic forms (p. 106), or the notion of conversational 'maxims' (p. 117)).

As a result of these overlapping areas of interest, several conflicting definitions of the scope of pragmatics have arisen. One approach focuses on the factors formally encoded in the structure of a language (honorific forms, *tu / vous* choice, and so on). Another relates it to a particular view of semantics: here, pragmatics is seen as the study of all aspects of meaning other than those involved in the analysis of sentences in terms of truth conditions (p. 107). Other approaches adopt a much broader perspective. The broadest sees pragmatics as the study of the principles and practice underlying *all* interactive linguistic performance – this including all aspects of language usage, understanding, and appro-

UNDERSTANDING MISUNDERSTANDING

The 1990s has seen the growth of a domain which can perhaps best be labelled 'applied pragmatics' – the use of a pragmatic perspective to analyse situations in which a conversation has not been successful, and to suggest solutions (p. 118). The general interest of this approach has been well illustrated by the success of Deborah Tannen's *That's Not What I Meant!* (1986) and *You Just Don't Understand* (1990), which focus on the different strategies and expectations people use when they try to talk to each other. There are a surprising number of everyday notions which can be illuminated by this kind of analysis, such as 'nagging', 'accusing', and 'being at cross-purposes'.

Here is one of Tannen's anecdotes and part of her associated commentary:

Loraine frequently compliments Sidney and thanks him for doing things such as cleaning up the kitchen and doing the laundry. Instead of appreciating the praise, Sidney resents it. 'It makes me feel like you're demanding that I do it all the time', he explains. ...

'In all these examples, men complained that their independence and freedom were being encroached on. Their early warning system is geared to detect signs that they are being told what to do ... Such comments surprise and puzzle women, whose early warning systems are geared to detect a different menace. ... If a man struggles to be strong, a woman struggles to keep the community strong.'

Applied pragmatics is not limited to family arguments. The same issues arise in the attempt to achieve successful communication in any setting at any level. A course in problems of business communication, advertising itself with the slogan 'Are you getting through to your customer?' is, in effect, an exercise in applied pragmatics.

priateness. Textbooks on pragmatics to date, accordingly, present a diversity of subject matter, and a range of partially conflicting orientations and methodologies, which proponents of the subject have yet to resolve. However, if we take diversity of opinion to be a sign of healthy growth in a subject, it must be said that few other areas of language study have such a promising future.

SPEECH ACTS

The British philosopher J. L. Austin (1911–60) was the first to draw attention to the many functions performed by utterances as part of interpersonal communication. In particular, he pointed out that many utterances do not communicate information, but are equivalent to actions. When someone says 'I apologize ...', 'I promise ...', 'I will' (at a wedding), or 'I name this ship ...', the utterance immediately conveys a new psychological or social reality. An apology takes place when someone apologizes, and not before. A ship is named only when the act of naming is complete. In such cases, to say is to perform. Austin thus called these utterances *performatives*, seeing them as very different from statements that convey information (*constatives*). In particular, performatives are not true or false. If A says 'I name this ship ...', B cannot then say 'That's not true'!

In speech act analysis, we study the effect of utterances on the behaviour of speaker and hearer, using a threefold distinction. First, we recognize the bare fact that a communicative act takes place: the *locutionary* act. Secondly, we look at the act that is performed as a result of the speaker making an utterance – the cases where 'saying = doing', such as betting, promising, welcoming, and warning: these, known as *illocutionary* acts, are the core of any theory of speech acts. Thirdly, we look at the particular effect the speaker's utterance has on the listener, who may feel amused, persuaded, warned, etc., as a consequence: the bringing about of such effects is known as a *perlocutionary* act. It is important to appreciate that the illocutionary force of an utterance and its perlocutionary effect may not coincide. If I warn you against a particular course of action, you may or may not heed my warning.

There are thousands of possible illocutionary acts, and several attempts have been made to classify them into a small number of types. Such classifications are difficult, because verb meanings are often not easy to distinguish, and speakers' intentions are not always clear. One influential approach sets up five basic types (after J. R. Searle, 1976):

- *Representatives* The speaker is committed, in varying degrees, to the truth of a proposition, e.g. *affirm, believe, conclude, deny, report.*
- *Directives* The speaker tries to get the hearer to do something, e.g. *ask, challenge, command, insist, request.*
- *Commissives* The speaker is committed, in varying degrees, to a certain course of action, e.g. *guarantee, pledge, promise, swear, vow.*
- *Expressives* The speaker expresses an attitude about a state of affairs, e.g. *apologize, deplore, congratulate, thank, welcome.*
- *Declarations* The speaker alters the external status or condition of an object or situation solely by making the utterance, e.g. *I resign, I baptize, You're fired, War is hereby declared.*

FELICITY CONDITIONS

Speech acts are successful only if they satisfy several criteria, known as 'felicity conditions'. For example, the 'preparatory' conditions have to be right: the person performing the speech act has to have the authority to do so. This is hardly an issue with such verbs as *apologize, promise,* or *thank,* but it is important constraint on the use of such verbs as *fine, baptize, arrest,* and *declare war,* where only certain people are qualified to use these utterances. Then, the speech act has to be executed in the correct manner: in certain cases there is a procedure to be followed exactly and completely (e.g. *baptizing*); in others, certain expectations have to be met (e.g. one can only *welcome* with a pleasant demeanour). And, as a third example, 'sincerity' conditions have to be present: the speech act must be performed in a sincere manner. Verbs such as *apologize, guarantee,* and *vow* are effective only if speakers mean what they say; *believe* and *affirm* are valid only if the speakers are not lying.

Ordinary people automatically accept these conditions when they communicate, and they depart from them only for very special reasons. For example, the request *Will you shut the door?* is appropriate only if (a) the door is open, (b) the speaker has a reason for asking, and (c) the hearer is in a position to perform the action. If any of these conditions does not obtain, then a special interpretation of the speech act has to apply. It may be intended as a joke, or as a piece of sarcasm. Alternatively, of course, there may be doubt about the speaker's visual acuity, or even sanity!

Each part of this notice conveys the directive illocutionary force intended by the writer. The perlocutionary effect, however, is not as anticipated!

INDIRECT SPEECH ACTS

Some speech acts directly address a listener, but the majority of acts in everyday conversation are indirect. For example, there are a very large number of ways of asking someone to perform an action. The most direct way is to use the imperative construction (*Shut the door*), but it is easy to sense that this would be inappropriate in many everyday situations – too abrupt or rude, perhaps. Alternatives stress such factors as the hearer's ability or desire to perform the action, or the speaker's reasons for having the action done. These include the following:

I'd be grateful if you'd shut the door.

Could you shut the door?

Would you mind shutting the door?

It'd help to have the door shut.

It's getting cold in here.

Shall we keep out the draught?

Now, Jane, what have you forgotten to do?

Brrr!

Any of these could, in the right situation, function as a request for action, despite the fact that none has the clear form of an imperative. But of course, it is always open to the hearer to misunderstand an indirect request – either accidentally or deliberately.

Teacher: Johnny, there's some chalk on the floor.
Johnny: Yes, there is, sir.
Teacher: Well, pick it up, then!

47 · HOW MANY LANGUAGES?

There is no agreed total for the number of languages spoken in the world today. Most reference books give a figure of 5,000 to 6,000, but estimates have varied from 3,000 to 10,000. To see why there is such uncertainty, we need to consider the many problems facing those who wish to obtain accurate information, and also the reasons (linguistic, historical, and cultural) which preclude a simple answer to the question 'What counts as a language?'

DISCOVERIES

An obvious reason for the uncertainty over numbers is that even today new peoples, and therefore languages, continue to be discovered in the unexplored regions of the world – especially in the Amazon basin (as the Transamazonica road system is extended), Central Africa, and New Guinea. However, only a few languages are likely to be encountered in this way; and it is much more usual to find parts of the world where the people are known, but the languages spoken in their area are not. There are in fact many countries where linguistic surveys are incomplete or have not even begun. It is often assumed that the people speak one of the known languages in their area; or that they speak a dialect of one of these languages; but upon investigation their speech is found to be so different that it has to be recognized as a separate language.

ALIVE OR DEAD?

Against this steady increase in the world language total, there is a major factor which decreases it. For a language to count as 'living', there obviously have to be native speakers alive who use it. But in many parts of the world, it is by no means an easy matter to determine whether native speakers are still living – or, if they are, whether they still use their mother tongue regularly.

The speed with which a language can die in the smaller communities of the world is truly remarkable. The Amazonian explorations led to the discovery of many new languages, but they also led to their rapid death, as the Indians became swallowed up by the dominant western culture. Within a generation, all traces of a language can disappear. Political decisions force tribes to move or be split up. Economic prospects attract younger members away from the villages. New diseases take their toll. In 1962, Trumai, spoken in a single village on the lower Culuene River in Brazil, was reduced by an influenza epidemic to a population of fewer than 10 speakers. In the 19th century, there were thought to be over 1,000 Indian languages in Brazil; today, there are only 200. A quarter of the world's languages have fewer than 1,000 speakers; half have fewer than 10,000. It is likely that most of these languages will die out in the next 50 years.

LANGUAGE – OR DIALECT?

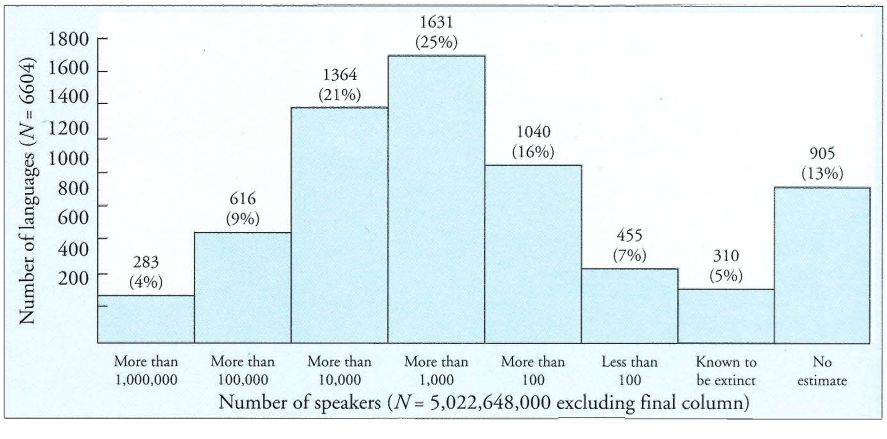
For most languages, the distinction between language and dialect is fairly clear-cut (p. 25). In the case of English, for example, even though regional vocabulary and local differences of pronunciation can make communication difficult at times, no-one disputes the existence of an underlying linguistic unity that all speakers identify as English, and which is confirmed by the use of a standard written language and a common literary heritage. But in hundreds of cases, considerations of this kind are in conflict with each other, or do not clearly apply.

The best-known conflicts occur when the criteria of national identity and mutual intelligibility do not coincide. The most common situation is one where two spoken varieties are mutually intelligible, but for political and historical reasons, they are referred to as different languages. For example, using just the intelligibility criterion, there are really only two Scandinavian languages: Continental (Swedish, Danish, and two standard varieties of Norwegian) and Insular (Icelandic, Faeroese). Swedes, Danes, and Norwegians can understand each other's speech, to a greater or lesser extent. But as soon as non-linguistic criteria are taken into account, we have to recognize at least five languages. To be Norwegian is to speak Norwegian; to be Danish is to speak Danish; and so on. In such cases, political and linguistic identity merge. And there are many other similar cases where political, ethnic, religious, literary, or other identities force a division where linguistically there is relatively little difference – Hindi vs Urdu, Bengali vs Assamese, Serbian vs Croatian, Twi vs Fante, Xhosa vs Zulu.

A new road cuts a swathe through the Brazilian rainforest.

NUMBER OF SPEAKERS

Number of speakers of the world's languages, based on the data provided in *the International Encyclopedia of Linguistics* (Bright, 1992). The total number of languages (including extinct ones) is 6,604. Most of the estimates were made during the 1980s, with some from the late 1970s. The world population total passed 5,000 million in July 1986, and had reached 5,111 million by mid-1988, which gives an indication of the order of magnitude unaccounted for in the final column.



The opposite situation is also quite common. Here we find cases where spoken varieties are mutually *unintelligible*, but for political, historical, or cultural reasons they are nonetheless called varieties of the same language. The three main 'dialects' of Lapp fall into this category, for example. Chinese is a case where linguistic criteria alone are in conflict with each other. From the viewpoint of the spoken language, the many hundreds of dialects in China can be grouped into eight main types (p. 314), which are mutually unintelligible to various degrees. But speakers of all these dialects share the same written language tradition, and those who have learned the system of Chinese characters are able to communicate with each other. Despite the linguistic differences, therefore, Chinese is considered by its speakers to be a single language.

In the above cases, the languages in question have been well studied, and many speakers are involved. When languages have been little studied, or have very few speakers, it is much more difficult for linguists to interpret all the factors correctly. For example, when two languages are in close proximity, they often borrow words from each other – sometimes even sounds and grammar. On first acquaintance, therefore, the languages may seem more alike than they really are, and analysts may believe them to be dialects of the same language. This has proved to be a real problem in such parts of the world as South America, Africa, and South-east Asia, where whole groups of languages may be affected in this way. Similarly, decisions about how to analyse all cases of dialect continua (p. 25) will affect our final total of languages.

LANGUAGE NAMES

A big problem, in working on lesser-known language areas, is deciding what credence to give to a language name. This issue does not arise when discussing the main languages of the world, which are usually known by a single name that translates neatly into other languages – as in the case of *Deutsch, German, Tedesco, Nemetskiy,* and *Allemand*, for instance. But in many cases the situation is not so straightforward.

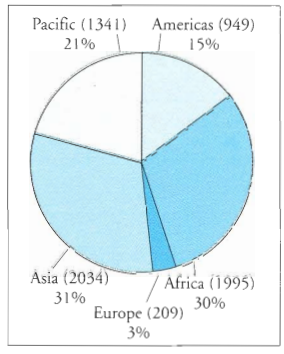
At one extreme, many communities have no specific name for their language. The name they use is the same as a common word or phrase in the language, such as the word for 'our language' or 'our people'. This is often so in Africa (where the name *Bantu*, which is given to a whole family of languages, means simply 'people'), and also in Meso- and South America. In the latter areas, we find such examples as *Carib* = 'people', *Tapuya* = 'enemy', and *Macu* = 'forest tribes'. Some tribes were called *chichimecatl* (= 'lineage of dogs'), *chontalli* (= 'foreigners') or *popoloca* (= 'barbarians'), and these labels led to the modern language names Chichimeca, Chontal, and Popoloca. Frequently, the name is the same as a river on which a tribe has been observed to live, as with the many groups of Land Dayak, in the West Indonesian family. In several Australian aboriginal languages, the name for the language is the word for 'this': for example, the nine languages within the Yuulngu family are known as *Dhuwala, Dhuwal, Dhiyakuy, Dhangu, Dhay'yi, Djangu, Djinang, Djining,* and *Nhangu.* Asking native speakers what language they speak is of little practical help, in such circumstances, if they only answer 'this'!

At the other extreme, it is quite common to find a community whose language has too many names. A South American Indian tribe, for instance, may have several names. A tribe, first of all, will have a name for itself (see above). But adjacent tribes may give the people a different name (e.g. *Puelche* means 'people from the east' in Araucanian). The Spanish or Portuguese explorers may have given them a third name – perhaps a characteristic of their appearance (e.g. *Coroado* means 'crowned' in Portuguese). More recently, anthropologists and other investigators may have used another name, often based on the geographical location of the tribe (e.g. 'up-river' vs 'down-river'). And lastly, the same language may be spelled differently in Spanish, Portuguese, English, or in its own writing system (if one has been devised). For example, Machacali, spoken in Minas Gerais, Brazil, is sometimes spelled Maxakali, sometimes Maxakari. When the initial letters vary (as when the Peruvian language Candoshi is spelled Kandoshi), indexing is especially awkward.

There are further complications. Sometimes, the same name is applied to two different languages, as when *mexicano* is used in Mexico to refer to Spanish (otherwise known as *español* or *castellano*) and to the main Indian language (*nahuatl*). Sometimes, speakers from different backgrounds may disagree about whether their ways of speaking should be related at all. Speakers of Luri, spoken in south-west Iran, say that their speech is a dialect of Persian; speakers of Persian disagree. Asking the native speakers is evidently no solution, for their perceptions will be governed by non-linguistic considerations, especially of a religious, nationalistic, or socioeconomic kind.

TO CONCLUDE

When all these factors are taken into account, it is plain that there will be no single answer to the question 'How many languages?' In some parts of the world, there has been a tendency to over-estimate, by taking names too literally and not grouping dialects together sufficiently – the Malayo-Polynesian languages are often cited in this connection. In other places, the totals are likely to have been underestimated – Indonesian languages, for example. There are over 37,000 language names listed in the 12th edition of *Ethnologue* (Grimes, 1992), and these have been grouped into 6,528 living languages. The number listed in the Index to the *Atlas of the World's Languages* (Moseley & Asher, 1994) is 6,796. The *International Encyclopedia of Linguistics* (Bright, 1992) lists 6,604, but this includes some 300 extinct languages. These surveys generally use data from the 1970s and 1980s. A total of 6,000 would seem to be a safe estimate for the 1990s.



WHERE ARE MODERN LANGUAGES SPOKEN?

The geographical distiribution of living languages, according to *Ethnologue* (Grimes, 1992), based on a total of 6,528 languages.

HOW MANY LANGUAGES HAVE THERE BEEN?

Based on what is known about the rate of language change at which new languages develop from a common origin (p. 331), it is possible to speculate about the number of languages which may have existed since the emergence of a human language faculty. Cautious estimates suggest 30,000; radical ones, over 500,000. A plausible 'middle of the road' figure is 150,000.

49 · THE ORIGINS OF LANGUAGE

For centuries, people have speculated over the origins of human language. What is the world's oldest spoken language? Have all languages developed from a single source? What was the language spoken in the Garden of Eden? How did words come to be, in the very beginning? These questions are fascinating, and have provoked experiments and discussion whose history dates back 3,000 years. The irony is that the quest is a fruitless one. Each generation asks the same questions, and reaches the same impasse – the absence of any evidence relating to the matter, given the vast, distant time-scale involved. We have no direct knowledge of the origins and early development of language, nor is it easy to imagine how such knowledge might ever be obtained. We can only speculate, arrive at our own conclusions, and remain dissatisfied. Indeed, so dissatisfied was one group of 19th-century scholars that they took drastic action: in 1866, the Linguistic Society of Paris published an edict banning discussion of the topic at their meetings. But the theorizing continues, and these days there is a resurgence of interest, as new archaeological finds and modern techniques of analysis provide fresh hints of what may once have been.

EARLY 'EXPERIMENTS'

The lengths to which some people have gone in order to throw light on the question are truly remarkable – if the accounts are to be believed. One of the best-known reports concerns the Egyptian king, Psamtik I, who reigned in the 7th century BC. According to the Greek historian, Herodotus, Psamtik wished to find out which of all the peoples of the world was the most ancient. His way of determining this was to discover the oldest language which, he thought, would be evidence of the oldest race. This is how Herodotus tells the story.

He gave two new-born babies of ordinary men to a shepherd, to nurture among his flocks after this manner. He charged him that none should utter any speech before them, but they should live by themselves in a solitary habitation; and at the due hours the shepherd should bring goats to them, and give them their fill of milk, and perform the other things needful. Thus Psammetichus did and commanded because he desired, when the babes should be past meaningless whimperings, to hear what tongue they would utter first.

And these things came to pass; for after the shepherd had wrought thus for a space of two years, when he opened the door and entered in, both the babes fell down before him, and cried *becos*, and stretched out their hands. Now when the shepherd heard it for the first time, he held his peace; but when this word was often-times spoken as he came to care for them, then he told his lord, and brought the children into his presence when he commanded. And when Psammetichus had also heard it, he enquired which nation called anything *becos*; and enquiring, he found that the Phrygians call bread by this name. Thus the Egyptians, guided by this sign, confessed that the Phrygians were elder than they. That so it came to pass I heard of the priests of Hephaestus in Memphis.

Phrygian is now extinct, but at the time it was spoken in an area corresponding to the north-western part of modern Turkey.

Psamtik's conclusion was wrong, for we know from philological studies that Phrygian is but one of several languages which had developed in that period of history. So why did the children say *becos*? Doubtless they had begun to babble naturally and repetitively to each other, in a similar way to twins (see p. 249), and this was one of the 'snatches' that the shepherd recognized. Some commentators have even suggested that they were imitating the sound of the goats!

Whether the Psamtik experiment ever took place is open to question. Possibly the origins of the story lie in a fiction invented by someone to discredit the Egyptians. But whatever the reality, the initiative credited to Psamtik has apparently had its parallels in several later times and places. At least two similar experiments have been reported – though again, there are doubts as to their authenticity (see also p. 230).

Psamtik I of Egypt (663–610 BC)

The Holy Roman Emperor, Frederick II of Hohenstaufen (1194–1250) also carried out an experiment with children. According to the chronicle of a Franciscan friar, Brother Salimbene:

He made linguistic experiments on the vile bodies of hapless infants, bidding foster-mothers and nurses to suckle and bathe and wash the children, but in no wise to prattle or speak with them; for he would have learnt whether they would speak the Hebrew language (which had been the first), or Greek, or Latin, or Arabic, or perchance the tongue of their parents of whom they had been born. But he laboured in vain, for the children could not live without clappings of the hands, and gestures, and gladness of countenance, and blandishments.

James IV of Scotland (1473–1513) is said to have carried out a similar experiment. The account given in the History of Robert Lindesay of Pitscottie runs as follows:

The king took a dumb woman and put her in Inchkieth, and gave her two young children in company with her, and furnished them of all necessary things pertaining to their nourishment, that is to say food, drink, fire and candle, clothes, with all other kinds of necessaries which is required to man or woman, desiring the effect hereof to come to know what language the children would speak when they came to lawful age. Some say they spoke good Hebrew, but as to myself I know not but by hearsay.

FIVE THEORIES OF THE ORIGINS OF LANGUAGE

The Danish linguist, Otto Jespersen (1860–1943), grouped commonly held theories about the origins of language into four types, and added a fifth of his own. They are often referred to by nicknames.

The 'bow-wow' theory
Speech arose through people imitating the sounds of the environment, especially animal calls. The main evidence would be the use of onomatopoeic words (p. 176), but as few of these exist in a language, and as languages vary so much in the way they represent natural sounds, the theory has little support.

The 'pooh-pooh' theory
Speech arose through people making instinctive sounds, caused by pain, anger, or other emotions. The main evidence would be the universal use of sounds as interjections (p. 91), but no language contains many of these, and in any case the clicks, intakes of breath, and other noises which are used in this way bear little relationship to the vowels and consonants found in phonology. The spelling is never a satisfactory guide.

The 'ding-dong' theory
Speech arose because people reacted to the stimuli in the world around them, and spontaneously produced sounds ('oral gestures') which in some way reflected or were in harmony with the environment. The main evidence would be the universal use of sounds for words of a certain meaning, but apart from a few cases of apparent sound symbolism (p. 176), the theory has nothing to commend it. Several fanciful examples have nonetheless been cited – *mama* is supposed to reflect the movement of the lips as the mouth approaches the breast, and *bye-bye* or *ta-ta* show the lips and tongue respectively 'waving' good-bye.

The 'yo-he-ho' theory
Speech arose because, as people worked together, their physical efforts produced communal, rhythmical grunts, which in due course developed into chants, and thus language. The main evidence would be the universal use of prosodic features (p. 171), especially of rhythm; but the gap between this kind of expression and what we find in language as a whole is so immense that an explanation for the latter would still have to be found.

The 'la–la' theory
Jespersen himself felt that, if any single factor was going to initiate human language, it would arise from the romantic side of life – sounds associated with love, play, poetic feeling, perhaps even song. But again, the gap between the emotional and the rational aspects of speech expression would still have to be accounted for.

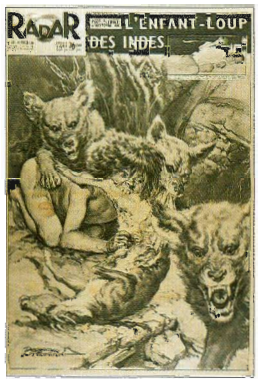
CHILDREN OF THE WILD

For several hundred years, cases have been reported of children who have been reared in the wild by animals or kept isolated from all social contact. These cases are listed below, adapted from Lucien Malson's *Wolf Children* (1972). Sometimes the information is based on little more than a brief press report. At other times, the cases have been studied in detail – in particular, the stories of Victor, Kaspar Hauser, Amala and Kamala, and Genie.

The ideas of Psamtik I receive no support at all from these children. Only some of the reports say anything about the children's language abilities, and the picture is quite clear: none could speak at all, and most had no comprehension of speech. Most attempts to teach them to speak failed. The cases of 1694, 1731, and 1767 (Fraumark) are said to have learned some speech, and Tomko of Hungary (also 1767) is reputed to have learned both Slovak and German. The 1717 girl and the 19th-century Bankipur child are both said to have learned some sign language. But of the well-attested cases, the results are not impressive. Victor, the 'Wild Boy of Aveyron', remained unable to speak, though he could understand and read to some extent. Kamala of Midnapore learned some speech and sign. The two most successful cases on record are Kaspar Hauser, whose speech became quite advanced, and Genie (p. 265), who learned a few words immediately after discovery, and whose subsequent progress in speech was considerable.

RECORDED CASES OF CHILD ISOLATION

	Date of discovery	Age at discovery		Date of discovery	Age at discovery
Wolf-child of Hesse	1344	7	Wolf-child of Sekandra	1872	6
Wolf-child of Wetteravia	1344	12	Child of Sekandra	1874	10
Bear-child of Lithuania	1661	12	Wolf-child of Kronstadt	?	23
Sheep-child of Ireland	1672	16	Child of Lucknow	1876	?
Calf-child of Bamberg	c1680	?	Child of Jalpaiguri	1892	8
Bear-child of Lithuania	1694	10	Child of Batsipur	1893	14
Bear-child of Lithuania	?	12	Child of Sultanpur	1895	?
Kidnapped Dutch girl	1717	19	Snow-hen of Justedal	?	12
Two boys of Pyrenees	1719	?	Amala of Midnapore	1920	2
Peter of Hanover	1724	13	Kamala of Midnapore	1920	8
Girl from Sogny	1731	10	Leopard-child of India	1920	?
Jean of Liège	?	21	Wolf-child of Maiwana	1927	?
Tomko of Hungary	1767	?	Wolf-child of Jhansi	1933	?
Bear-girl of Fraumark	1767	18	Leopard-child of Dihungi	?	8
Victor of Aveyron	1799	11	Child of Casamance	1930s	16
Kaspar Hauser of Nuremberg	1828	17	Assicia of Liberia	1930s	?
Sow-girl of Salzburg	?	22	Confined child of Pennsylvania	1938	6
Child of Husanpur	1843	?	Confined child of Ohio	1940	?
Child of Sultanpur	1843	?	Gazelle-child of Syria	1946	?
Child of Sultanpur	1848	?	Child of New Delhi	1954	12
Child of Chupra	?	?	Gazelle-child of Mauritania	1960	?
Child of Bankipur	?	?	Ape-child of Teheran	1961	14
Pig-boy of Holland	?	?	Genie, U.S.A.	1970	13½
Wolf-child of Holland	?	?			



Ramu the 'wolf-child' found near Lucknow, India. The magazine is dated 21 February 1954.

Peter the 'wild boy' found in a wood in Hanover and brought to England by King George II.

SCIENTIFIC APPROACHES

By contrast with the informal discussion and speculation of preceding centuries, serious attempts have been made in recent years to see if modern science can throw any light on the question of the origins of language. The study is sometimes called *glossogenetics* – the study of the formation and development of human language, in both the child and the race. The main sciences involved are biology (especially sociobiology), anthropology, psychology, semiotics (p. 403), neurology (for the study of brain evolution), primatology, and linguistics.

THE EVIDENCE FROM PALAEONTOLOGY

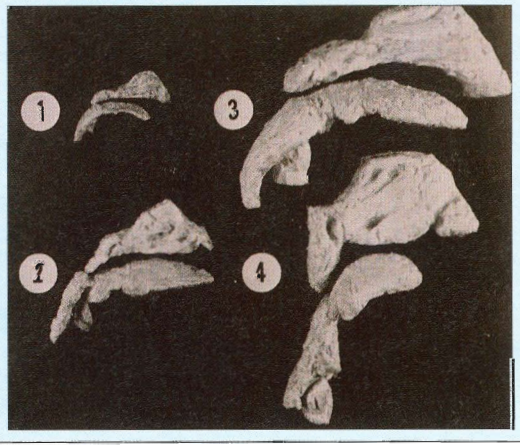
Might it be possible to deduce, from the fossil record of early man, the point at which speech began? The matter has been well investigated, but the results are not conclusive.

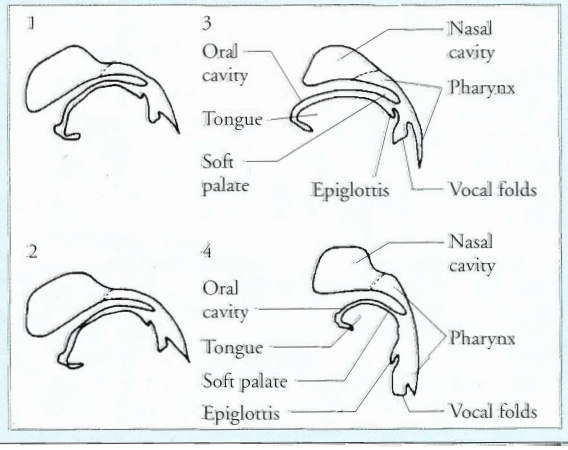
It is possible to make plaster casts of the bony cavities within the skulls which have been found. It can be shown, for example, that both Neanderthal man and Cro-Magnon man (pre-30,000 BC) had similar brain sizes to that of modern man. But this information is of limited value. The relative size and shape of the brain can be established, but none of the more relevant detail (such as the orientation of the various furrows, or sulci (§45)). In any case, there is no direct correlation between the size of a brain and the use of language: in modern man, language is found in people with small brains, such as nanocephalic dwarfs, or children who have had large areas of brain removed – and some gorillas have a brain size close to these. It is plausible that an increase in the number of brain cells increases intellectual or linguistic capacity, but no correlation has been established.

Another way of looking at the problem is to ask whether primitive man had the physiological capacity to speak, and this has led to a great deal of interesting research. The problem is that only the shape of the jaws and the oral cavity are preserved in fossils; there is no direct information about the size and shape of the soft tissues of tongue, pharynx or larynx, nor about the ability to move these organs (§22). Most of the reasoning has therefore had to be based on reconstruction using plaster casts, and comparison with the physiology and vocalization of present-day primates and human infants.

It is possible to say with some conviction, using this kind of argument, that the older hominids did not possess speech; but the position of the more recent remains is unclear. It is unlikely that *Australopithecus* (who appeared around 4–5 million BC) could speak, but the evidence is ambiguous for Neanderthal man (70–35,000 BC). Linguists and anatomists have compared the reconstructed vocal tract of a Neanderthal skull with those of a newborn and an adult modern human. The newborn and Neanderthal vocal tracts are remarkably similar. Neanderthal man would have been able to utter only a few front consonant-like sounds and centralized vowel-like sounds, and may have been unable to make a contrast between nasal and oral sounds. This is well below what is found in the phonologies of the world's languages today (p. 167). It would have been possible to construct a linguistic code out of these limited sounds, but it would have required a level of intellectual ability apparently lacking at that evolutionary stage. On the other hand, these phonetic abilities are far ahead of modern primates. It has thus been concluded that Neanderthal man represents an intermediate stage in the gradual evolution of speech. Cro-Magnon man (35,000 BC), by contrast, had a skeletal structure much more like that of modern man.

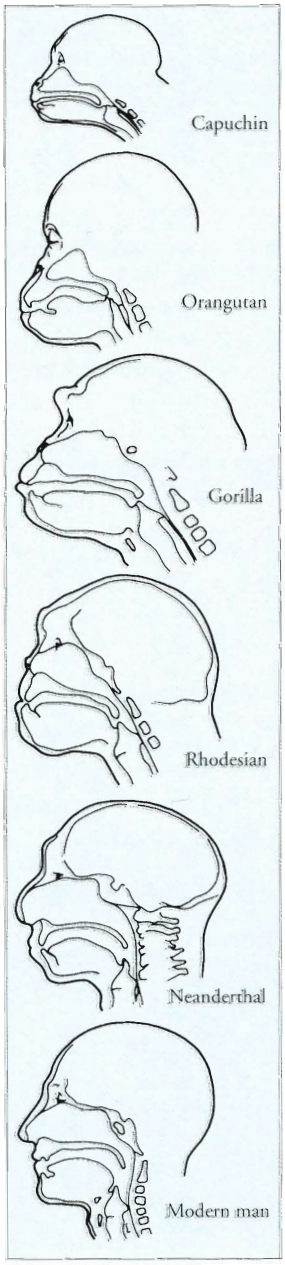
Casts of the nasal, oral, and pharyngeal air passages of (1) a newborn baby, (2) an adult chimpanzee, (3) a Neanderthal reconstruction, and (4) an adult man. The differences in dimensions can be clearly seen (below right) when the four tracts are drawn so that they are nearly equal in size. (After P. Lieberman, 1972.)





PRIMATE VOCAL TRACTS

The vocal tracts of primates are very different from that of modern man. They have long, flat, thin tongues, which have less room to move. The larynx is higher, and there is little sign of a pharynx. There is no evidence of ability to change the configurations of the vocal tract, to produce the range of sounds required in speech. In the course of evolution, posture becomes erect, and the head moves forward. The larynx descends and a long, flexible pharyngeal cavity develops. (From V. E. Negus, 1949.)



HOMO LOQUENS

It would seem that the human vocal tract evolved from a non-human primate form to enable fast and efficient communication to take place. Speech is not merely the incidental result of a system designed for breathing and eating. The changes that took place in the larynx, pharynx, and mouth came about at the cost of less efficient breathing, chewing, and swallowing. Modern man can choke from food lodged in the larynx; monkeys cannot. The survival value of speech must be considerable to compensate for such deficiencies. The human being, in short, seems to have evolved as a speaking animal – *homo loquens.*

Some hominids had a human-like vocal tract as far back as 200,000 BC, but they probably did not have a sufficiently developed nervous system to control it. There is general agreement on a time-scale from 100,000 to 20,000 BC for the development of speech. If the Neanderthal evidence is accepted, this scale narrows to 50–30,000 BC, in the latter part of the Upper Palaeolithic period.

This is the conclusion regarding speech. But the lack of physical similarities with modern man does not prove that there was no language in an abstract sense, or other modes of communication. At the time indicated above, cultural development was relatively advanced, and there must have been some efficient way of transmitting information about skills from one generation to the next. Any degree of social interdependence – as found in tribal grouping, religious activity, or group hunting techniques – would seem to require a communication system. Cave drawings of the period also suggest the existence of an intellectual capability such as would be required for language.

An elaborate gesture system is one possibility. The early development of language may well have been assisted by some kind of signing, which would have been the simplest way of communicating basic meanings – such as how to use tools. Hands were no longer necessary for locomotion, so they could be used for other activities. Perhaps primitive people who were skilful in using signs stood a better chance of survival. Natural selection could then have led to the development of the intellectual faculties prerequisite for speech.

Learning to use tools, and to pass the skills on, would be most efficiently done through language. It has even been suggested that learning to use tools and learning language are interrelated skills. They are localized in the same general area of the brain (§45); and both tool using and gesture require sophisticated use of the hands. However, some non-human primates can use tools, and it is unlikely that the hands could have been used for two such different purposes for long. Nonetheless, in an indirect way, tools could have promoted the development of speech. Sounds made at the same time as the gestures might have come to be associated with various activities. The idea has been proposed that, as tools came to be used for more advanced purposes, food would be stored, so that there would be intervals between meals, and thus more time available for the mouth to be put to other uses – such as the development of spoken language.

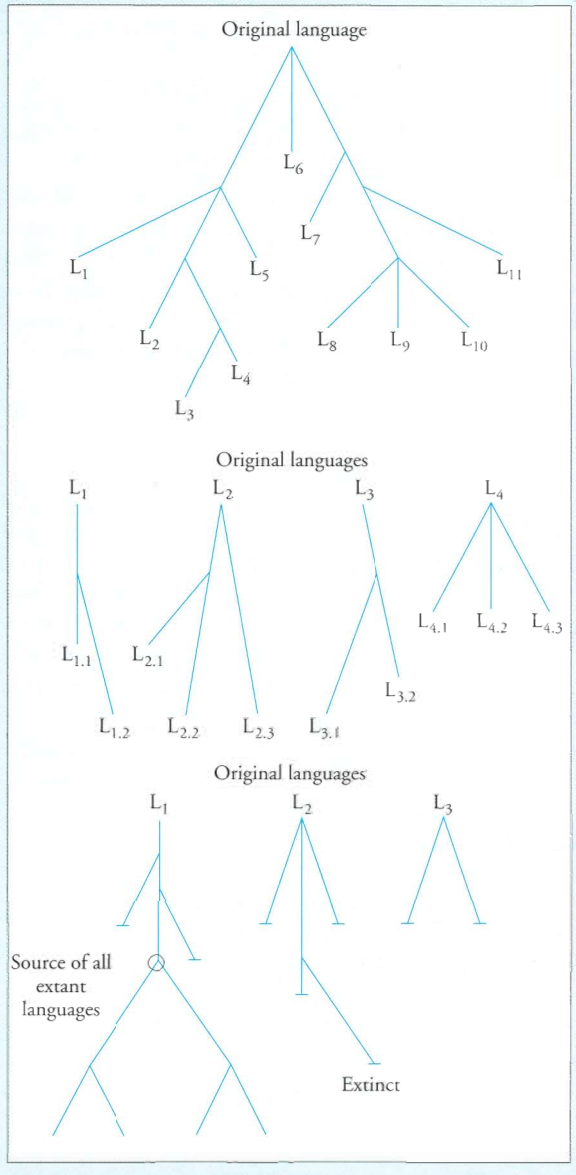
We can only speculate about the link between oral and gestural language. Similarly, the gap between human language and the communication systems of the nearest primates remains vast, and there is no sign of a language-like increase in communicative skills as one moves from lower to higher mammals. Human language seems to have emerged within a relatively short space of time, perhaps as recently as 30,000 years ago. But that still leaves a gap of over 20,000 years before the first unequivocal evidence of written language (p. 198).

WAS THERE EVER AN ORIGINAL LANGUAGE?

Right: The view that all languages have diverged from a common source, the result of cultural evolution or divine intervention, is known as *monogenesis*. The existence of differences between languages is then explained as a result of people moving apart, in waves of migration around the world. In this view, language universals (§14) would be interpreted as evidence of common origin.

Centre: The opposite view, that language emerged more or less simultaneously in several places, is known as *polygenesis*. Language universals, and other similarities between languages, are then explained by pointing to the similar constraints which must have operated upon the early speakers (in terms of both their physiology and their environment), and by the likelihood that, as groups came into contact, their languages would influence each other – a process known as *convergence*.

Bottom: There is also a third possibility, given the vast time-scale involved. All of the languages that now exist may indeed have diverged from a common source, but this may have been just one line of descent from an earlier era when several independent languages emerged.



50 · FAMILIES OF LANGUAGES

The first scientific attempts to discover the history of the world's languages were made at the end of the 18th century. Scholars began to compare groups of languages in a systematic and detailed way, to see whether there were correspondences between them. If these could be demonstrated, it could be assumed that the languages were related – in other words, that they developed from a common source, even though this might no longer exist.

Evidence of a common origin for groups of languages was readily available in Europe, in that French, Spanish, Italian, and other Romance languages (p. 303) were clearly descended from Latin – which in this case is known to have existed. The same reasoning was applied to larger groups of languages, and by the beginning of the 19th century, there was convincing evidence to support the hypothesis that there was once a language from which many of the languages of Eurasia have derived. This proto-language came to be called Proto-Indo-European (p. 298). Very quickly, other groups of languages were examined using the same techniques.

The main metaphor that is used to explain the historical relationships is that of the language family, or family tree. Within the Romance family, Latin is the 'parent' language, and French, Spanish, etc. are 'daughter' languages; French would then be called a 'sister' language to Spanish and the others. The same approach is used with larger groups. Within the Indo-European family, Proto-Indo-European is the parent language, and Latin, Greek, Sanskrit, and others are the daughter languages. In a large family, it will be necessary to distinguish various 'branches' , each of which may contain several languages, or 'sub-families' of languages.

This way of talking must not be taken too literally. A 'parent' language does not live on after a 'daughter' language is 'born', nor do languages suddenly appear in the way implied by the metaphor of birth. Nor is it true that, once branches of a family begin to emerge, they develop quite independently, and are never afterwards in contact with each other. Languages converge as well as diverge. Furthermore, stages of linguistic development are not as clear-cut as the labels on a family tree suggest, with change operating smoothly and uniformly throughout. Linguistic change, we now know, is much more uneven, with different social groups responding to change in different ways (p. 330).

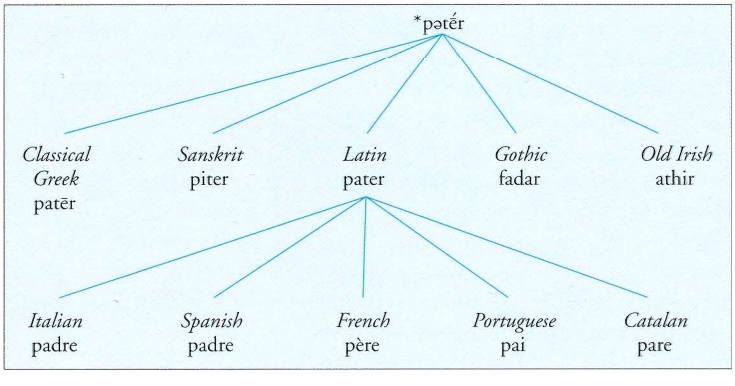
Since the 19th century, other classificatory terms have come into use. *Family* is still used as a general term for any group of languages where there is a likelihood of a historical relationship (and this is the way the terms is used in this encyclopedia). But in some classifications, a distinction is drawn in terms of how definite the relationship is. If there is clear linguistic evidence of a close relationship, the term *family* continues to be used; but where the relationship is less definite, or more remote, the grouping is referred to as a *phylum*. Sometimes the term *macro-phylum* is used for yet more general and less definite groupings. It is evident, for example, that all the Aboriginal languages of Australia (p. 326) are related, but as there is no clear-cut historical evidence which bears on the matter, and little typological work, scholars often refer to the Australian '(macro)phylum' rather than to the Australian 'family'.

THE COMPARATIVE METHOD

In historical linguistics, the *comparative method* is a way of systematically comparing a series of languages in order to prove a historical relationship between them. Scholars begin by identifying a set of formal similarities and differences between the languages, and try to work out (or 'reconstruct') an earlier stage of development from which all the forms could have derived. The process is known as *internal reconstruction*. When languages have been shown to have a common ancestor, they are said to be *cognate*.

The clearest cases are those where the parent language is known to exist. For example, on the basis of the various words for 'father' in the Romance languages, given below, it is possible to see how they all derived from the Latin word *pater*. If Latin no longer existed, it would be possible to reconstruct a great deal of its form, by comparing large numbers of words in this way. Exactly the same reasoning is used for cases where the parent language does not exist, as when the forms in Latin, Greek, Sanskrit, Welsh, etc. are compared to

August Schleicher (1821–68) The 'family tree' theory (*Stammbaumtheorie*) was introduced by the German linguist Schleicher, who thought of language as an organism which could grow and decay, and whose changes could be analysed using the methods of the natural sciences.



reconstruct the Indo-European form, **pəter*. The asterisk in front of a form, in historical linguistics, shows that the form in question is a reconstruction which has not been attested in written records. Exactly how such reconstructed forms were pronounced is a matter of (at times fierce) debate: some scholars are happy to assign phonetic values to the forms, and pronounce them as if they were part of a real language; others argue that the forms are little more than abstract formulae, summarizing the sets of correspondences which have been noted (§54).

TYPES OF LINGUISTIC CLASSIFICATION

There are two main ways of classifying languages: the *genetic* (or *genealogical*) and the *typological* (§14). Both are used in contemporary language work, but the former has received far more investigation, and has the better developed procedures and frame of reference. A further approach (an *areal* classification) is reviewed in §8.

GENETIC CLASSIFICATION

This is a historical classification, based on the assumption that languages have diverged from a common ancestor. It uses early written remains as evidence, and when this is lacking, deductions are made using the comparative method to enable the form of the parent language to be reconstructed. The approach has been widely used, since its introduction at the end of the 18th century, and provides the framework within which all world-wide linguistic surveys to date have been carried out. The success of the approach in Eurasia, where copious written remains exist, is not matched in most other parts of the world, where a classification into families is usually highly tentative.

TYPOLOGICAL CLASSIFICATION

This is based on a comparison of the formal similarities which exist between languages. It is an attempt to group languages into structural types, on the basis of phonology, grammar, or vocabulary, rather than in terms of any real or assumed historical relationship. For example, it is possible to group languages in terms of how they use sounds – how many and what kinds of vowels they have, whether they use clicks, whether they use tones, and so on. Languages can also be classified in terms of whether their word order is fixed or free, and which order is favoured (p. 98). The earliest typologies, however, were in the field of morphology (p. 90). These, propounded by August von Schlegel (1767–1845) and others in the early 19th century, recognized three main linguistic types, on the basis of the way a language constructs its words.

Isolating, analytic, or root languages

All the words are invariable: there are no endings Grammatical relationships are shown through the use of word order. Chinese, Vietnamese, and Samoan are clear cases. For example, 'I bought some oranges to eat' in Beijing Chinese would be:

	Wǒ	*mǎi júzi*	*chī*
literally,	I	buy orange	eat

Inflecting, synthetic, or fusional languages

Grammatical relationships are expressed by changing the internal structure of the words – typically by the use of inflectional endings (p. 90) which express several grammatical meanings at once. Latin, Greek, and Arabic are clear cases. For example, the *-o* ending of Latin *amo* 'I love' simultaneously expresses that the form is in the first person singular, present tense, active, and indicative.

Agglutinative or agglutinating languages

Words are built up out of a long sequence of units, with each unit expressing a particular grammatical meaning, in a clear one-to-one way. A sequence of five affixes might express the meaning of *amo*, for example – one for each category of person, number, tense, voice, and mood. Turkish, Finnish, Japanese, and Swahili form words in this way. 'He who gets water for me' in Swahili is *anayenipatia maji*, which can be analysed as:

a – na	*– ye*	*– ni*	*– pat*	*– i*	*– a*	(*maji*)
he	PRESENT TENSE who	me	gets	for		(water)

Polysynthestic or incorporating languages

Words are often very long and complex, containing a mixture of agglutinating and inflectional features, as in Eskimo, Mohawk, and Australian languages. For example, the aboriginal language Tiwi expresses 'I kept on eating' as *ngirruunthingapukani*, which is analysable as:

ngi – rru	*– unthing*	*– apu – kani*	
I	PAST TENSE for some time	eat	repeatedly

Some linguists, however, do not regard this as a separate typological category.

PHILOLOGIST – OR LINGUIST?

People who study the history of languages are sometimes called *comparative philologists* (or just 'philologists') and sometimes *historical linguists*. The difference lies partly in the training, partly in the subject matter. The philological tradition is one of painstaking textual analysis, often related to literary history, and using a fairly traditional descriptive framework. The newer, linguistic approach tends to study historical data more selectively, as part of the discussion of broader issues in linguistic theory, in the process using the conceptual apparatus of modern linguistics. Needless to say, proponents of the two approaches do not always see eye to eye. Philologists are often still sceptical of the new science, remembering the days when linguists considered historical topics to be of secondary importance (§65). Historical linguists, similarly, are often impatient with the philologist's reluctance to develop general explanatory theories of language change. But nowadays there are many signs that the skills of these two categories of scholar are being seen as complementary, not in opposition.

WHAT SORT OF LANGUAGE IS ENGLISH?

English is a Germanic language, according to the genetic method of classification. But from other points of view, the picture alters. Culturally, it displays many similarities with Romance, in view of the large number of loan words (p. 332) it has taken in from French and Italian, and the way these languages have even exercised some influence on grammar (e.g. *chicken supreme*) and phonology (e.g. the use of final /ʒ/ in words like *garage*). If we consider the place names of North America, then we have to allow a relationship with Amerindian languages (*Chappaquiddick, Susquehanna*). From a typological viewpoint, English is in fact more similar to an isolating language like Chinese than Latin: there are few inflectional endings, and word-order changes are the basis of the grammar.

Three-in-one

Isolating: The boy will ask the girl.
The girl will ask the boy.
Inflecting: The biggest boys have been asking.
Agglutinating: anti-dis-establish-ment-arian-ism.

THE PROBLEM OF CLASSIFICATION

These days, typological questions are of undoubted interest – especially in relation to the search for language universals (§14). But some of the early classifications have been severely criticized because of the way they were interpreted. No one would now follow the early tendency of typologists, under the influence of Darwin, to evaluate languages as if they were points on an evolutionary scale – that isolating languages are 'not as well developed' as inflecting languages, for example. Nor is there any evidence that languages of a particular type are inevitably associated with particular geographical areas, or with people of a particular ethnic or cultural group. It must also be appreciated that there is no such thing as a 'pure' instance of one of the above types. Languages seem to have these characteristics to various degrees.

Is a typological classification possible therefore? Everything depends on how we evaluate the variables which provide the basis of the classification. Morphology is only one variable. When we take into account *all* the features of language – syntax, phonology, discourse, and language use (§13) – the nature of the problem is evident. There are a vast number of possible classifications, and how should we decide which criteria are the most important? If two languages are 90% similar in phonology and 50% similar in grammar, are they more or less closely related than two languages which are 50% similar in phonology and 90% in grammar? Linguistic theory has hardly begun to answer such questions.

Both typological and genetic classifications ignore the relevance of cultural links between languages – the fact that languages influence each other by contact, such as by borrowing words from each other. Sometimes languages that have no historical relationship can converge so that they seem to be members of the same family. Conversely, related languages can be influenced by other languages so much that the differences become more striking than the similarities. The role of cultural contact is a real problem in studying many language families, where it is often totally unclear whether two languages are similar because they share a common origin, or because they have borrowed from each other (p. 332).

Some linguists have tried to move away from a classification into general types, proposing instead to rank languages in terms of individual structural criteria. One criterion could be the number of morphemes (p. 90) per word in a language (an 'index of synthesis'). In the sentence, 'The boys saw the girls', there are 5 words but 8 morphemes, producing a synthetic index of 1.6. Using this criterion, according to one study, the average for English was 1.68, compared with 1.06 for Annamese and 3.72 for Eskimo. There are several other grammatical ratios which could be investigated in this way.


GREENLAND
ALASKA
CANADA
UNITED STATES OF AMERICA
NORTH KOREA
SOUTH KOREA
JAPAN
PACIFIC OCEAN
ATLANTIC OCEAN
TAIWAN
HAWAII
PHILIPPINES
MARSHALL ISLANDS
MEXICO
CUBA
DOMINICAN REP
HAITI
BELIZE
JAMAICA
HONDURAS
GUATEMALA
EL SALVADOR
NICARAGUA
COSTA RICA
PANAMA
VENEZUELA
GUYANA
SURINAME
FRENCH GUIANA
COLOMBIA
CUBA
ECUADOR
PAPUA NEW GUINEA
INDONESIA
SOLOMON ISLANDS
TUVALU
KIRIBATI
MELANESIA
WESTERN SAMOA
SOCIETY ISLANDS
VANUATU
FIJI
TONGA
FRENCH POLYNESIA
PERU
BRAZIL
BOLIVIA
PARAGUAY
EASTER ISLAND
AUSTRALIA
CHILE
ARGENTINA
URUGUAY
TASMANIA
NEW ZEALAND
Oto-Manguean
Paleosiberian
Penutian
Sino-Tibetan
Siouan
Tai
Uralic
Uto-Aztecan

51 · THE INDO-EUROPEAN FAMILY

'Indo-European' is the name scholars have given to the family of languages that first spread throughout Europe and many parts of southern Asia, and which are now found, as a result of colonialism, in every part of the world. The parent language, generally known as 'Proto-Indo-European', is thought to have been spoken before 3000 BC, and to have split up into different languages during the subsequent millennium. The differences were well-established between 2000 and 1000 BC, when the Greek, Anatolian, and Indo-Iranian languages are first attested.

WHO WERE THE INDO-EUROPEANS?

Archaeological evidence shows the existence of a semi-nomadic population living in the steppe region of southern Russia around 4000 BC, who began to spread into the Danube area of Europe and beyond from around 3500 BC. The people are known as the Kurgans, because of their burial practices (*kurgan* being the Russian for 'burial mound'). Kurgan culture seems to have arrived in the Adriatic region before 2000 BC, and this coincides well with the kind of time-scale needed to produce large amounts of linguistic change. The ancestors of the Kurgans are not known, though there are several similarities between Proto-Indo-European and the Uralic family of languages (p. 306), spoken further east, and these may well have had a common parent, several thousand years before.

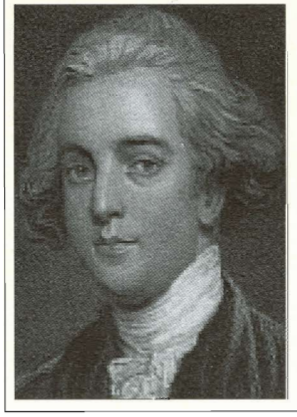
By comparing the similar vocabulary of the extant Indo-European languages, it is possible to draw some conclusions about the geographical origins and life-style of the people. For instance, many family words (such as 'mother', 'husband', 'brother') can be reconstructed for Proto-Indo-European. These include several words for 'in-laws', which seem to have been used solely with reference to the bride. Evidence of this kind suggests that it was the wife who was given a position within the husband's family, rather than the other way round, and that the society must therefore have been patriarchal in character.

The reconstructed language has words for horses, dogs, sheep, pigs, and other animals; there is a word for some kind of vehicle, and this vehicle definitely had wheels; there are many words for parts of the body; there are several words relating to farming, and a few words relating to tools and weapons; many abstract notions are attested, relating to such fields as law, religious belief, and social status; numerals went to at least 100. Words relating to fauna and flora are of particular interest, for they can provide clues as to the place of origin of the people. There are no words for 'palm tree' or 'vine', for example, which suggests, independently of any archaeological evidence, that the migrations did not begin in the Mediterranean area. But other clues often seem contradictory. The word for 'beech tree' is widely attested, and, as this tree does not grow in Asia, it has been suggested that the Indo-Europeans must have originated in north-central Europe. On the other hand, there is little evidence of a common word for 'oak', which is also a European tree, and if this word was not known to the Indo-Europeans, the view is supported that their migration must have begun in Asia after all. Indo-European philology (§50) raises many fascinating questions of this kind.

THE DISCOVERY OF PROTO-INDO-EUROPEAN

It was not possible to deduce the existence of this family of languages until scholars became aware of the systematic resemblances which can be found between European languages and Sanskrit, the oldest-attested language of the Indian sub-continent. When these were first noticed, in the 16th century, many people thought that Sanskrit was the parent of the European languages; but towards the end of the 18th century the systematic studies began which showed conclusively that this was not the case.

Following an early statement of the common origin hypothesis in 1786, by Sir William Jones, the early 19th century produced several major works which laid the foundation of Indo-European philology. In 1816, the German philologist Franz Bopp published a study, whose scope is well illustrated by its title (translated): *On the conjugation system of the Sanskrit language, in comparison with those of the Greek, Latin, Persian and Germanic languages*. The relationship of Germanic to Latin, Greek, Slavic, and Baltic was demonstrated in a work written in 1814 by the Danish linguist, Rasmus Rask, but not published until 1818: *Investigation on the Origin of the Old Norse or Icelandic Language*. Further philological treatises followed, mainly written by Germans, such as Jakob Grimm and August Schleicher. In 1833, Bopp began the publication of the first major Indo-European grammar: *Comparative Grammar of Sanskrit, Zend, Greek, Latin, Lithuanian, Old Slavic, Gothic, and German*. It took 19 years to complete, and by its third edition incorporated Celtic and Albanian. In due course, this work and its contemporaries became out of date, as a result of the vast amount of philological study undertaken in the second half of the 19th century. A further publishing landmark was Karl Brugmann's *Outline of Comparative Indo-European Grammar* (1897–1916). A new *Indo-European Grammar*, the outcome of a project directed by the Polish linguist, Jerzy Kuryłowicz, commenced publication in 1968.



William Jones (1746–94) British orientalist and jurist, whose presidential address to the Bengal Asiatic Society in 1786 contained the following observation, generally quoted as the first clear statement asserting the existence of Indo-European:

The Sanskrit language, whatever be its antiquity, is of a wonderful structure; more perfect than the Greek, more copious than the Latin, and more exquisitely refined than either, yet bearing to both of them a stronger affinity, both in the roots of verbs, and in the forms of grammar, than could possibly have been produced by accident; so strong, indeed, that no philologer could examine them all three, without believing them to have sprung from some common source, which, perhaps, no longer exists.

Jacob Grimm (1785–1863) Well known to children everywhere for the collection of fairy tales and songs which he compiled with his brother. To linguists and philologists, he is also remembered for his major works in Germanic philology, especially his explanation of how the consonants of different Indo-European languages relate to each other. There is, for example, a regular relationship between words beginning with *p* in Latin and *f* in Germanic languages (as in *pater* and *father*), or between initial *t* in Greek and initial *th* in English (as in *treis* and *three*). The rules governing these sound shifts became known as 'Grimm's law' (p. 330).

WHAT DID PROTO-INDO-EUROPEAN SOUND LIKE?

There are no written records relating to this period. The Kurgans must have been illiterate – unlike the people of Egypt and Mesopotamia of the time. So the entire character of Proto-Indo-European has been the result of painstaking reconstruction on the part of philologists, using the methods outlined on p. 294.

There is general agreement about the number of contrasts in the consonant system (p. 167), though the status of some of the less well-attested sounds (such as /b/) is disputed. This system seems largely to have been composed of plosives (p. 159), organized into three series: voiceless, voiced, and (less definitely) voiced aspirate. Four main places of articulation were used: labial, dental, palatal or velar, and labio-velar. There was a single fricative, which was voiced or voiceless according to context. In addition, there were probably one or more laryngeal consonants (see below). There were two nasals, two continuants, and two semi-consonants (p. 154), all of which could occur at the centres of syllables as well as at syllable edges. This system may be summarized as follows:

	Labial	Dental	Palatal/ Velar	Labio-velar
Plosives				
Voiceless	p	t	k	k^w
Voiced	b	d	g	g^w
Voiced aspirate	bh	dh	gh	g^wh
Fricatives		s(z)		
Nasals	m	n		
Continuants		l r		
Semi-consonants	w		j	

There is more disagreement over the vowel system – vowels, as always (p. 169), being more difficult to analyse. Four main contrasts are generally recognized: mid-front, mid-back, open, and central, the first three occurring both in long and short forms (though how far these were independent contrasts, as opposed to laryngeally controlled variants, is unclear). In addition, some scholars recognize two further contrasts in close position, /i/ and /u/, but the overlap with the use of these sounds as semi-consonants makes this analysis less certain also. The possible vowel system can thus be summarized as follows:

(i) (u)
e/e: ə o/o:
a/a:

THE LARYNGEAL THEORY

Towards the end of the 19th century, the Swiss linguist Ferdinand de Saussure (p. 411) put forward the view that, in order to explain various anomalies in early Indo-European forms, an extra set of sounds would have to be postulated as occurring in Proto-Indo-European. Saussure did not suggest any phonetic details for these sounds, but later they came to be called *laryngeals*, a term taken from the study of Semitic languages (p. 318), where consonants in the region of the larynx were known to occur. Laryngeal consonants did not occur in any Indo-European language known at the time, but the previous existence of some kind of sound, it was argued, was indicated by the way they had caused the changes to take place in adjacent vowels (altering their length and quality) that had long been noticed in the extant languages.

The laryngeal theory was immediately controversial, and received little support for many years. But attitudes changed after 1927, when it was found that Hittite (discovered several years after the theory was postulated) had a sound, represented by *h*, that occurred in some of the places where Saussure had predicted the laryngeals should be. However, the phonetic character of these laryngeals is still quite unclear, and philologists disagree on how many laryngeal sounds there were, whether their phonetic properties can (or should) be defined, and whether better analyses can be found. It is generally recognized that there must have been three (some say four) types, pronounced somewhere in the back part of the mouth, probably as fricatives or glottal stops (p. 159). They are usually symbolized by *H* or schwa (ə), and numbered with subscripts (H_1, H_2, etc. or $ə_1$, $ə_2$, etc.) Alternative analyses which postulate an earlier vowel, rather than a laryngeal, have also been proposed.

Laryngeal theory can be illustrated in this way. Most Proto-Indo-European basic forms (or 'roots') had a structure of Consonant–Vowel–Consonant (CVC, which is often written as CeC, when discussing this language), e.g. **bher-* 'bring', **med-*, 'measure'. But several forms had only one consonant, e.g. **es-* 'be', **dō-* 'give'. It is argued that these roots can be reconstructed as having the regular CVC structure, by postulating a laryngeal as the 'missing' consonant, e.g. **Hes-*, **doH-*. In roots such as **doH-*, with a preceding vowel, when the laryngeal finally disappeared, it caused the vowel to lengthen, as is attested in Latin *dōnum* 'gift', and elsewhere. Using these techniques, it is possible to show that almost all the roots of the proto-language (there are still a few exceptions, such as numerals) had a CVC structure.

SOME GRAMMATICAL FEATURES

People often think that the oldest languages must have been simpler than their modern counterparts (§49). The noun inflections of Proto-Indo-European clearly show this not to be so. It is possible to reconstruct three genders (masculine, feminine, and neuter) and up to eight cases (nominative, vocative, accusative, genitive, dative, ablative, locative, instrumental). Adjectives agreed in case, number, and gender with the noun. The verb system was also rich in inflections, used for aspect, mood, tense, voice, person, and number (p. 93). Different grammatical forms of a word were often related by the feature of *ablaut*, or *vowel gradation*: the root vowel would change systematically to express such differences as singular and plural, or past and present tense, as is still the case in English *foot/feet* or *take/took*.

55 INDO-EUROPEAN VARIETIES

Our father, who art in Heaven...

Celtic
Ein Tad, yr hwn wyt yn y nefoedd (Welsh)
Ár n-atheir, atá ar neamh (Irish Gaelic)
Ar n-athair a tha air nèamh (Scottish Gaelic)
Ayr ain, t' ayns niau (Manx)
Agan tas ny, us yn nef (Cornish)

Germanic
Unser Vater, der Du bist im Himmel (German)
Undzer voter, vos bist im himl (Yiddish)
Fæder ure, þū þe eart on heofonum (Old English)
Onze vader, die in de hemelen zijt (Dutch)
Fader vår, du som er i himmelen (Norwegian)
Fader vår, som är i himmelen (Swedish)
Vor Fader, du som er i himlene (Danish)

Italic
Pater noster, qui es in caelis (Latin)
Notre père, qui es aux cieux (French)
Padre nuestro, que estás en los cielos (Spanish)
Pai nosso, que estás nos céus (Portuguese)
Pare nostre, que estau en lo cel (Catalan)

Albanian
Ati ynë që je në qiell

Greek
Páter 'ēmōn, 'o en toīs ouranoīs (New Testament)
Patéra mas, poù eisai stoùs ouranoús (Modern)

Baltic
Teve mūsų, kurs esi danguje (Lithuanian)
Mūsu tēvs debesīs (Latvian)
Tāwa noūson, kas tu essei en dangon (Old Prussian)

Slavic
Otĭče našĭ iže jesi na nebesĭchŭ (Old Church Slavonic)
Ótče naš, súščij na nebesách (Russian)
Ojča naš, katory jëść u nebe (Belorussian)
Otče naš, ščo na nebi (Ukrainian)

Ojcze nasz, któryś jest w niebiesiech (Polish)
Otče náš kterýž jsi v nebesích (Czech)
Otče náš, ktorý si v nebesiach (Slovak)

Oče naš, što si na neboto (Macedonian)
Oče naš, koji si na nebesima (Serbo-Croat)
Otče naš, kojto si na nebesata (Bulgarian)
Oče naš, ki si na nebesih (Slovene)

Armenian
Mer hayr or erknk'umn (East)
Ov hayr mer or erkink'n es (West)

Iranian
Max fyd, kæcy dæ ærvty midæg (Ossetic)
Yā bāwk-i ēma, ka la āsmān-ā-y (Kurdish)
Ei pedar-e-mā, ke dar āsmān hasti (Persian)
Phiṯḥ manī, ki bihishtā asti (Baluchi)
Aj jmuğplāra, če pa āsmān kxe ye (Pashto)

Indo-Aryan
Bho asmākham svargastha pitaḥ (Sanskrit)
Saggaṭha no pitā (Pali)
He hamāre svargbāsī pitā (Hindi)

He sādę pitā, jihṛā surg vic hai (Panjabi)
E asān-jā piu, jo āsmāna men āhe (Sindhi)
Ai sāni māli, yus asmānas paṭh chu (Kashmiri)
He hāmrā svargavāsī pitā (Nepali)

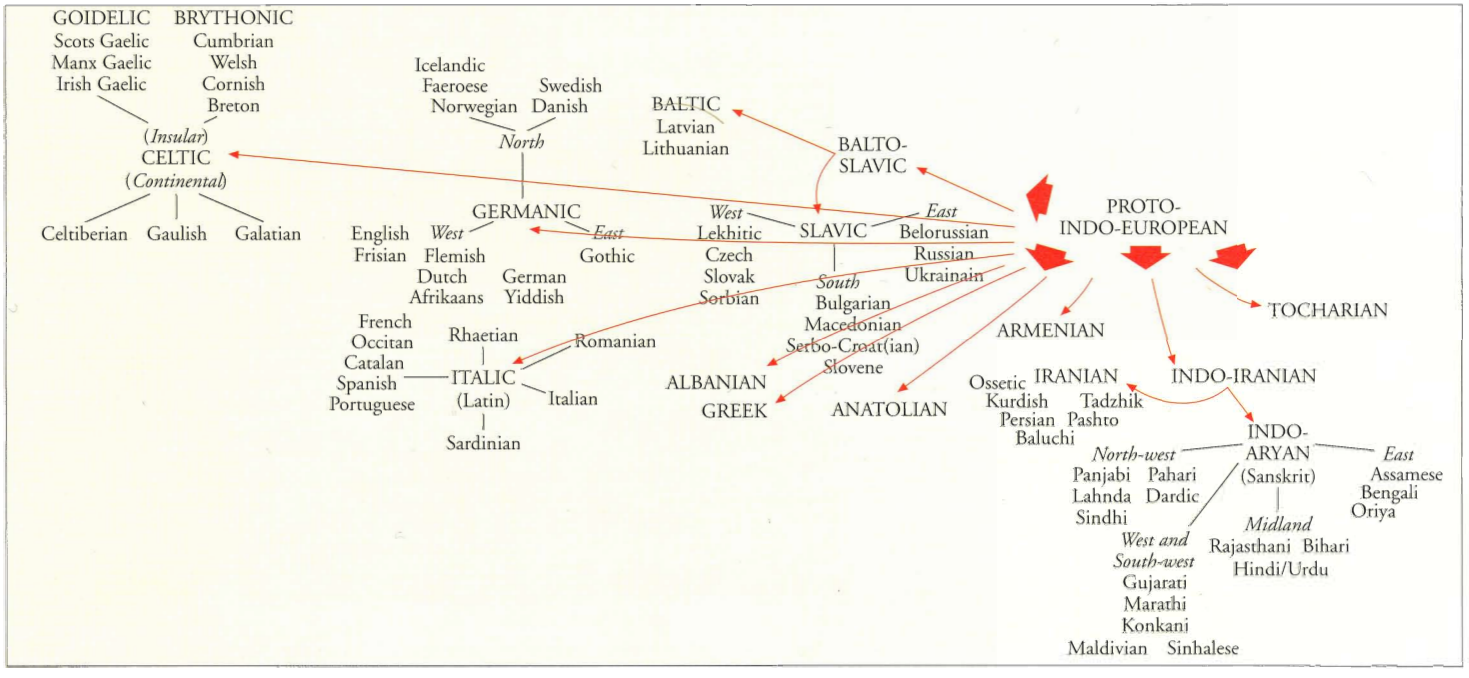
O ākāśamānnā amārā bāpa (Gujarati)
He āmacya svargātila pityā (Marathi)
Svargayehi vāḍasiṭina apagēpiyāṇeni (Sinhalese)

He āmār svargat thakā pitri (Assamese)
He āmāder svargastha pitā (Bengali)
He āmbhamānaṅka svargasha pitā (Oriya)

Dáde amaré, kaj isién k'o devlé (Romani)

GEOGRAPHICAL DISTRIBUTION OF THE INDO-EUROPEAN FAMILY OF LANGUAGES

THE INDO-EUROPEAN FAMILY TREE, REFLECTING GEOGRAPHICAL DISTRIBUTION




Kola Peninsula
White Sea
SIBERIA
YAKUT
Chukchi Peninsula
RUSSIA
URAL MOUNTAINS
CHUVASH
TATAR
SAYAN MTS
Kamchatka Peninsula
SAKHALIN ISLAND
KAZAKHSTAN
MONGOLIA
GEORGIA
ARMENIA
AZ.
AZ.
UZBEKISTAN
KYRGYZSTAN
XINJIANG
TURKMENISTAN
TAJIKISTAN
JAPAN
NORTH KOREA
SOUTH KOREA
SYRIA
IRAQ
IRAN
AFGHANISTAN
CHINA
JORDAN
PAKISTAN
NEPAL
BHUTAN
Ryukyu Islands
BANGLA DESH
SAUDI-ARABIA
INDIA
TAIWAN
MYANMAR
PHILIPPINES
SRI LANKA
MALAYSIA
INDONESIA
Albanian
Armenian
Balto-Slavic
Celtic
Germanic
Greek
Indo-Iranian
Italic
Non-Indo-European

THE INDO-EUROPEAN FAMILY

ALBANIAN

This language forms a single branch of the Indo-European family, spoken by nearly 6 million people in Albania, and nearby parts of the Balkans, Greece, Turkey, and in Italy. Albanian has two main dialects, known as Gheg (in the north) and Tosk (in the south), but these contain many further dialect divisions, not all of which are mutually intelligible. The history of the language is obscure, and it is not possible to demonstrate a clear relationship with any other Indo-European group. This is partly because of the many loan words which have shaped the modern language, and partly because so few written remains of earlier times exist, dating only from the 15th century, largely on religious themes. An official alphabet was not introduced until 1909, using roman characters. Since the Second World War, the official language has been based on the Tosk dialect.

ANATOLIAN

A group of languages, now extinct, spoken from around 2000 BC in parts of present-day Turkey and Syria. The main Anatolian language is Hittite, shown to be Indo-European only as recently as 1915. Its written remains, consisting of tablets inscribed with cuneiform writing (p. 200), date from the 17th century BC. The earliest forms of Hittite ('Old Hittite') are the oldest Indo-European texts so far discovered. Most of the texts have religious themes, but they also contain a great deal of historical and social information. Other languages of the group are Palaic, Lydian, Lycian, and Luwian (represented in cuneiform and hieroglyphic systems). Also grouped under this heading are certain languages which do not belong to the Indo-European family (Hurrian, Urartian) or where the relationship is not certain (Phrygian).

ARMENIAN

This branch of Indo-European consists of a single language, spoken in many dialects by between 5 and 6 million people in the Armenian republic and Turkey, and (through emigration) in parts of the Middle East, Europe, and the United States. The spoken language may have been established soon after 1000 BC, but there was no written form until after the introduction of Christianity. Classical Armenian, or *Grabar*, is the language of the older literature, and the liturgical language of the Armenian church today. The oldest writings date from the 5th century, and the 38-letter alphabet, invented by St Mesrop, is still widely used. Modern literary Armenian exists in two standard varieties: East Armenian is the official language of Armenia; West Armenian is the dominant variety elsewhere. Because of the large numbers of loan words (see p. 332) which have come into the language, its basic Indo-European character is often obscured.

BALTO-SLAVIC

Baltic languages and Slavonic languages are often placed together as a single branch of Indo-European, because of their similarities, though there is some dispute over whether these constitute evidence of common origin rather than of more recent mutual influence. Taken together, these languages are spoken by about 300 million people, more than half of whom speak Russian.

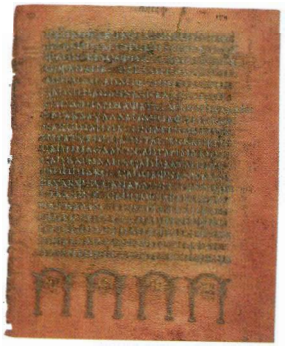
The main *Baltic* languages are Latvian (also known as Lettish) and Lithuanian, with written texts dating from the 14th century. There are around 4 million speakers in the Baltic area, with a further million abroad, mainly in the United States. Both languages have standard forms, and many dialects. Several other languages of this group are now extinct, though there are a few written remains of Old Prussian.

The *Slavonic* (or *Slavic*) languages are more numerous, and are usually divided into three groups: *South Slavonic*, found in Bulgaria, the countries of former Yugoslavia, and parts of Greece, includes Bulgarian, Macedonian, Serbian, Croatian, and Slovene; *West Slavonic*, found in the Czech and Slovak republics, Poland, and eastern areas of Germany, includes Czech, Slovak, Sorbian, and Polish; *East Slavonic*, found in the countries which replaced the USSR, includes Russian, Belorussian, and Ukrainian. The name *Lekhitic* is sometimes given to a group of West Slavonic languages originally spoken along the Baltic between the Vistula and the Oder, including Polish, Kashubian, Polabian (died out in the 18th century), and Slovincian. Each of the main Slavonic languages has an official status as a standard (pp. 38, 366); but there are numerous dialect differences within these groupings. Old Church Slavonic is evidenced in texts dating from the 9th century, and its later form (Church Slavonic) is still used as a liturgical language in the Eastern Orthodox Church. The distinctive Cyrillic alphabet (p. 204), attributed to Saints Cyril and Methodius in the late 9th century, is still used for writing Bulgarian, Serbian, Macedonian, and all the East Slavonic languages. In modified forms, it is also used for about 100 non-Slavonic minority languages of Russia.

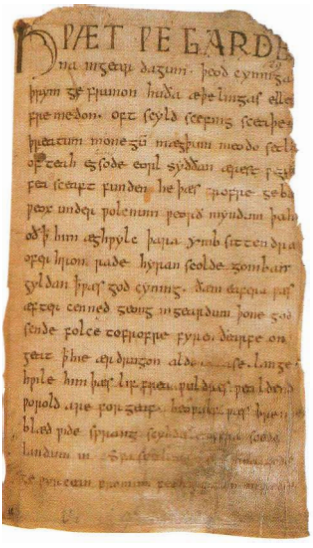
GERMANIC

The various branches of the Germanic family of languages derive from the migrations of the Germanic tribes who lived in northern Europe during the 1st millennium BC. Some Germanic words are recorded by Latin authors, and Scandinavian inscriptions in the runic alphabet (p. 205) are recorded from the 3rd century AD. The earliest main texts is the Gothic Bible of Bishop Ulfilas (or Wulfila), translated around AD 350, using an alphabet of his own devising (the Gothic alphabet: p. 188). Anglo-Saxon and Old High German are recorded from the 8th century, and the oldest forms of Scandinavian languages from the 12th century.

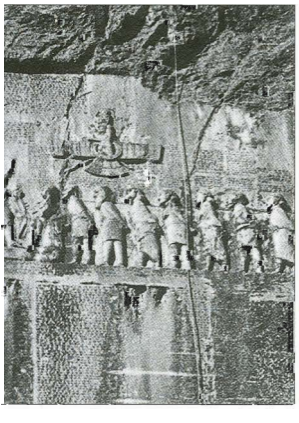
A bilingual tablet in Hittite and Luwian, dating from around 1400 BC, on which is written a ritual against the plague. The tablet was found in Hattusas, modern Bogazköy, Turkey. Inscriptions from this area provided some of the earliest evidence for the classification of Hittite as an Indo-European language.



A page from the Codex Argenteus, a 5th- or 6th-century copy of the Bible from Ulfilas; its name derives from the lettering, which is in gold and silver on a purple parchment. It is kept at Uppsala, in Sweden, not far from the Goths' homeland.



A page from the oldest epic poem in English, *Beowulf* The work was probably composed in the 8th century, but the only surviving manuscript dates from around AD 1000. It tells the story of a Scandinavian hero, Beowulf, who fights and kills a monster, Grendel, in Denmark. He is later made king of the Geats, in southern Sweden, where, as an old man, he kills a dragon, in a fight that leads to his own death.



This inscription, carved in an almost impossible position at the top of a steep cliff in Behistun (modern Bisitun), Iran, recounts the feats of King Darius the Great of Persia (522–486 BC). It is in three languages, Old Persian, Akkadian, and Elamite, and proved to be of particular value in deciphering the cuneiform writing system.

Germanic languages are used as a first language by over 550 million people, largely because of the worldwide role of English (§59). They are usually classified into three groups. *East Germanic* languages are all extinct, and only Gothic is preserved in manuscript to any extent – most recently, in a few words recorded in the Crimea in the 16th century. *North Germanic* includes the Scandinavian languages of Swedish and Danish (*East Scandinavian*), Norwegian, Icelandic, and Faeroese (*West Scandinavian*), and the older states of these languages, most notably the literary variety of Old Icelandic known as Old Norse – the language of the Icelandic sagas. *West Germanic* comprises English and Frisian (often grouped as Anglo-Frisian), and German, Yiddish, Netherlandic or Dutch (including local, Flemish dialects in Belgium), and Afrikaans (often grouped as Netherlandic–German). Dialect similarities often blur the distinctions suggested by these labels (§§8, 47).

GREEK

The branch of Indo-European consists of a single language, represented in many dialects, and attested from around the 14th century BC. The earliest evidence of the language is found in the inscriptions discovered at Knossos and other centres in Crete, written mainly on clay tablets in a syllabic script known as Linear B, and discovered to be Greek only as recently as 1952 (p. 203). This period of the language is referred to as Mycenaean Greek, to be distinguished from the later, classical Greek, dating from the 8th century BC, when texts came to be written in the Greek alphabet (p. 204) – notably the epic poems, *Iliad* and *Odyssey*. The great period of classical drama, history, philosophy, and poetry ended in the 4th century BC. A later variety of Greek, known as *koine* (or 'common') Greek, was spoken throughout the eastern Mediterranean from around the 4th century BC for nearly a thousand years. In its written form, it was the language of the New Testament (p. 388). The modern varieties of Greek, now spoken by over 11 million people in Greece, Cyprus, Turkey, the United States, and other localities, derive from this *koine* (p. 43).

INDO-IRANIAN

This branch of Indo-European comprises two large groups, known as Indo-Aryan (or Indic) and Iranian. There are over 200 Indo-Aryan languages, spoken by over 825 million people in the northern and central parts of the Indian subcontinent. They may be divided into several groups, on a broadly geographical basis: the *Midland* group mainly includes Hindu/Urdu (p. 286), the Bihari languages, and the Rajasthani languages; the *Eastern* group includes Assamese, Bengali, and Oriya; the *West* and *South-west* groups include Gujarati, Konkani, Maldivian, Marathi, and Sinhalese; and the *North-west* group includes Panjabi, Sindhi, Lahnda, the Dardic languages, and the Pahari languages. The Romani languages of the gypsies is also a member of Indo-Iranian. The early forms of Indo-Aryan, dating from around 1000 BC, are collectively referred to as Sanskrit – the language in which the Vedas, the oldest sacred texts, are written (p. 388). Later forms, the Prakrits, lasted 1,000 years, and were the medium of Buddhist and Jain literature.

During the same period, the Iranian languages were being spoken in an area centred on modern Afghanistan and Iran – especially Old Persian and Avestan (the sacred language of the Zoroastrians), both of which have texts dating from the 6th century BC. The group has over 70 languages spoken by over 75 million people, but many of these languages, and innumerable dialects, have not received a definite classification. Major languages include the closely related Persian (or Farsi) and Tadzhik, as well as Pashto, Ossetic, Kurdish, and Baluchi.

ITALIC

The main language of this family is Latin, the language of Rome and of its surrounding provinces, preserved in inscriptions from the 6th century BC, and most systematically in literature from the 3rd century BC. Other languages of the period include Faliscan, Oscan, Umbrian, and Venetic, spoken in and to the north-east of modern Italy. From the spoken, or 'vulgar' form of Latin, used throughout the Roman empire, developed the *Romance* languages – French, Spanish, Portuguese, Italian, Romanian, Sardinian, Occitan (in southern France), Rhaetian (various dialects in northern Italy and Switzerland), Galician (in north-west Spain), and Catalan (predominantly in north-east Spain). A Romance language known as Dalmatian, spoken along the Croatian coast, became extinct when its last known speaker died in 1898. But the main Romance languages have spread, as a result of colonialism, throughout the world, so that today around 650 million people speak a Romance language, or one of the creoles based on French, Spanish, or Portuguese (pp. 336–41).

TOCHARIAN

This language, now extinct, was spoken in the northern part of Chinese Turkistan during the 1st millennium AD. The first evidence of Tocharian was discovered only in the 1890s, in the form of various commercial and Buddhist religious documents, dating from around the 7th century, and on the basis of these discoveries two dialects were established – an eastern variety, from the Turfan region, which was labelled Tocharian A, and a western variety, from the Kucha region, which was labelled Tocharian B. The functions of these dialects, and the identity of their speakers, have been sources of controversy in comparative philology, as has the very name of the language (based on that of the Tochari people, who lived further east, and who were probably speakers of an Iranian language). But the status of Tocharian as an independent Indo-European language is not in doubt.

For Celtic, see pp. 304–5

53 · LANGUAGE ISOLATES

A language isolate is a language which has no known structural or historical relationship to any other language (p. 295). Most of the world's languages can be grouped into families using comparative linguistic techniques. But occasionally one encounters a language where resemblances to other languages are few or non-existent. Sometimes, the few points of contact are sufficient to motivate a tentative classification – thus some scholars place the Scots language Pictish within Celtic, the African languages Fur and Songhai within the Nilo-Saharan group, the Mexican language Huave within Penutian, and Tasmanian and Andamanese within Indo-Pacific. However, others see the differences as more important than the points of similarity, and list these languages as isolates.

Many languages have been classified as isolates simply because little is known about them, linguistically or historically. For example, preliminary research into South American Indian languages has brought to light several possible isolates, but further study may well indicate relationships with other languages – provided the cultures survive long enough for these studies to be carried out (p. 324). Examples are Callahuaya in Bolivia, and Aricapu, Baenna, Hixkaryana, Juma, and Natu in Brazil. Then, from a historical point of view, there are several languages of ancient Asia Minor which are known only from passing references to them in classical Greek literature, or occasional place names and inscriptions – examples include Bithynian, Cappadocian, Carian, Cataonian, Cilician, Gergito-Salymean, Hattic, Isaurian, Lycaonian, Myriandynian, Ördek-Burnu, Paphlagonian, Pisidian, Pontic, and Sidetic. It is unlikely that their affiliations will ever be known.

The diagram gives some information about several of the languages which have been proposed as isolates. It includes languages which remain undeciphered, languages where there is insufficient material available to establish a family relationship, and languages where, despite a great deal of data, the relationship is undetermined. Two of the best-known isolated languages, Korean and Japanese, are discussed on p. 308.

1) Iberian This language was spoken in parts of southern and eastern Spain, especially around the Ebro River, in pre-Roman times. It may formerly have been used throughout a much wider area of western Europe. It is known mainly through inscriptions on stones and artefacts of the period, few of which can be interpreted. Its 28-letter alphabet shows the influence of both the Greek and the Phoenician alphabets, but for the most part its history is unclear.

2) Basque is the only language remaining of those which must have been spoken in south-west Europe before the advent of the Indo-European invasions. Estimates of the number of speakers vary, from 500,000 to over 700,000. Most Basques live in a 4,000-square-mile area of northern Spain and south-west France, but many went into exile in the USA after the Spanish Civil War. Attempts have been made to show a relationship with Caucasian languages (p. 307), with North African languages, and also with Iberian, the now extinct language of many inscriptions found along the Mediterranean coasts; but none has been convincing. The written history of the language can be traced to Roman times, through various inscriptions. There is now intensive local concern to develop the language, and introduce it into primary education; but for many abroad, the language and culture are more associated with the violence of the political separatist movement, Euzkadi ta Azkatasuna (ETA). (*Euskara* is the Basque word for their language (p. 34).)

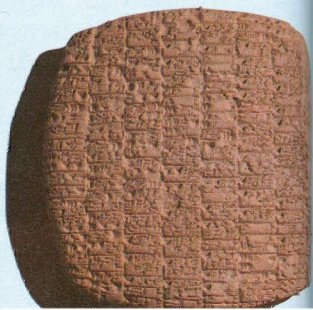
3) Etruscan The area of Tuscany in modern Italy is the site of the ancient country of Etruria, where the Etruscan civilization was at its height in the 6th century BC. The language is known from about 10,000 inscriptions, mainly short epitaphs and dedications, written in an alphabet probably derived from the Greek, and from which in due course came the Latin alphabet. The language may still have been spoken as late as the 4th century AD.

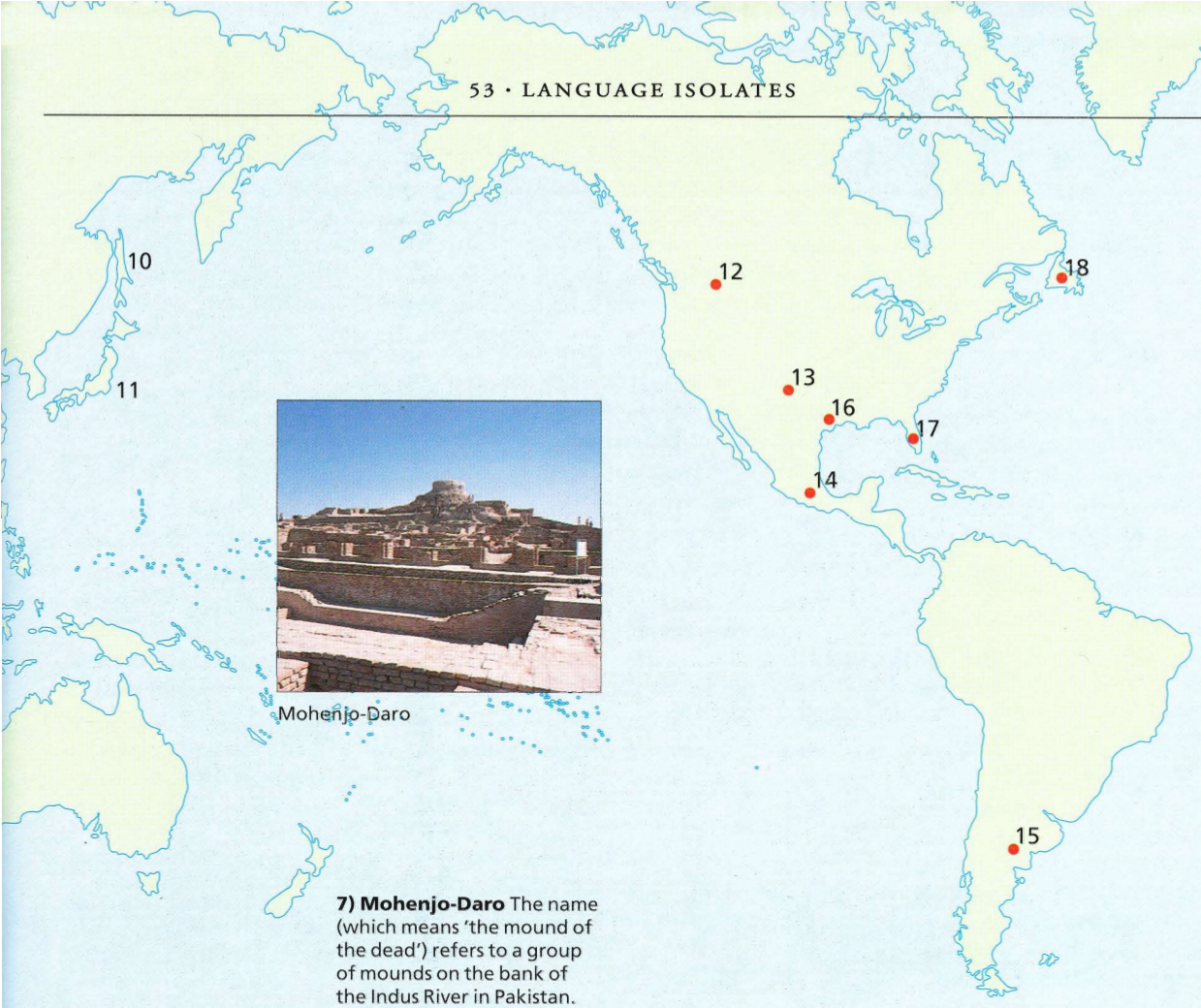
Only a few words of the language have been deciphered: no contemporary translations seem to have survived, and little progress has been made using philological methods, because Etruscan seems to bear no relationship to any other language. There is no extant literature or historical record of the civilization. Why this should be so remains one of the great unanswered questions of classical studies.

4) Linear A This is the name given to a Cretan script used in the middle of the 2nd millennium BC. It has still not been deciphered, and the language it represents is therefore not known, though some believe it to be Minoan (or Eteocretan). The name refers to the way the script is written in lines, probably from left to right – a contrast with previous hieroglyphic writing. The label 'A' distinguishes the script from Linear B, which was used to write Greek later in the same millennium (p. 303).

5) Sumerian This is the oldest known language to be preserved in written form. Inscriptions date from around 3100 BC, written in cuneiform script (p. 200). The existence of Sumerian was not recognized until cuneiform was deciphered in the 19th century, when it was realized that this language was quite different from others written in the same script. Sumerian was spoken in southern Mesopotamia (part of modern Iraq) until the 2nd millennium BC. It was then supplanted by a Semitic language (Akkadian) – though the written form of Sumerian continued to be used for nearly 2,000 years. There are many records of the language – business, legal, religious, administrative, and private texts and inscriptions. Literary work is preserved from the later period, in a range of forms including hymns, rituals, proverbs, and myths. Several dialect forms are known. Attempts have been made to relate the language to many other families, including Altaic and Dravidian, but none has been successful.

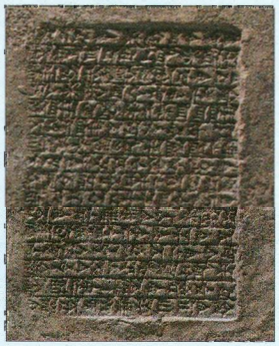
A Sumerian account listing the amount of grain paid to officials and servants of the temple of Baal, c. 2400 BC.





Mohenjo-Daro

6) Elamite This extinct language was spoken in the ancient country of Elam – an area now corresponding to Khuzistan in south-west Iran. The oldest writings are in the form of pictographic inscriptions from the 3rd millennium BC. Later writing is in cuneiform script. The language was still in use in the 1st millennium AD. A relationship with Dravidian has been proposed.


An Elamite inscription stamped on a baked clay brick, dating from the 12th century BC. It describes the rebuilding of the temple of the 'Great King' by King Shilhak-Inshushinak I.

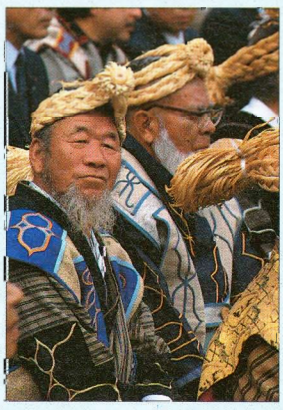
7) Mohenjo-Daro The name (which means 'the mound of the dead') refers to a group of mounds on the bank of the Indus River in Pakistan. Excavations at the site since the 1920s have brought to light the remains of a major city, dating from the 3rd millennium BC. The many finds contain evidence of a script, which so far is undeciphered.

8) Burushaski This language is spoken in north-west Kashmir, India, and in a small part of adjoining Pakistan, by over 50,000 people belonging to the Burusho tribe. It has no written form.

9) Nahali About 5,000 people speak this language in a small area in south-west Madhya Pradesh, in India. Some scholars have related the language to the Munda group of Austro-Asiatic (p. 311), but most view it as independent.

10) Gilyak This language is spoken by some 400 people in north-east Russia, on the island of Sakhalin and on parts of the mainland opposite. Gilyak (or Nivkhi, the name used by the people themselves) is often listed along with the neighbouring Palaeosiberian languages (p. 308), but proposed links with these and other languages of the area (especially Korean and the Altaic languages (p. 309)) have not been accepted.

11) Ainu About 16,000 Ainu tribespeople live in Hokkaido, Japan, and in the Sakhalin and Kuril Islands, but in recent years, the culture as well as the language has lost ground to the Japanese, and there are now probably no native speakers left. The traditional Ainu are unlike the Japanese in physical appearance, representing a different line of descent from the prehistoric (Jômonese) peoples of Japan.


An Ainu tribesman

12) Kutenai There are many spellings and names for this language – Kootenay, Cootenais, Skalzi, Arc-a-plat, and Flatbow are some of those recorded. It is spoken by a North American Indian tribe, mainly in south-east British Columbia and Alberta, but also in northern parts of Idaho, Washington, and Montana. Their numbers are decreasing–fewer than 200 in the 1980s. Some scholars have postulated relationships with other Amerindian languages (p. 322), but none of the proposals is generally accepted.

13) Keres Also known as Keresan, or Queres, this language is spoken in two main varieties by 8,000 speakers in the 1980s. Originally thought to be a member of the Hokan-Siouan family, it is now considered an isolate.

14) Tarasca This language was spoken by around 60,000 in the late 1960s, in parts of south-west-central Mexico. It goes under several names and spellings, including Tarascan, Porepecha, and Mechoacan. A relationship to other languages of the area has been proposed (under the general heading of Penutian, p. 322), but is unclear. In recent years, the number of speakers has been decreasing, with many Tarascan Indians becoming assimilated within the mixed European culture dominant in Mexico.

15) Het This South American Indian language, also known as Chechehet, became extinct at the end of the 18th century. It was spoken in Argentina, and is known from only a few words and place names.

16) Karankawa This language, also known as Clamcoets, was spoken by Indian tribes living along the Texan coastline in the 18th century. They seem to have died out by the mid-19th century, with the influx of white settlers into the area.

17) Calusa An extinct tribe of American Indians who lived in the south-west part of Florida until the end of the 18th century, and perhaps later. Many families emigrated to Cuba, to escape from the invasions of other tribes, and, ultimately, the British.

18) Beothuk This language, spoken by an Indian tribe on the island of Newfoundland, is now extinct. Its last known speaker died in 1829. Some scholars have argued that it should be classified as an Algonquian language, but the opinion is controversial. The Beothuk rubbed red ochre on their bodies – a practice which may well be the reason for the European name 'Red Indians'.

GRAMMATICAL CHANGE

The most noticeable way in which grammatical systems change is known as *analogy*. In this process, irregular grammatical patterns are changed in accordance with the regular patterns which already exist in the language.

A well-studied case is the verb system in the history of English. Several of the irregular verbs of Anglo-Saxon have fallen under the influence of the regular verbs in the past 1,000 years. For example, *helpan* (help) had *healp* as a past tense and *holpen* as a past participle; but by the 14th century, the verb had become regular, using the normal *-ed* ending – *helped*. During the early Middle English period, over 40 other verbs (including *walk, climb, burn,* and *step*) were influenced in the same way. Social factors, such as the development of the standard language, and the growth of printing, slowed the change down, so that present-day English still has many irregular verbs. But the force of analogy can still be heard, when people use non-standard forms (such as *knowed*), or when children, learning the language, experiment with such forms as *goned*. The tension between regular and irregular forms is also illustrated by problems of modern usage, such as the choice of *strove* vs *strived, chid* vs *chided*, or *sown* vs *sowed*.

Analogy does not operate only in word forms. Syntactic constructions can also be affected. In Anglo-Saxon, for example, the Subject–Verb–Object pattern applied only to main clauses; in subordinate clauses, the object preceded the verb. In Modern English, both clause types show the same order (§14).

Analogy does not create new grammatical patterns: it simply extends the range of a pattern which already exists in the language. Other processes of change have a more radical role, creating new patterns and eliminating old ones. For example, in Latin, the relationship between subject and object was shown by inflectional endings, and the order of the elements was not important; but in the modern Romance languages, these relations are expressed by word order. In early Indo-European, there were three grammatical genders for nouns – masculine, feminine, and neuter; these have been retained in modern German and Greek, but are reduced to two in modern Swedish (common vs neuter) and French (masculine vs feminine), and have been completely lost in modern English.

SEMANTIC CHANGE

This is perhaps the most obvious area of linguistic change, and the one which many people find the most fascinating. Semantic change is profoundly connected with the life, literature, and culture of a community. Innumerable examples can be found in the pages of old books, or simply by careful watching and listening to everyday usage. But plotting the history of the changes in the form, meaning, and use of words and morphemes is difficult work, because the evidence is often lacking.

To find out about lexical history, or *etymology*, the best source of information is a dictionary which has been written on historical principles, such as the *Oxford English Dictionary*. Many languages also have specialized etymological dictionaries.

New words and old

The two most obvious factors in semantic change are the arrival of new words and the loss of old ones. In most languages, the vast majority of new words are in fact *borrowings* from other languages – though this term is not a very appropriate one, as new words are not given back at a later stage! Borrowing proceeds in all directions. *Weekend* and *parking* have been borrowed by French from English; *chic* and *savoir-faire* have been borrowed by English from French. Some languages have borrowed so extensively that native words are in a minority.

A special type of borrowing is known as a *loan translation* or *calque*. In this process, a word is not borrowed whole, but its parts are translated separately and a new word formed – as when German produced the equivalent of English *telephone* in *Fernsprecher* (literally, *fern* 'distant' + *sprecher* 'speaker').

When a word or sense ceases to be used, it is said to be *obsolescent* or *obsolete*. This often happens because an object or concept is no longer of value to a community (other than to the historian or literary scholar); but a word or sense may become obsolescent if it develops unpleasant associations, or is replaced by another word which is felt to be more modern. *Wight* (person), *leman* (sweetheart), and *hie* (hasten), are examples from Elizabethan English which are now no longer used; *humour* (= 'temperament') and *conceit* (= 'idea') illustrate obsolete senses from the same period.

SOME TYPES OF SEMANTIC CHANGE

Extension. A word widens its meaning.
Example: In Latin, *virtue* was a male quality (cf. *vir* 'man'); today, it applies to both sexes.

Narrowing. A word becomes more specialized in meaning.
Example: In Old English, *mete* referred to food in general (a sense which is retained in *sweetmeat*); today, it refers to only one kind of food.

Shift. A word moves from one set of circumstances to another.
Example: *Navigator* once applied only to ships, but it now applies to planes, and even to cars.

Figurative use. A shift in meaning based on an analogy or likeness between things.
Example: *Crane*, a bird with a long neck, has led to the use of *crane* as a piece of equipment for lifting weights.

Amelioration. A word loses an original sense of disapproval.
Example: *Mischievous* has lost its strong sense of 'disastrous', and now means the milder 'playfully annoying'.

Pejoration. A word develops a sense of disapproval.
Example: *Notorious* once meant 'widely known', and now means 'widely and unfavourably known'.

SOME SURPRISING ETYMOLOGIES

The words in the left-hand column once had the meaning given on the right.

treacle	← wild animal
villain	← farm labourer
taxation	← fault finding
bonnet	← a man's hat
furniture	← equipment
pretty	← ingenious
cheater	← rent collector
naughty	← worth nothing
vulgar	← ordinary
sly	← wise
publican	← public servant
orchard	← garden (without fruit trees)

SOME SOURCES OF ENGLISH WORDS

ballot	Italian
banshee	Scots Gaelic
chow mein	Chinese
garage	French
gong	Javanese
goulash	Hungarian
junta	Spanish
kiosk	Turkish
llama	Quechua
marmalade	Portuguese
robot	Czech
schmaltz	Yiddish
slim	Dutch
sofa	Arabic
tomato	Nahuatl
tycoon	Japanese
veranda	Hindi
window	Old Icelandic
yen (= 'desire')	Chinese

BOYFRIENDS AND GIRLFRIENDS

Whether a language will borrow a word whole, or translate its parts, is never predictable. As the words *girlfriend* and *boyfriend* spread from the west to the east, they were handled differently. The Chinese loan-translated the words as *nan pengyou* (male friend) and *nü pengyou* (female friend). The Japanese, however, borrowed the words as wholes, adapting them to their sound system: the result was *bōifurendo* and *gārufurendo*.

55 · PIDGINS AND CREOLES

PIDGIN LANGUAGES

A pidgin is a system of communication which has grown up among people who do not share a common language, but who want to talk to each other, for trading or other reasons. Pidgins have been variously called 'makeshift', 'marginal', or 'mixed' languages. They have a limited vocabulary, a reduced grammatical structure, and a much narrower range of functions, compared to the languages which gave rise to them. They are the native language of no-one, but they are nonetheless a main means of communication for millions of people, and a major focus of interest to those who study the way languages change.

It is essential to avoid the stereotype of a pidgin language, as perpetrated over the years in generations of children's comics and films. The 'Me Tarzan, you Jane' image is far from the reality. A pidgin is not a language which has broken down; nor is it the result of baby talk, laziness, corruption, primitive thought processes, or mental deficiency. On the contrary: pidgins are demonstrably creative adaptations of natural languages, with a structure and rules of their own. Along with creoles (p. 338), they are evidence of a fundamental process of linguistic change, as languages come into contact with each other, producing new varieties whose structures and uses contract and expand. They provide the clearest evidence of language being created and shaped by society for its own ends, as people adapt to new social circumstances. This emphasis on processes of change is reflected in the terms *pidginization* and *creolization.*

Most pidgins are based on European languages – English, French, Spanish, Dutch, and Portuguese – reflecting the history of colonialism. However, this observation may be the result only of our ignorance of the languages used in parts of Africa, South America, or South-east Asia, where situations of language contact are frequent. One of the best-known non-European pidgins is Chinook Jargon, once used for trading by American Indians in north-west USA. Another is Sango, a pidginized variety of Ngbandi, spoken widely in west-central Africa.

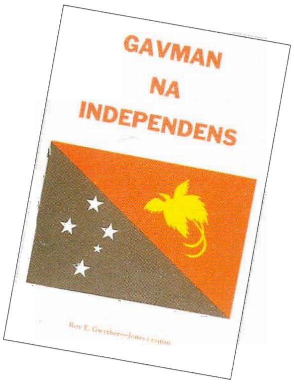
Because of their limited function, pidgin languages usually do not last for very long – sometimes for only a few years, and rarely for more than a century. They die when the original reason for communication diminishes or disappears, as communities move apart, or one community learns the language of the other. (Alternatively, the pidgin may develop into a creole.) The pidgin French which was used in Vietnam all but disappeared when the French left; similarly, the pidgin English which appeared during the American Vietnam campaign virtually disappeared as soon as the war was over. But there are exceptions. The pidgin known as Mediterranean Lingua Franca, or Sabir, began in the middle ages and lasted until the 20th century.

Some pidgins have become so useful as a means of communication between languages that they have developed a more formal role, as regular auxiliary languages. They may even be given official status by a community, as lingua francas. These cases are known as 'expanded pidgins', because of the way in which they have added extra forms to cope with the needs of their users, and have come to be used in a much wider range of situations than previously. In time, these languages may come to be used on the radio, in the press, and may even develop a literature of their own. Some of the most widely used expanded pidgins are Krio (in Sierra Leone), Nigerian Pidgin English, and Bislama (in Vanuatu). In Papua New Guinea, the local pidgin (Tok Pisin) is the most widely used language in the country.

POSSIBLE ORIGINS OF THE WORD *PIDGIN*

All of the following have been suggested as sources for the word *pidgin*, which is first attested in print in 1850:

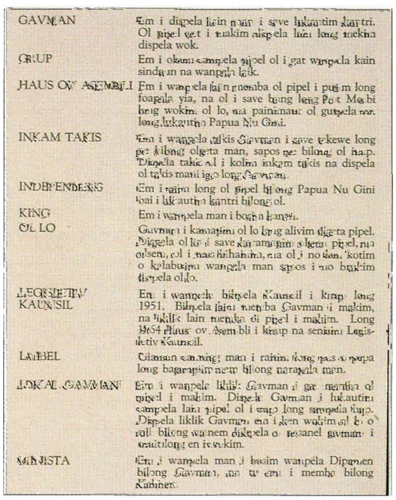
- A Chinese mispronunciation of the English word *business.*
- The Portuguese word *ocupação* (business).
- The Hebrew word *pidjom* (barter).
- A Yayo word *pidians* meaning *people.*
- Portuguese *pequeno* (little, child)–cf. 'baby-talk'.
- English *pigeon*–suitable for carrying simple messages.



An extract from a glossary of political terms listed in a Tok Pisin booklet on government and independence. It was produced by the Political Education Committee in Port Moresby, Papua New Guinea, in August 1972.

Right: The front page of *Wantok* ('Friend'), a Papua New Guinea weekly newspaper written entirely in pidgin (Tok Pisin) (with an English sports section).

Right, below: A street poster from Freetown, Sierra Leone, written in Krio: 'Electricity has no legs: it's Kabel-metal cable that carries it.'



GAVMAN	Em i dispela lain man i save lukautim kantri. Ol pipel yet i makim dispela lain long mekim dispela wok.
GRUP	Em i olsem sampela pipel ol i gat wanpela kain sindaun na wanpela laik.
HAUS OV ASEMBLI	Em i wanpela lain memba ol pipel i putim long foapela yia, na ol i save bung long Pot Mosbi bing wokim ol lo, na painimaut ol gutpela rot long lukautim Papua Nu Gini.
INKAM TAKIS	Em i wanpela takis Gavman i save tekewe long pe bilong olgeta man, sapos pe bilong ol inap. Dispela takis ol i kolim inkam takis na dispela ol takis mani igo long Gavman.
INDEPENDENS	Em i taim long ol pipel bilong Papua Nu Gini bai i lukautim kantri bilong ol.
KING	Em i wanpela man i bosim kantri.
OL LO	Gavman i kamapim ol lo long alivim olgeta pipel. Dispela ol lo i save karamapim olgeta pipel, na olsem, ol i mas bihainim, na ol i no ken kotim o kalabusim wanpela man sapos i no brukim dispela ol lo.
LEGISLETIV KAUNSIL	Em i wanpela bikpela Kaunsil i kirap long 1951. Bikpela lain memba Gavman i makim, na liklik lain memba ol pipel i makim. Long 1964 Haus ov Asembli i kirap na senisim Legisletiv Kaunsil.
LAIBEL	Giaman sampela man i raitim long pas o pepa long bagarapim nem bilong narapela man.
LOKAL GAVMAN	Em i wanpela liklik Gavman i gat memba ol pipel i makim. Dispela Gavman i lukautim sampela lain pipel ol i stap long sampela hap. Dispela liklik Gavman em i ken wokim ol lo o rul bilong wanem bikpela o nesanel gavman i tantilong en i wokim.
MINISTA	Em i wanpela man i bosim wanpela Dipatmen bilong Gavman, na tu em i memba bilong Kabinet.

GOVERNMENT This is a group [line] of people who look after [look out for] a country. All the people elect this group in order to do this work.
GROUP It is a number of people who share the same activities [one kind sit down] and the same interests.
HOUSE OF ASSEMBLY It is a group of members that the people put in every four years, and they regularly meet [know how to meet] in Port Moresby to make [work] the laws and to find out the best ways [the good roads] of looking after Papua New Guinea.
INKAM TAKIS It is a tax that the government takes from the pay of every person whose pay is of a certain level. This tax is called income tax and the money goes to the government.
INDEPENDENS It is the time when the people of Papua New Guinea will look after their own country.
KING He is a man who rules a country.

WANTOK

NIUSPEPA BILONG YUMI OL PAPUA NIUGINI STRET - 26 YIA NAU

2 pes | Namba 1,138 | Wik i stat long Fonde, Epril 18, 1996 | 50 toea

Robinson
rusoe muvi
bai kamaut
long Ogas

William Takaku
ithan) wantaim
erce Brosnan long
spela muvi nem
ong em 'Robinson
usoe.' Brosnan em
npela ekta husat i
ve ektim dispela
mes Bond muvi
oldeneye. Takuku na
osnan i wokim dis-
la muvi long
dang long 1994.
uvi ya i soim wei
ong ol waitman na
i bilong ol blek
n na long wanem
i tupela i ken
map long wanpela
gting long traim
ng save gut long
l na pasin bilong
pela yet. Dispela
uvi nau bai kam aut
ng 23 Ogus long
spela yia. LUKIM
TORI LONG PES 16.
Poto: William Takaku.

KANAGE — EM NA TASOL!

Belo na Kanage wantaim sampela wanwok bilong em i go baim buai long tupela meri Manus long 3 Mail insait long Mosbi siti na wok long kaikai buai na paitim tokpisin wantaim tupela meri ya i stap.

Ol i tokpisin i stap na tripela yangpela meri i wokabaut i kam na abrusim ol na i go. Em nau wanpela bilong dispela tupela meri Manus ya i kirap na tokim Kanage wantaim ol wanwok bilong em: Ol olsem yet, olgeta belo ol i save painim kam i go olsem.

Wanpela wanwok bilong Kanage kirap na askim dispela meri Manus: Ol i save painim wanem samting?

Arapela meri Manus ya kirap na tokim dispela wanwok bilong Kanage: Olsem nau tasol mama karim yu na yu no save na askim.

Wanwok bilong Kanage kirap na tokim meri wansolwara bilong em: Nogat wanpela samting ol i save kolim samting...olgeta samting i gat nem ya.

Meri kramsel ya lukluk long wanwok bilong Kanage na tokim em: Ol i painim sosis na hot dok.

Kanage sanap isi tru na kaikai buai bilong em pinis na laitim wanpela smok na putim long maus. Na em i pulim strongpela win pinis na tok: Kain pasin bilong dabolim dabolim na populesen bilong Papua Niugini i wok long bikpela hariap.

Joke Boss
MOSBI

Pablik Solisita laik Ombusu i mas fri

ERONICA HATUTASI i raitim

Kot bin givim 'death penalty' long em

YA bilong Charles Ombusu,
ajo Koeget bilong Pablik
isita i laikim kot long larim
busu i go olsem fri man. Dispela
bikos i nogat witnes long em i
n man na bagarapim meri.
Ombusu wantaim 21 krismas
ng Oro provins em i namba
n man long Papua Niugini
at Suprim kot i bin laik kilim em
g sas bilong kilim dai man na tu
garapim pikinini meri bilong
n ya long las yia.
oya Koeget i laikim Ombusu
g go fri bikos suprim kot i no
witnes long kotim Ombusu
long mekim ol dispela pasin.

Narapela samting tu em Mista Koeget i tok planti ripot na nius i bin kamap pinis long redio, telivisen na ol niuspepa long dispela samting. Na ol pablik i ting Ombusu tasol i bin wokim dispela rong. Olsem na dispela i no gutpela long pulim kot long dispela samting i go yet. Insait long ripot bilong em, em i askim kot long stapim kot long dispela mak nau bikos apil bilong em i orait pinis.

Ombusu i bin apil egensim olpela disisen bilong Nesenel Kot long kilim em bikos em (Ombusu) i kilim man na tu kisim 5-pela krismas long bagarapim meri.

Suprim Kot i bin oraitim dispela apil bilong Ombusu bikos Nesenel Kot i bin sasim Ombusu long tupela sas wantaim insait long wanpela kot. Suprim Kot i ting kot bilong kilim man i mas kamap long narapela taim na kot bilong bagarapim meri i mas kamap long narapela taim.

Pablik Prosekyuta, Panuel Mogish, long wankain taim tu i strong olsem kot bilong Ombusu i mas kamap gen long Nesenel Kot gen.

Na dispela i ken mekim ol i harim wan wan sas na i no tupela sas wantaim olsem pastaim.

Em i tokim Suprim kot jas long taim bilong harim trail kot bilong Ombuso olsem kot i bin asua long wanem em i no bin bihainim rot bilong brukim tupela sas na tu painim aut as tru long Ombusu i rong o nogat. Olsem na kot i mas kamap gen long harim dispela samting insait long nesenel kot.

Ol narapela poin long sabmisen bilong pablik prosekyuta long dispela kot em:

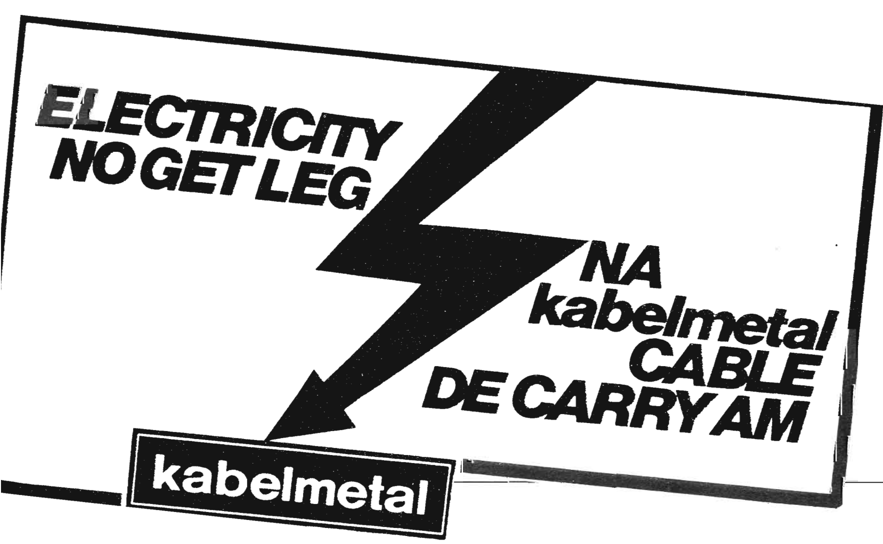
• Prosekyuta i tokaut long wanem tupela sas em bai i go het wantaim

I go moa long pes 2.

INSAIT

4-PES RAGBI LIG NIUS

TOYOTA LAND CRUISER
SAPOS ROT IBAGARAP NA YU PAINIM HAT LONG IGO...
GO WANTAIM 4WD LAND CRUISER!
LUKIM MIPELA NAU LONG ELA MOTORS
ISTAP OLGETA HAP LONG PAPUA NIUGINI!
TOYOTA
Ela Motors



SHAKESPEARE IN PIDGIN

The range of pidgin English is well illustrated by the translations which have been made of such works as the Bible and Shakespeare. Here is an extract from *Julius Caesar* (Act III, Scene 2), translated into Krio pidgin and Tok Pisin.

Friends, Romans, countrymen, lend me your ears;
I come to bury Caesar, not to praise him.
The evil that men do lives after them;
The good is oft interred with their bones;
So let it be with Caesar. The noble Brutus
Hath told you Caesar was ambitious.
If it were so, it was a grievous fault;
And grievously hath Caesar answer'd it.
Here, under leave of Brutus and the rest –
For Brutus is an honourable man;
So are they all, all honourable men –
Come I to speak in Caesar's funeral.

Krio

Padi dem, kohntri, una ohl wey dey na Rom. Meyk una ohl kak una yeys. A kam ber Siza, a noh kam preyz am. Dem kin memba bad wey pohsin kin du lohng tem afta di pohsin kin dohn dai. Boht plenti tem di gud wey pohsin du kin ber wit im bon dem. Meyk i bi so wit Siza. Bra Brutohs dohn tel una sey Siza na bin man wey want pas mak. It i tohk tru, na badbad ting dis ya. En Siza dohn get im bad pey foh dat. A tayk pamishohn frohm Bra Brutohs dem foh kam tohk na Bra Siza im berin. En Bra Brutohs na ohnareybul O! Dem ohda wan sef na ohnareybul.
(From T. Decker, 1965, p. 74.)

Tok Pisin

Pren, man bolong Rom, Wantok, harim nau. Mi kam tasol long plantim Kaesar. Mi noken beiten longen. Sopos sampela wok bolong wampela man i stret; sampela i no stret; na man i dai; ol i wallis long wok i no stret tasol. Gutpela wok bolongen i slip; i lus nating long giraun wantaim long Kalopa. Fesin bolong yumi man. Maski Kaesar tu, gutpela wok i slip.

Brutus ia tokim yu long Kaesar i mangal. Sopos olosem, bikpela pekato tru. Tasol Kaesar Kalopa bekim pinis long virua belongen. Tru Brutus, na ol pren bolongen, gutpela man. I orait. Ol i gipim mi orait long mi toktok sore hia long Kaesar.
(From J. J. Murphy, 1966, pp. 19–20.)

A page from a New Guinea road safety handbook *Rot Sefti Long Niugini* (1972), with instructions in English (top), Tok Pisin (middle), and Hiri Motu (bottom).

If you have an accident, get the other driver's number, if possible his name and address and report it to the police. Do not fight him or abuse him.

Sapos yu kisim bagarap kisim namba bilong narapela draiva, sapos yu ken, kisim naim bilong em na adres tu, na tokim polis longen. Noken paitem em o tok nogut long em.

Bema kerere davaria neganai, taraka o motuka taria tauna ena ladana oi abia bona ena noho o gaukara gabuna danu abia. Taraka o motuka ena naba danu abia vadaeni Police hamaoroa. Oi heai bona hereva dika lasi.

CREOLE LANGUAGES

A creole is a pidgin language which has become the mother tongue of a community – a definition which emphasizes that pidgins and creoles are two stages in a single process of linguistic development. First, within a community, increasing numbers of people begin to use pidgin as their principal means of communication. As a consequence, their children hear it more than any other language, and gradually it takes on the status of a mother tongue for them. Within a generation or two, native language use becomes consolidated and widespread. The result is a creole, or 'creolized' language.

The switch from pidgin to creole involves a major expansion in the structural linguistic resources available – especially in vocabulary, grammar, and style, which now have to cope with the everyday demands made upon a mother tongue by its speakers. There is also a highly significant shift in the overall patterns of language use found in the community. Pidgins are by their nature auxiliary languages (§58), learned alongside vernacular languages which are much more developed in structure and use. Creoles, by contrast, are vernaculars in their own right. When a creole language develops, it is usually at the expense of other languages spoken in the area. But then it too can come under attack.

The main source of conflict is likely to be with the standard form of the language from which it derives, and with which it usually co-exists. The standard language has the status which comes with social prestige, education, and wealth; the creole has no such status, its roots lying in a history of subservience and slavery. Inevitably, creole speakers find themselves under great pressure to change their speech in the direction of the standard – a process known as *decreolization.*

One consequence of this is the emergence of a continuum of several varieties of creole speech, at varying degrees of linguistic 'distance' from the standard – what has been called the 'post-creole continuum'. Another consequence is an aggressive reaction against the standard language on the part of creole speakers, who assert the superior status of their creole, and the need to recognize the ethnic identity of their community. Such a reaction can lead to a marked change in speech habits, as the speakers focus on what they see to be the 'pure' form of creole – a process known as *hypercreolization.* This whole movement, from creolization to decreolization to hypercreolization, can be seen at work in the recent history of African-American English in the USA.

The term *creole* comes from Portuguese *crioulo,* and originally meant a person of European descent who had been born and brought up in a colonial territory. Later, it came to be applied to other people who were native to these areas, and then to the kind of language they spoke. Creoles are now usually classified as 'English based', 'French based', and so on – though the genetic relationship of a creole to its dominant linguistic ancestor is never straightforward, as the creole may display the influences of several contact languages in its sounds, vocabulary, and structure.

Today, the study of creole languages, and of the pidgins which gave rise to them, attracts considerable interest among linguists and social historians. To the former, the cycle of linguistic reduction and expansion which they demonstrate, within such a short timescale, provides fascinating evidence of the nature of language change. To the latter, their development is seen to reflect the process of exploration, trade, and conquest which has played such a major part in European history over the past 400 years.

Members of Radio Nous Mêmes ('Ourselves Radio'), a local radio station in the Maroni River area of French Guiana.

CREOLES COMPARED

The similarities between European-based creoles are striking, as can be seen from this comparison of the verb phrase in the French-based creole of Guiana and the English-based Krio of Sierra Leone (after L. Todd, 1984, p. 24).

French	Guianese Créole	Krio	English
Mangez	Mãʒe	Chɔp	Eat
J'ai mangé	Mo mãʒe	A chɔp	I ate
Il/Elle a mangé	Li mãʒe	I chɔp	He/She ate
Je mange/Je suis en train de manger	Mo ka mãʒe	A de chɔp	I am eating
J'avais mangé	Mo te mãʒe	A bin chɔp	I ate/had eaten
Je mangeais	Mo te ka mãʒe	A bin de chɔp	I was eating
Je mangerai	Mo ke mãʒe	A go chɔp	I shall eat
Il/Elle est plus grand que vous	Li gros pas u	I big pas yu	He/She/It is bigger than you

WHERE DO PIDGINS AND CREOLES COME FROM?

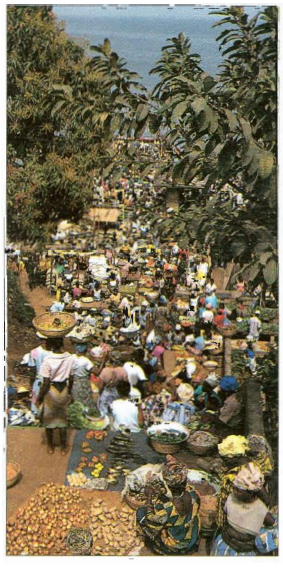
The world's pidgins and creoles display many obvious differences in sounds, grammar, and vocabulary, but they have a remarkable amount in common. Two opposed theories have attempted to explain these differences.

MANY SOURCES?

A long-standing view is that every creole is a unique, independent development, the product of a fortuitous contact between two languages. On the surface, this 'polygenetic' view is quite plausible. It seems unlikely that the pidgins which developed in South-east Asia should have anything in common with those which developed in the Caribbean. And it is a general experience that these varieties come into use in an apparently spontaneous way – as any tourist knows who has faced a souvenir seller. Would not the restricted features of the contact situations (such as the basic sentence patterns and vocabulary needed in order to trade) be enough to explain the linguistic similarities around the world?

The view is tempting, but there are several grounds for criticism. In particular, it does not explain the *extent* of the similarities between these varieties. Common features such as the reduction of noun and pronoun inflec-

Saturday Market, Freetown, Sierra Leone.



tions, the use of particles to replace tenses, and the use of repeated forms to intensify adjectives and adverbs are too great to be the result of coincidence. Why, then, should the pidginized forms of French, Dutch, German, Italian, and other languages all display the same kind of modifications? Why, for example, should the English-based creoles of the Caribbean have so much in common with the Spanish-based creoles of the Philippines? How could uniformity come from such diversity?

ONE SOURCE?

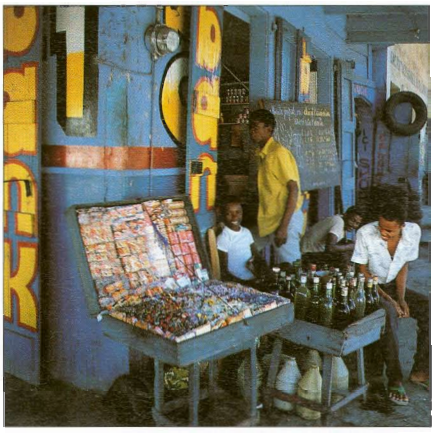
The opposite view argues that the similarities between the world's pidgins and creoles can be explained only by postulating that they had a common origin (i.e. are 'monogenetic'), notwithstanding the distance which exists between them. Moreover, a clear candidate for a 'proto'-language has been found – a 15th-century Portuguese pidgin, which may in turn have descended from the Mediterranean lingua franca known as Sabir (p. 340). The Portuguese are thought to have used this pidgin during their explorations in Africa, Asia, and the Americas. Later, it is argued, as other nations came to these areas, the simple grammar of this pidgin came to be retained, but the original Portuguese vocabulary was replaced by words taken from their own languages. This view is known as the *relexification* hypothesis.

There is a great deal of evidence to support the theory, deriving from historical accounts of the Portuguese explorations, and from modern analyses of the languages. For instance, every English-based pidgin and creole has a few Portuguese words, such as *savi* 'know', *pikin* 'child', and *palava* 'trouble'. In Saramaccan, an English-based creole of Suriname, 38% of the core vocabulary is from Portuguese. Early accounts of Chinese pidgin refer to a mixed dialect of English and Portuguese. And on general grounds, relexification of a single 'proto'-pidgin seems a more plausible hypothesis than one which insists on a radical parallel restructuring of several languages.

The shift in approach, implicit in the relexification theory, is fundamental: it is not the case that English, and the other languages, were 'creolized', but that an original (Portuguese) creole was 'Anglicized'. However, not all the facts can be explained in this way. Pitcairnese creole has no Portuguese influence, and yet has much in common with other varieties. What accounts for those similarities? Then there are several pidgins and creoles which have developed with little or no historical contact with European languages – Sango and Chinook, for instance. And there seem to be many structural differences between European and non-European pidgins and creoles, which the common origin hypothesis finds difficult to explain.

The evidence is mixed. Disentangling the structural similarities and differences between these varieties is a difficult task, and the evidence could be taken to support either a monogenetic or a polygenetic theory. Far more descriptive studies are needed before we rule out one view or the other.

Meanwhile, other theories have been proposed, in an attempt to explain these similarities and differences. Other forms of simplified speech have been noted, such as that used by children (§41), in telegrams and headlines, and in talking to foreigners (p. 377). It is possible that the processes underlying pidgins and creoles reflect certain basic preferences in human language (such as fixed word order, or the avoidance of inflections). In this connection, these languages provide fresh and intriguing evidence in the search for linguistic universals (§14).



Street scene in Port-au-Prince, Haiti.

PIDGINS COMPARED

Lexical similarities and differences between pidgins are clearly illustrated in this list of items collected by F. G. Cassidy in the 1960s, taken from the set of 'basic words' used in glottochronology (p. 333). The English element predominates in Tok Pisin and Chinese Pidgin; in Sango, the vast majority of the words are African; in Chinook, most words are from Chinook or other Amerindian languages (but note the influence of both French and English). French names for parts of the body have emerged in Sango and Chinook. Though there is no historical connection between the languages, note also the coincidences of thought which have produced the figurative phrases for *feather* (grass-of-bird (Tok Pisin), hair-of-bird (Sango), and leaf-of-bird (Chinook)), and the words for *heart* in Tok Pisin and Chinook, both of which stress the notion of heartbeat.

English	Tok Pisin	Chinese Pidgin	Sango	Chinook Jargon
bell	bɛl	bell	ngbéréná	tíntin
big	bɪgfɛlə	big	kótá	hyás
bird	pɪgɪn	bird(ee)	ndɛkɛ	kalákala
bite	kajkajɪm	bitee	tɛ	múckamuck
black	blækfɛlə	black	(zo)vɔkɔ́	klale
blood	blʊt	blood	méné	pilpil
cold	kilfɛlə	colo	dé	cole, tshis
come	kəm	li	ga	chahko
die	daj	dielo	kúi	mémaloost
dog	dɔg	doggee	mbo	kámooks
drink	drɪŋk	dlinkee, haw	yç	muckamuck
ear	ir	ear	mé	kwolánn
earth	grawn	glound	sése	illahie
eat	kajkaj	chowchow	kóbe, tɛ	múckamuck
fat	gris	fat, glease	mafuta	glease
feather	gras bɪlɔŋ pɪgɪn	fedder	kɔ́á tí ndɛkɛ	kalákala yaka túpso
fish	fiš	fishee	susu	pish
give	gɪvɪm	pay	fú	pótlatch
green	grinfɛlə	gleen, lu	vɔkɔ́ kété	pechúgh
hair	gras bɪlɔŋ hɛd	hair	kɔ́á	yákso
hand	hæn	hand, sho	mabɔ́kɔ	le mah
head	hɛd	headee	li	la tet
heart	klak	heart	coeur	túmtum
know	save	savvy	hínga	kumtuks
man	mæn	man	kɔ́lî	man
no	no	na	non	wake
nose	nos	peedza	hɔ̄	nose
one	wənfɛlə	one piecee	ɔ́kɔ́	ikt
small	lɪklɪk	likki	kété	ténas
sun	sən	sun	lá	sun, ótelagh
talk	tɔk	talkee	tɛnɛ	wáuwau
two	tufɛlə	two	óse	mokst
warm	hɔtfɛlə	warm	wá	waum

BEING BILINGUAL

Research into bilingualism usually distinguishes between large-scale analyses of multilingual societies ('societal' bilingualism, p. 362) and small scale analyses of the settings in which bilingual speakers interact ('individual' bilingualism). Several fundamental questions have to be dealt with under the latter heading – in particular, how bilingualism is to be identified and defined, and what its purpose is within the speech community. Both questions have 'obvious' answers, neither of which is adequate.

WHAT IS A BILINGUAL?

The obvious answer is: someone who speaks two languages. But this answer will not suffice. It does not allow for those who make irregular use of one or other language, or those who have not used the language at all for many years (so-called 'dormant' bilinguals). Nor does it allow for the many people who have developed a considerable skill in comprehending a foreign language, but who do not speak it; or those who have learned to read in another language, but who cannot speak or write it. It leaves unclear the relationship between different languages and different dialects, styles, or levels of the same language (as in the case of diglossia, p. 43). And above all, this definition says nothing about the level of proficiency that has to be attained before speakers can legitimately claim to be bilingual.

The notion of proficiency raises some very complex issues. Again, the 'obvious' answer is to say that people are bilingual when they achieve native-like fluency in each language. But this criterion is far too strong. People who have 'perfect' fluency in two languages do exist, but they are the exception, not the rule. The vast majority of bilinguals do not have an equal command of their two languages: one language is more fluent than the other, interferes with the other, imposes its accent on the other, or simply is the preferred language in certain situations. For example, a child of French/English parents went to school and university in France. She became a geography teacher, married a British doctor, and came to live in England, where she had her first child. In general conversation, she could cope with ease in either language; but she found herself unable to teach geography in English, and she was extremely reluctant to discuss baby care in French. In each case she knew the slang, jargon, and phrasing which is naturally assimilated when learning a new skill – but this was available in only one of her languages. Her linguistic competence certainly did not resemble that of monolingual teacher-mothers.

This situation seems to be typical. Studies of bilingual interaction have brought to light several differences in linguistic proficiency, both within and between individuals. Many bilinguals fail to achieve a native-like fluency in either language. Some achieve it in one (their 'preferred' or 'dominant' language), but not the other. For such reasons, scholars now tend to think of bilingual ability as a continuum: bilingual people will find themselves at different points on this continuum, with a minority approaching the theoretical ideal of perfect, balanced control of both languages, but most being some way from it, and some having very limited ability indeed. However, the notion is a difficult one to make precise, because so many different abilities are involved – in speaking, listening, reading, and writing, as well as in phonology, grammar, vocabulary, and pragmatics (Parts III–V).

WHY USE TWO LANGUAGES?

Here, the 'obvious' answer is: to communicate with people of different language backgrounds. And once again, the obvious answer will not account for the remarkable range of linguistic behaviour that can be observed in adult bilinguals. The 'easy' cases are those where a bilingual meets different monolingual people within a multilingual society, and changes from one language to the other in order to communicate with them. Somewhat more complex are cases where a bilingual chooses to use one language knowing that the listener would prefer the other (for example, electing to be tried in the language of a minority group, in order to embarrass the authorities). Here, language choice is a symbol of national identity.

But such bilingual/monolingual interactions and confrontations account for only a minority of cases. More often, in a multilingual society, bilinguals interact with other bilinguals, and opt to use their different languages in a complex network of interaction that proves extremely difficult to describe and explain. The choice of language will vary depending on the type of person addressed (e.g. members of the family, schoolmates, colleagues, superiors, friends, shopkeepers, officials, transport personnel, neighbours), and on the location or social setting (e.g. a family may vary their language use depending on whether they are at home, in the street, or in church; at the office, someone may talk to a colleague in language X, but over lunch talk to the same person using language Y). Even more complex, and not well understood, are the many cases when a bilingual talks to another bilingual with the same language background, and yet changes from one language to another in the course of the conversation – a phenomenon known variously as 'language mixing', 'language switching', or simply 'code switching'.

DORMANT LANGUAGES

There is no clear indication as to whether there is a limit to human multilingual ability. Cardinal Giuseppe Mezzofanti (1774–1849), librarian at the Vatican, is reputed to have been able to speak 50 languages (most with great fluency), to understand a further 20, and to translate 114. The Victorian diplomat Sir John Bowring (1792–1872) was said to have spoken 100 languages and read another 100. Unfortunately, there is no way of knowing exactly what proficiency level was achieved by these remarkable language learners.

It is in fact highly unusual to maintain proficiency in more than two or three languages at a time. Most multilinguals have a single dominant language, others being 'dormant' to varying degrees. The typical situation can be illustrated by a case study that was made in the field of aphasia (p. 272). It emerged that the person had learned seven languages during his life, but five had become dormant. His mother tongue had been Hungarian. At the age of 4, he moved to Poland, learned Polish, and stopped using Hungarian. When he was 6 he returned to Hungary, and had to relearn Hungarian. At the age of 10, he moved to Romania, using Romanian in school and Yiddish socially. Two years later he returned to Hungary, where in school he learned German, English, and Hebrew. This was followed by six years in Germany, during which time German became his dominant language. At 25, he moved to the U.S., where English became dominant. At the time of the study, only English and Hungarian were regularly used (his wife is Hungarian). The others were dormant, and in some cases almost forgotten. (L. Galloway, 1978.)

Cardinal Giueseppe Mezzofanti

Language switching

Switching between languages is extremely common and takes many forms. A long narrative may switch from one language to the other. Sentences may alternate. A sentence may begin in one language, and finish in another. Or phrases from both languages may succeed each other in apparently random order (though in fact grammatical constraints are frequently involved). Such behaviour can be explained only by postulating a range of linguistic or social factors such as the following.

- Speakers cannot express themselves adequately in one language, so switch to the other to make good the deficiency. This may trigger a speaker to continue in the other language for a while. An example from a Spanish/English study (G. Valdés Fallis, 1976): *Porque alli hay cashews. You don't like them?* (Because here are some cashews...'). This tends to happen a great deal when the speaker is upset, tired, or otherwise distracted.
- Switching to a minority language is very common as a means of expressing solidarity with a social group. The language change signals to the listener that the speaker is from a certain background; if the listener responds with a similar switch, a degree of rapport is established. The same switch may of course also be used to exclude other people, who do not know the language, from the group.
- The switch between languages can signal the speaker's attitude towards the listener – friendly, irritated, distant, ironic, jocular, and so on. Monolinguals can communicate these effects to some extent by varying the level of formality of their speech; bilinguals can do it by language switching. If two bilinguals normally talk to each other in language X, the choice of Y is bound to create a special effect. A common example is for a mother to tell her child to do something in one language, and then, if the child fails to obey, to switch to another language, thereby showing her stronger emphasis or displeasure.

These are but some of the sociolinguistic functions that language switching can perform. The phenomenon is evidently a complex and subtle one, with speakers usually being totally unaware of the extent to which they have been switching in a conversation. If interrupted, they may even be unable to say which language they were using in their last sentence. Monolinguals often dismiss or satirize language switching, using such pejorative labels as 'Franglais', 'Spanglish', or 'Tex-Mex'. Perhaps because of this kind of criticism, many bilingual people come to be very self-conscious about their switching, and try to avoid it in talking to strangers or on formal occasions. But in informal speech, it is a natural and powerful communicative feature of bilingual interaction, which presents linguists with one of their most intriguing analytical challenges.

BILINGUAL VERBAL STRATEGIES

Language switching is a major feature of this conversation between two native Americans of Mexican ancestry. *E* is a university teacher, who is working as a volunteer in a day care centre where *M* is a social worker. The Spanish passages are translated in parentheses.

E: What do you dream in?
M: I don't think I ever have any conversations in my dreams. I just dream. Ha. I don' hear people talking: I jus' see pictures.
E: Oh, they're old-fashioned, then. They're not talkies yet, huh?
M: They're old-fashioned. No, they're not talkies yet. No, I'm tryin' to think. Yeah, there too have been talkies. Different. In Spanish and English both. An' I wouldn't be too surprised if I even had some in Chinese. (*Laughter*) Yeah, Ed. Deveras. ('Really') (*M offers E a cigarette which is refused*) Tu no fumas, verdad? Yo tampoco. Deje de fumar ('You don't smoke, do you? I don't either. I stopped smoking') and I'm back to it again.

...

M: An' – an' – an' they tell me, 'How did you quit, Mary? I di'n' quit. I – I just stopped. I mean it wasn't an effort that I made. Que voy a dejar de fumar porque me hace daño o ('That I'm going to stop smoking because it's harmful to me, or') this or tha', uh-uh. It just – that – eh – I used to pull butts out of the – the wastepaper basket. Yeah (*Laughter*) I used to go look in the (*unclear speech*). Se me acababan los cigarros en la noche ('My cigarettes would run out on me at night'). I'd get desperate, y ahi voy al basurero a buscar, a sacar, you know? ('And there I go to the wastebasket to look for some, to get some, you know?') (*Laughter*) Ayer los (*unclear speech*) no había que no traia cigarros Camille, no traia Helen, no traia yo, el Sr. de Leon ('Yesterday the – there weren't any. Camille didn't have any, Helen, I, Mr. de Leon didn't have any') and I saw Dixie's bag crumpled up, so I figures she didn't have any, y ahi ando en los ceniceros buscando a ver onde estaba la – ('And there I am in the ashtrays looking to see where there was the –') I din' care whose they were.

The authors of this study point out that *M*'s language switching is not random. *M* is ambivalent about her smoking, and she signals this through her choice of language. Spanish sentences in this conversation reflect her embarrassment and personal involvement; English is used for more general or detached statements.
(J. Gumperz, 1970.)

BILINGUAL ACQUISITION

There is a widespread popular impression that the children of bilingual parents are linguistically at risk. It is said that their brains will not be able to cope, and that they will grow up 'semilingual', confused, or retarded. There is no justification for this pessimism, as is evident from the confident fluency displayed by millions of bilingual and trilingual children all over the world. By the time these children arrive in school, the vast majority have reached the same stage of linguistic development as have their monolingual peers.

But the process of learning two languages is not exactly the same as the process of learning one (Part VII). Three main stages of development have been noted:

1. The child builds up a list of words, as does a monolingual child, but the list contains words from both languages. It is rare for these words to be translation equivalents of each other.
2. When sentences begin to contain two or more elements, words from both languages are used within the same sentence, e.g. (from a 2-year-old German/English child) *ein* ('a') *big cow, from up in Himmel* ('sky'). The amount of mixing rapidly declines. In one study, at the beginning of the third year, nearly 30% of the sentences contained mixed vocabulary; by the end of the year, it was less than 5%.
3. As vocabulary grows in each language, translation equivalents develop. But the acquisition of separate sets of grammatical rules takes longer. For a while, a single system of rules seems to be used for both languages, until finally the two grammars diverge.

When bilingual children reach this stage, usually in the fourth year, they have become aware that the two languages are not the same. They typically use each language to the parent who speaks it, and not to the other. Indeed, if one parent uses the language of the other to the child, there may be quite a reaction. The child may be surprised, embarrassed, fail to understand, think it funny, or become upset. An extract from a recent bilingual-acquisition study illustrates this last reaction. Lisa (nearly 4 years old) has an Italian father and a German mother. The father uses a short German sentence to her, to which she replies:
Lisa: No, non puoi. ('No, you can't')
Father: Ich auch – spreche Deutsch. ('I also speak German.')
Lisa: No, tu non puoi! ('No, you cannot.')
(V. Volterra & T. Taeschner, 1978.)

Not surprisingly, it is at this age that children try to play their parents off against each other. One child would always switch into French when he saw his English father approach him purposefully at bedtime!

61 · LANGUAGE PLANNING

Language, sooner or later, proves to be a thorn in the flesh of all who govern, whether at national or local level. Different social groups wish to see their linguistic identities and interests maintained, and may actively – and often violently – campaign for recognition (§9). Governments have to react to these differences, officially or unofficially: they may wish to reconcile them, or try to eliminate them. With the pace of change increasing, and countries becoming more heterogeneous, cosmopolitan, and internationally aware, it is not possible to rely on the slow course of natural linguistic evolution to resolve the many pressures and conflicts that arise. Many governments, accordingly, try to solve their problems by engaging in conscious, principled 'language planning', or 'linguistic engineering'.

Language planning involves the creation and implementation of an official policy about how the languages and linguistic varieties of a country are to be used. Decisions of a fundamental nature may need to be made, especially in the developing countries. But planning issues are to be found in all countries, as people debate such topics as the place of minority languages, the role of an academy in safeguarding standards (§1), the influence of the media on usage (p. 396), the value of spelling reform (p. 217), the avoidance of sexist language (p. 46), the modernization of religious language (p. 388), the need for plain English (p. 382), stylistic standards in publishing (p. 392), and the maintenance of oracy and literacy levels in school (§44).

Language planning is carried out by a variety of government departments and agencies, academies, committees, popular societies, and individuals. Activities range from the political and judicial, at one extreme, to the unofficial and illegal, at the other. Popular attitudes towards planning proposals include everything from complete support, through partial approval, general indifference, and mild antagonism, to total antipathy. Historical, political, economic, religious, educational, judicial, and social factors all have to be disentangled. As a consequence, it is hardly surprising that those who study this subject have not yet reached the stage when they can explain why some planning proposals succeed, whereas others fail. The field of language planning, which dates only from the 1960s, is still largely at the stage of descriptive enquiry, with a continuing need for detailed case studies of the widely differing situations in individual countries; few general theoretical principles have been proposed. However, the area continues to attract a great deal of interest, for both applied and theoretical reasons.

Most obviously, its findings and analyses may assist those (politicians, educators, lawyers, etc.) whose responsibility it is to make decisions about the development of languages in society, many of whom have no specialized knowledge of linguistic issues. But it also presents a fresh perspective for our understanding of linguistic change (§54). Many linguists have held the view that language change is a natural, spontaneous phenomenon, the result of underlying social and/or linguistic forces that it is impossible or undesirable to tamper with. We should 'leave our language alone' (p. 180). However, language planning studies have shown that it is quite possible for social groups to alter the course of a language, and that the question of desirability is a highly controversial one. It is still unclear how far languages can be permanently influenced by social manipulation, but there is now strong evidence that such factors must be taken seriously when considering historical linguistic matters.

TWO KINDS OF LANGUAGE PLANNING

Many analysts recognize a binary classification of language-planning activities, based on whether the changes affect primarily linguistic structure or linguistic use (§13). In *corpus planning*, the changes are introduced into the structure (or 'corpus') of a language/variety – as when changes are proposed in spelling, pronunciation, grammar, or vocabulary. In *status planning*, changes are proposed in the way a language/variety is to be used in society (thus altering its status) – as when it is permitted for the first time in law courts or in official publications. The distinction is not clear-cut, because not all kinds of planning activity can be neatly classified in this way, but it is widely encountered in language planning research.

PLANNING IN PRACTICE

Selecting the norm
If several languages are spoken within a country, it is usually necessary to choose a single language as a norm for official, educational, and other purposes. It may prove possible to use one of the indigenous languages, but intergroup rivalry may make it necessary to introduce a language from elsewhere as a lingua franca (e.g. Hindi in India, English in Ghana), in which case the relative merits of these languages will need to be debated. In addition, it may be necessary to choose a particular variety of a language (Part II), or to construct a new variety, taking into account such factors as formality, social class, regional dialect, and previous literary use.

Codification
If an indigenous language is chosen, it will need to be developed to meet the demands placed upon it as a medium of national or international communication. If the language has previously existed only in spoken form, or in an unusual writing system, an alphabet will have to be devised, along with rules of spelling and punctuation. An early aim will be the codification of the pronunciation, grammar, and vocabulary to provide a set of norms for standard use, especially if there is a great deal of local variation.

Modernization
The vocabulary will need to be modernized, to enable foreign material (in such areas as science, medicine, or the consumer society) to be translated in a consistent way. Principles will have to be agreed for the introduction of new terms; for example, should they be loan words, or coinages based on native roots? New styles of discourse may need to be developed, for use on radio or in the press. Decisions will need to be made about new or uncertain usages, especially in technical contexts (e.g. how to abbreviate scientific terms).

Implementation
The chosen standard will need to be officially implemented, by using it for government publications, in the media, and in schools. Inevitably, it will come to be viewed as the 'best' form of language in the speech community (§1), because it will be associated with educational progress and social status. It will also provide the norm for literary style, and may be associated with factors of a nationalistic, cultural, or religious kind. In due course, it is likely to be promulgated as a norm through an official body, such as an academy, or through prescriptive grammars, dictionaries, and manuals of usage.

INVENTING AN ALPHABET

One of the first tasks facing explorers, missionaries, and administrators, when they encounter a new language, is to devise a means of writing it down. The basic linguistic task is to ensure that each phoneme is represented by a grapheme (§§28, 33). But there are hundreds of possible graphemic shapes: /ʧ/, for example, could be written as *c, č, ć, ch, ts, tch,* and in many other ways. The choice between them involves factors of a psychological, historical, social, and educational kind. Language-planning principles thus need to be borne in mind from the outset.

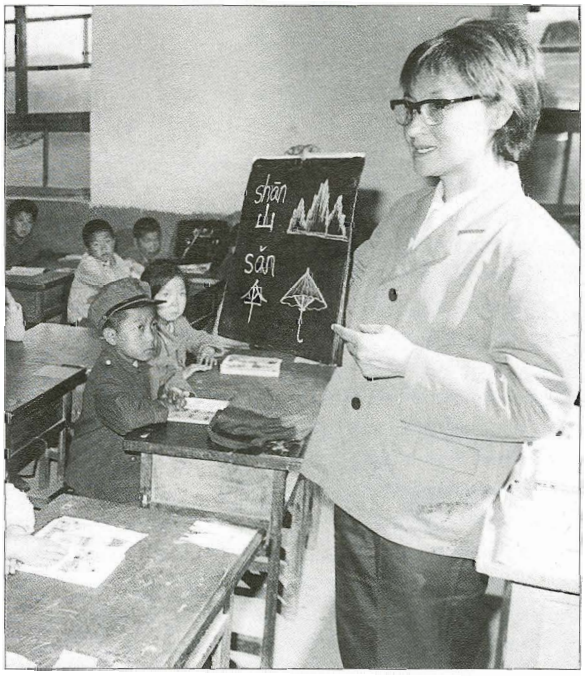
Political, religious, and other considerations may affect the choice of which kind of alphabet to adopt. A community may wish to 'align' with countries that use Roman, Cyrillic, Arabic, or other alphabets. It may also be important to choose a set of characters that can be used by all the languages throughout an area (as in the case of the All-India Alphabet). Written uniformity is often a powerful political symbol. It is also an economical measure, as it reduces the costs of printing and word processing.

For a language where there are many new sounds, a decision has to be made about whether to invent new letters, combine letters into digraphs, or go in for diacritics (such as accents). If the first path is taken, there is still the question of whether the new forms should be adaptations of familiar letters, or totally fresh inventions (as in the use of some phonetic symbols).

Many other questions need to be considered. For example, if some features of a language are only occasionally used to contrast meanings (as often happens with the tones of a tone language, p. 174), should they be systematically represented by some form of symbol, or can they be ignored? Should grammatical differences be represented in the spelling (as in the case of English *find* vs *fined*)? And how should loan words, with their distinctive phonology, be written down? Even a well-established writing system can be faced with problems of this kind, as in the continuing debate over whether French loans in English should keep their accents (*rôle, cliché, résumé*, etc.).

A PLANNING MYTH

Probably the best-known myth in the history of language planning is the story that German nearly became the national language of the U.S. in the 19th century, losing to English by only one vote in the legislature (the 'Muhlenberg' legend). In fact, all that was involved was a request, made by a group of Virginia Germans, to have certain laws issued in German *as well as* in English. The proposal was rejected by one vote, apparently cast by a German-speaking Lutheran clergyman, Frederick Muhlenberg (1750–1801). But the general status of English as the major language was never in doubt. (After S. B. Heath & F. Mandabach, 1983.)



Chinese language planning Some of the most ambitious programmes of language planning ever conceived have taken place in China since the 1950s, with hundreds of millions of people affected. The two main developments have been the provision of a romanized alphabet (*pīn-yīn*), and the promotion of a common spoken language, *pǔtónghuà*, to provide a means of communication between the various regional languages (p. 314). Reports of early progress in the campaign are illustrated by Datian county in Fujian province, which has over a dozen dialects, and where it was said that 'people separated by a blade of grass could not understand each other'. A group of officials from the north on one occasion needed as many as seven interpreters to make a speech to the people in this area. But after an active teaching campaign, officials using *pǔtónghuà* were able to address large crowds without any interpreter being needed. The picture shows a *pīn-yīn* class taking place in an experimental school in Ningwu County, Shanxi.

ALPHABETS IN CONFLICT

The Roman alphabet has been so successful that it has begun to threaten the status of other alphabets. A question mark hangs over the future of Chinese characters, now that the romanized system known as *pīn-yīn* has been brought into use (p. 315). And in India, there is a body known as Roman Lipi Parishad (RLP) campaigning for the adoption of the Roman alphabet for the main languages of the country.

The arguments are complex ones, as can be seen from the situation in India.

- The RLP argue that the country cannot afford the luxury of making machines for each of the alphabetic scripts used in India (p. 205). Already, some 70% of mechanical typewriters in India are made for English, and the rest for all the other scripts. Electronic typewriters are made only for English.
- The RLP point to the need to anticipate the use of computers, in relation to the country's economy. The Roman script is easier to adapt to electronic screens and keyboards than the various Devanagari scripts. A larger dot-matrix system (p. 195) would be needed, to cope with the diacritics that are used above, below, preceding, and following the Devanagari letters.
- It is claimed that there is a greater demand for material in the Roman alphabet. In Bombay, for example, there was an experiment in which telephone directories were printed using both the English and Devanagari alphabets. There was a demand of 300,000 for the former; but less than 50% of the 5,000 Devanagari copies were sold.
- On the other hand, the Roman script is not accepted as an alternative by any of the 22 Indian languages recognized by the Sahitya Akadami, the highest body devoted to literature.
- The cultural identity of the main groups in India is very much bound up with the use of an individual alphabet.

Opponents therefore argue that the adoption of Roman script would diminish one of the most important symbols of identity (§§9–10), and perhaps be the thin end of the wedge towards the eventual supplanting of indigenous scripts. These are highly emotive issues, and it remains to be seen whether the economic arguments will be able to make much progress, given the highly charged atmosphere of linguistic debate in present-day India.

CAPITALS IN FRISIAN?

Frisian, spoken in several dialects in the northern part of Schleswig-Holstein, provides a good example of the way the invention of spelling rules can reflect social forces. In devising an orthography for the language, the question arose as to whether nouns should be written with a capital letter, as in German, or with a lower-case letter, as in other languages. Support for the capital letter proposal came from those who wished to see Frisian's ties with Germany strengthened. Opposition came from those who wished to see a more autonomous future for Frisian. The issue remains unresolved, with both groups arguing the relative merits of each position, and producing publications that follow their favoured orthographic principle. (After A. Walker, 1984.)

THEORIES OF LANGUAGE LEARNING

As with the study of first language acquisition (p. 236), several theories of the nature of the FLL process have been propounded, with similar issues being addressed. Indeed, comparisons are frequently made with the way children learn their first language (L1), as a means of providing hypotheses to guide FL research.

THE BEHAVIOURIST VIEW

A great deal of language learning and teaching in the 1950s and 1960s was influenced by the tenets of behaviourism (pp. 236, 412). In this view, FLL is seen as a process of imitation and reinforcement: learners attempt to copy what they hear, and by regular practice they establish a set of acceptable habits in the new language. Properties of the L1 are thought to exercise an influence on the course of L2 learning: learners 'transfer' sounds, structures, and usages from one language to the other. A widely used typology distinguishes two kinds of transfer. Similarities between the two languages cause 'positive transfer': it proves acceptable to use the L1 habits in the L2 setting (e.g. the assumption that the subject goes before the verb satisfactorily transfers from English to French). Differences cause 'negative transfer', generally known as 'interference': the L1 habits cause errors in the L2 (e.g. the same assumption about subject–verb order does not satisfactorily transfer into Welsh). Typical interference errors include: *I wait here since 3 hours* (from French) and *How long must my hand in plaster stay?* (from German). Problems of negative transfer are thought to provide a major source of FLL difficulty. The main aim of behaviourist teaching is thus to form new, correct linguistic habits through intensive practice, eliminating interference errors in the process.

There are several problems presented by this account of FLL. Imitation alone does not provide a means of identifying the task facing learners, who are continually confronted with the need to create and recognize novel utterances that go beyond the limitations of the model sentences they may have practised. Nor does imitation suffice as an explanation of the way learners behave: not many of the errors that are theoretically predicted by the differences between L1 and L2 in fact occur in the language of learners; and conversely, other errors are found that seem unrelated to the L1. In a frequently-cited early study (H. C. Dulay & M. K. Burt, 1973), 145 Spanish-speaking children aged 5 to 8 were observed while learning English. Six structures were selected and the error patterns analysed. It emerged that interference errors (such as *They have hunger* from *Ellos tienen hambre*) accounted for only 3% of the errors made. The majority of the errors (85%, with a further 12% unclear) were thought to resemble those that appear in the course of L1 acquisition (e.g. *They hungry*). Analyses of this kind have proved to be controversial (largely because of difficulties in validating the error analysis – see below), but their general conclusion is widely supported. The systematic comparison of L1 and L2, in order to predict areas of greatest learning difficulty – a procedure known as *contrastive analysis* – explains only a small part of what goes on in FLL.

THE COGNITIVE VIEW

The main alternative to the behaviourist approach sees as central the role of cognitive factors in language learning (pp. 236–7). In this view, learners are credited with using their cognitive abilities in a creative way to work out hypotheses about the structure of the FL. They construct rules, try them out, and alter them if they prove to be inadequate. Language learning, in this account, proceeds in a series of transitional stages, as learners acquire more knowledge of the L2. At each stage, they are in control of a language system that is equivalent to neither the L1 nor the L2 – an *interlanguage* (L. Selinker, 1972).

Error analysis plays a central role in this approach. Errors are likely to emerge when learners make the wrong deductions about the nature of the L2, such as assuming that a pattern is general, when in fact there are exceptions. The errors provide positive evidence about the nature of the learning process, as the learner gradually works out what the FL system is. For example, learners who say *vous disez* instead of *vous dites* 'you say' have assumed, wrongly, that the *-ez* ending found after *vous* in most other French verbs (*marchez, donnez*, etc.) also applies to *dire* 'say'. The error in this case indicates that a faulty generalization (or analogy, p. 236) has been made.

Since the 1970s, cognitive approaches to FLL have been in the ascendant, and error analysis in particular has attracted a great deal of attention. However, the analysis of errors turns out to be a highly complex matter, involving other factors than the cognitive. Some errors are due to the influence of the mother tongue, as contrastive analysis claims. Some come from external influences, such as inadequate teaching or materials. Some arise out of the need to make oneself understood by whatever means possible (e.g. replacing words by

TWO MODELS OF FOREIGN LANGUAGE LEARNING

Behaviourist	**Cognitive**
L2 input obtained from controlled, formal, instruction	Exposure to authentic use of L2 in near-natural situations
↓	↓
Imitation and reinforcement (conscious) strategies	Input processed using natural (universal, unconscious) strategies
↓	↓
L2 habits established	Transitional stages of learning (interlanguage)
↓	↓
L2 output	L2 output

THE MONITOR MODEL

In the 1970s, an influential view of the relationship between acquisition and learning was propounded by the American linguist, Stephen Krashen (1941–). This account recognizes a subconscious, natural process ('acquisition'), which is the primary force behind FL fluency. 'Learning' is seen as a conscious process that monitors, or edits, the progress of acquisition and guides the performance of the speaker. Its role is – or should be – minor, being used only to correct errors in speech or to give speech a more 'polished' appearance.

The emphasis on acquisition leads Krashen to propose an 'input hypothesis', which suggests that teachers should try to replicate in the classroom the conditions which occur in L1 acquisition. The parallel is drawn between the input of teacher to student and that from mother (or caretaker) to child (see facing page).

In fact, traditional FLT provides learners with a great deal of conscious knowledge of linguistic rules. As a result, they may come to rely too much on this knowledge, so that it actually gets in the way of their ability to communicate. People who worry too much about making a mistake, and who thus are reluctant to use their FL ability, are in this view 'overusing' their monitor.

Theories of this kind are inevitably controversial, given our limited knowledge of the psychological processes involved in speech production. There is plainly a need to take into account the distinction between conscious and subconscious awareness in language processing, and between formal and informal settings, but the way these variables interact, it has been argued, is more complex than anything which has so far been proposed.

In particular, since the 1980s, attention has been focused on applying a psycholinguistic perspective to FLL, in which the varying demands of information processing by the brain are used as an explanation for variability of errors. The student who can say *Mrs Brown lives in Reading* but **Mrs Brown who live in Reading has just won the lottery* makes the error because of the extra load involved in processing the sub-

gestures). Moreover, not all errors are equally systematic, disruptive, or unacceptable. Errors of vocabulary, for example, are less general and predictable than errors of grammar, but they are usually more disruptive of communication. Some errors, indeed, become so acceptable that they do not disappear: they become 'fossilized' – tolerated by learners (insofar as they are conscious of them) because they do not cause major problems of communication (e.g. the pronunciation errors that constitute a foreign accent).

Above all, error analysis is complicated by the fact that it is often unclear what the learner intended to say, and thus how to identify the error that has been made. For example, does *The lady eat it* display an error of the noun (*ladies*) or verb – and if the latter, should the correct form be *eats*, *is eating*, *ate*, or some other variant? And even if we assume that the speaker intended to say *eats*, we are still left with the question of whether the error is one of pronunciation (the speaker having difficulty with the [ts] cluster) or grammar – and, within the latter heading, whether the difficulty is one of morphology (lack of awareness of the ending) or syntax (lack of awareness of number agreement between subject and verb) (§16).

Despite the difficulties, research into errors continues to provide a fruitful way of investigating the processes underlying FL acquisition. However, as with contrastive analysis, the approach cannot provide a complete explanation. Most FLL settings do not constitute the kind of 'pure', natural linguistic situation that is presupposed by the cognitive approach, but contain elements of formal teaching, in which learners are systematically introduced to fragments of the L2 (e.g. one tense at a time). To understand the way languages come to be learned in these 'mixed' settings, it is thus proving necessary to devise more sophisticated models, which focus on the relationship between the processes of natural acquisition and those of formal learning, and which pay adequate attention to the needs and aims of the students, and to the nature of the social setting in which FLL interaction takes place.

There is something in the poetry of will
Wordsworth which ~~is~~ always ~~to~~ live. He

ERRORS IN LANGUAGE LEARNING

The error in this sentence, written by a Swedish student, seems straightforward, but it is not easy to say exactly what the error is, why it was made, and whether the teacher has made the best correction. Is the student confusing *be to* and *will*? Or has he learned the past tense use of *be to* in this context (as in *There was something in the poetry which was to live forever*), and assumed that the present tense would work in the same way? If so, is there not an additional error in the position of *always*? And would not *forever* be a more idiomatic word? The corresponding construction in Swedish is *som alltid skall leva*, but this will not explain all that is going on.

THE NATURAL ORDER HYPOTHESIS

During the 1970s, several studies drew attention to the fact that different FL learners make similar errors, regardless of their language background. Such errors as *I going* and *this a book* were observed in Spanish, Russian, Japanese, and several other learners of English. The conclusion was drawn that there must be a universal creative process at work; learners were said to be following a natural 'internal syllabus' (as opposed to the 'external' syllabus of the classroom). Several of the errors closely resembled those made by children learning their mother tongue. Analogies were therefore drawn with the 'language acquisition device' postulated by some child-language analysts (p. 234), and a parallel was proposed between the natural order of L1 acquisition and the way people acquired a foreign language.

Particular attention was focused on the way in which foreign learners of English used a set of grammatical morphemes (§16), such as *-ing*, *-ed*, and plural *-s*, which L1 studies had already found to be acquired in a certain order (p. 244). The errors learners made with each item were counted, and the morphemes were ranked on the basis of how accurately they were used. This ranking was then assumed to reflect the order in which the learners were acquiring these morphemes. Similar orders were found in several different FLL contexts, in both spoken and written language, thus supporting the idea of a natural, universal sequence of acquisition that was independent of the influence of the learner's first language.

If natural order exists, there would be major implications for external syllabuses, which would presumably be modified in that direction. However, criticisms have been made of this kind of approach. Order of acquisition as based on a cross-sectional study of speech samples may not correspond to the order of acquisition that would emerge from a longitudinal study (p. 231). The findings are of limited generality: only a very small number of grammatical items have been analysed, and there have been very few studies (most of which to date have focused on English, so that it is unclear how genuine the claimed universals are). And differences in acquisition order have already begun to emerge, casting doubt on the universality of the natural order hypothesis.

CHILD VS ADULT ACQUISITION

The similarities between L1 and L2 acquisition errors are striking, but there are many differences between the two kinds of learning situation (over and above issues of neurological development, p. 265), which makes it difficult to see a parallel between adult foreign language learners and young children acquiring their mother tongue.

- The adult has a set of formed cognitive skills and strategies that should make the FLL task easier (e.g. the ability to memorize, imitate, and use dictionaries). A major asset is the ability of most adults to read and write.
- Adults already have a language, and this inevitably reduces their motivation to learn another beyond minimal levels. Migrants, for example, generally learn only enough to enable them to survive in their new country.
- There are several emotional differences between adults and children when it comes to learning. In particular, adults are more self-conscious about FLL, and are less able to assimilate cultural differences.
- Adults meet a greater variety of L2 situations than do children learning their L1. Children's needs are also very different (e.g. they need language for play and emotional expression). Accordingly, the range of teaching objectives will differ in each case.
- The adult has less time and opportunity than the child for FLL. Some estimates suggest that it takes well over a year to accumulate as much L2 experience as a young child gets from the L1 in a month.
- Adults invariably find themselves in a less natural learning environment than children. It is rarely possible to devise a teaching situation which closely resembles that encountered by the L1 child, with its one-to-one interaction and strong emotional (caregiver) support.
- There is an uncertain parallel between the way in which mothers talk to their children and the way in which people talk to adults using a foreign language ('foreigner talk'). Certainly, adult L1 speakers adapt to learners, and (often unconsciously) try to help them by speaking slower and louder, repeating words, simplifying their grammar, and using stereotyped expressions (of which pidgin *savvy* is probably the most famous). They also ignore many errors. But it is unclear how universal or how systematic these input strategies are.
- Similarly, it is unclear how far teacher language displays correspondences with motherese (p. 237); the differences, at present, are more striking than the similarities. To facilitate learning, in the early stages, teachers need to keep their input relatively simple, interesting, comprehensible, relevant to the learning task, sufficiently repetitive to enable patterns to be perceived, and capable of providing appropriate feedback. Generalization proves difficult, given the great variation that exists among teaching methods (p. 378).

64 · LANGUAGE AND OTHER COMMUNICATION SYSTEMS

A widely recognized problem with the term 'language' is the great range of its application. This word has prompted innumerable definitions. Some focus on the general concept of 'language', some on the more specific notion of 'a language'. Some draw attention to the formal features of phonology (or graphology), grammar, and semantics (Parts III–VI). Some emphasize the range of functions that language performs (Parts I, II). Some stress the differences between language and other forms of human, animal, or machine communication (see below). Some point to the similarities. At one extreme, there are definitions that are highly technical in character; at the other, there are extremely general statements, reflecting the way in which the notion has been applied figuratively to all forms of human behaviour, such as the 'language' of music, cookery, or the cinema.

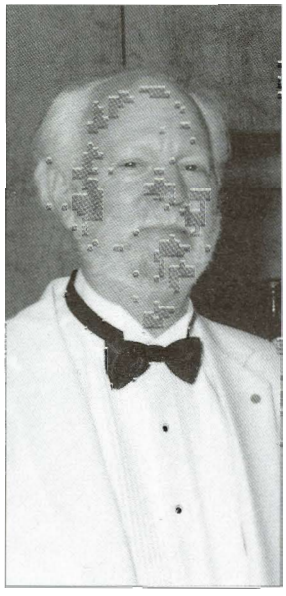
Most textbooks in the subject avoid the problem, preferring to characterize the notion of language rather than define it. They recognize that the question of identifying an individual language has no single, simple answer, because formal and social criteria are often in conflict (§47). Similarly, they note the correspondingly complex problems that arise when attempting to construct a definition of language in general that makes a precise and comprehensive statement about formal and functional universal properties. The set of definitions given below exemplifies the way different writers have attempted to tackle the problem, and illustrates some of the difficulties involved. There seems little to be gained by trying to summarize the content of the present volume in a single sentence – unless it is the banal observation that 'language' is what this encyclopedia is about!

A more useful approach to language, and one used by most modern linguists, is to identify the various properties that are thought to be its essential defining characteristics. The aim is to determine what 'counts' as a human language, as opposed to some other system of communication. Two main kinds of enquiry have been used. One focuses upon identifying the universal structural properties of language, and this is discussed in Part III (§§13–15). The other is to contrast language with non-human forms of communication and with other forms of human communication.

DESIGN FEATURES OF COMMUNICATION

The most widely acknowledged comparative approach has been that proposed by the American linguist Charles F. Hockett (1916–), who used a zoological mode of enquiry to identify the main points of connection between language and other systems of communication, especially those found in animals. His set of 13 design features of communication using spoken language were as follows:

- *Auditory–vocal channel* Sound is used between mouth and ear, as opposed to a visual, tactile, or other means (pp. 405–7).
- *Broadcast transmission and directional reception* A signal can be heard by any auditory system within earshot, and the source can be located using the ears' direction-finding ability (p. 142).
- *Rapid fading* Auditory signals are transitory, and do not await the hearer's convenience (unlike animal tracks, or writing, §31).



Charles F. Hockett

LANGUAGE DEFINITIONS

'Language is a purely human and non-instinctive method of communicating ideas, emotions and desires by means of voluntarily produced symbols.' (E. Sapir, 1921.)

'A language is a system of arbitrary vocal symbols by means of which the members of a society interact in terms of their total culture.' (G. Trager, 1949.)

A language is 'a set (finite or infinite) of sentences, each finite in length and constructed out of a finite set of elements'. (A. N. Chomsky, 1957.)

Language is 'the institution whereby humans communicate and interact with each other by means of habitually used oral–auditory arbitrary symbols'. (R. A. Hall, 1964.)

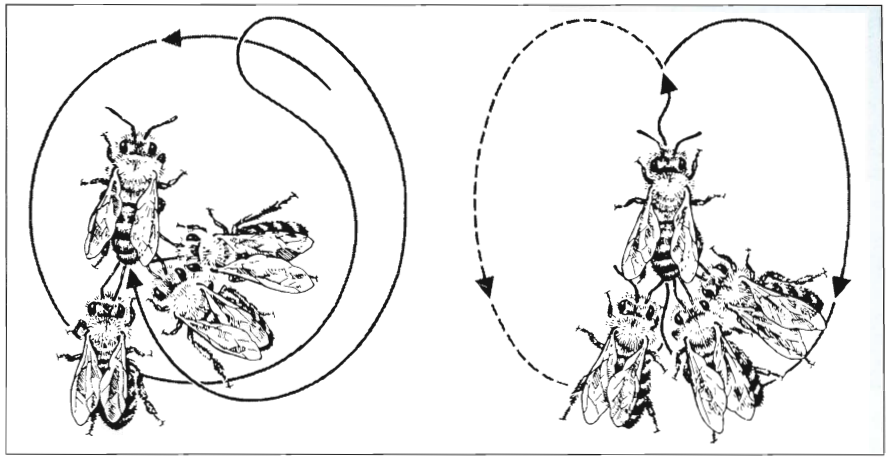
A dictionary definition

1. the words, their pronunciation, and the methods of combining them used and understood by a considerable community and established by long usage.
2a. audible, articulate, meaningful sound as produced by the action of the vocal organs.
2b. a systematic means of communicating ideas or feelings by the use of conventionalized signs, sounds, gestures, or marks having understood meanings.
2c. an artificially constructed primarily formal system of signs and symbols (as symbolic logic) including rules for the formation of admissible expressions and for their transformation.
2d. the means by which animals communicate or are thought to communicate with each other.
3. the faculty of verbal expression and the use of words in human intercourse ... significant communication.
4. a special manner or use of expression.
(*Webster's Third New International Dictionary*, 1961.)

And a comment

'The question "What is language?" is comparable with – and, some would say, hardly less profound than – "What is life?", the presuppositions of which circumscribe and unify the biological sciences ... it is not so much the question itself as the particular interpretation that the biologist puts upon it and the unravelling of its more detailed implications within some currently accepted theoretical framework that nourish the biologist's day-to-day speculations and research. So it is for the linguist in relation to the question "What is language?"'
(J. Lyons, 1981, p. 1.)

- *Interchangeability* Speakers of a language can reproduce any linguistic message they can understand (unlike the differing courtship behaviour of males and females in several species).
- *Total feedback* Speakers hear and can reflect upon everything that they say (unlike the visual displays often used in animal courtship, which are not visible to the displayer).
- *Specialization* The sound waves of speech have no function other than to signal meaning (unlike the audible panting of dogs, which has a biological purpose).
- *Semanticity* The elements of the signal convey meaning through their stable association with real-world situations (unlike dog panting, which does not 'mean' a dog is hot; it is 'part of' being hot).
- *Arbitrariness* There is no dependence of the element of the signal on the nature of the reality to which it refers (unlike the speed of bee 'dancing', which directly reflects the distance of the nectar from the hive).
- *Discreteness* Speech uses a small set of sound elements that clearly contrast with each other (unlike growling, and other emotional noises, where there are continuous scales of variation in strength).



- *Displacement* It is possible to talk about events remote in space or time from the situation of the speaker (unlike most animal cries, which reflect immediate environmental stimuli).
- *Productivity* There is an infinite capacity to express and understand meaning, by using old sentence elements to produce new sentences (unlike the limited, fixed set of calls used by animals).
- *Traditional transmission* Language is transmitted from one generation to the next primarily by a process of teaching and learning (unlike the bee's ability to communicate the source of nectar, which is passed on genetically).
- *Duality of patterning* The sounds of language have no intrinsic meaning, but combine in different ways to form elements (such as words) that do convey meaning (unlike animal calls, which cannot be analysed into two such levels of structure).

The 'language' of bees
One of the most closely investigated forms of animal communication is the 'dance' performed by a honey bee when it returns to the hive, which conveys precise information about the source and amount of food it has discovered. Several kinds of movement pattern have been observed. In the 'round dance' (above, left) used when the food source is close to the hive, the bee moves in circles alternately to left and right. In the 'tail-wagging dance' (above, right), used when the source is further away, the bee moves in a straight line while wagging her abdomen from side to side, then returns to her starting point. The straight line points in the direction of the food, the liveliness of the dance indicates how rich a source it is, and the tempo of the dance provides information about its distance. For example, in one study, an experimental feeding dish 330 metres from the hive was indicated by 15 complete runs through the pattern in 30 seconds, whereas when the dish was moved to 700 metres distance, only 11 runs were carried out in that time. No other animal communication system seems able to provide such a quantity of precise information – except human language. (After K. von Frisch, 1962.)

The applicability of the 13 design features to six systems of communication (after C. F. Hockett, 1960, pp. 10–11). The music column refers only to western music since the time of Bach. A question mark indicates that it is unclear or unknown whether a system has a particular feature. A blank space indicates that a feature cannot be determined because other information is lacking.

	Bee dancing	Stickleback courtship	Western meadowlark song	Gibbon calls	Language	Instrumental music
The vocal-auditory channel	no	no	yes	yes	yes	auditory, not vocal
Broadcast transmission and directional reception	yes	yes	yes	yes	yes	yes
Rapid fading	?	?	yes	yes, repeated	yes	yes
Interchangeability	limited	no	?	yes	yes	?
Total feedback	?	no	yes	yes	yes	yes
Specialization	?	in part	yes?	yes	yes	yes
Semanticity	yes	no	in part?	yes	yes	no (in general)
Arbitrariness	no		if semantic, yes	yes	yes	
Discreteness	no	?	?	yes	yes	in part
Displacement	yes, always		?	no	yes, often	
Productivity	yes	no	?	no	yes	yes
Traditional transmission	probably not	no?	?	?	yes	yes
Duality of patterning	no		?	no	yes	

CHIMP COMMUNICATION

The formal and functional complexity of language is such a distinctive human trait that many scholars think the designation *homo loquens* ('speaking man') to be a better way of identifying the species than any other single criterion that has been suggested (such as tool using) (p. 293). This is not to disregard the complex patterns that have been observed in the natural communicative systems of birds, insects, apes, and other animals (the subject matter of the field of *zöosemiotics*). But no animal system remotely compares with the level of sophistication found in human language. The evolutionary gap is very wide. Only the recent experiments in teaching language to chimpanzees have suggested that this gap may be somewhat narrower than has traditionally been assumed.

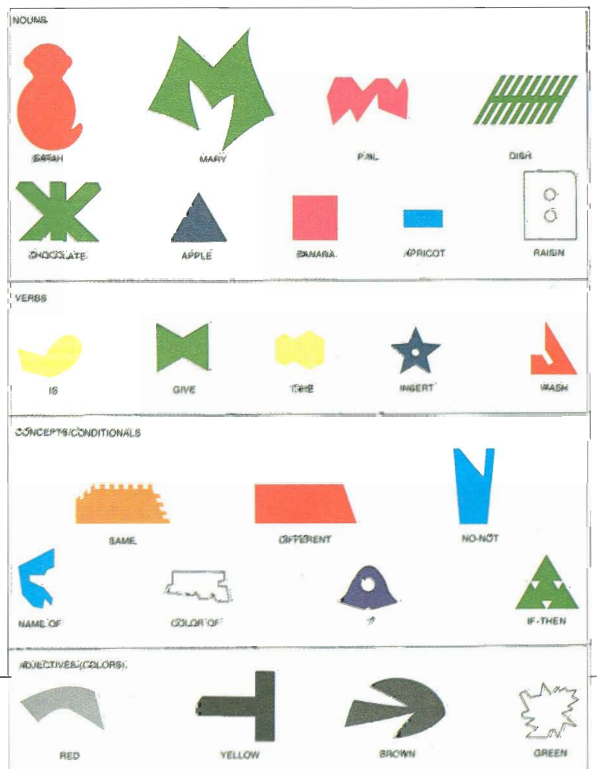
Early experiments to teach chimpanzees to communicate with their voices failed because of the insufficiencies of the animals' vocal organs (p. 292). However, when attempts were made to communicate with them using the hands, by teaching a selection of signs from American Sign Language (ASL, see Part VI), dramatic progress was claimed. The first subject was a female chimpanzee named Washoe, whose training began in 1966 when she was less than a year old. It took her just over four years to acquire 132 ASL signs, many of which bore striking similarities to the general word meanings observed in child language acquisition (Part VII). She also began to put signs together to express a small set of meaning relations, which resembled some of the early sentences of young children, such as *want berry, time drink, there shoe* (B. T. & R. A. Gardner, 1975).

Since then, several other chimpanzees (and also gorillas) have acquired a vocabulary of signs, and alternative teaching procedures have been tried. For example, in the case of the chimps Moja and Pili, sign language teaching began soon after birth, and training was carried out by native signers. Both chimps began to sign when they were about 3 months old, and had over a dozen signs by the age of 6 months – a marked contrast with Washoe, who had only 2 signs after 6 months of training.

A quite different way of proceeding was introduced in the case of a 5-year-old chimpanzee called Sarah, in a research programme that began in 1954 (D. & A. J. Premack, 1983). She (and, later, several others) was taught a form of written language – to arrange and respond to vertical sequences of plastic tokens on a magnetic board. Each token represented a word, e.g. small blue triangle = *apple*, small pink square = *banana*. In due course, the trainer was able to teach Sarah to respond correctly to several basic semantic sequences (e.g. 'give Mary apple'), including a number of more abstract notions, such as 'same/different' and 'if/then' (e.g. ? *apple different banana*).

Chimp language research attracted considerable media publicity in its early years, with reporters focusing on the implications of the work. What would chimps say if they could use language? What would they think of the human race? Would they claim civil rights? Such speculations were wholly premature, given the limited findings of the research to date. These findings are in any case controversial, receiving a range of reactions extending from total support to total antipathy. A variety of interpretations seems possible. It is evident that chimps can learn to imitate signs, combine them into sequences, and use them in different contexts, but the explanation of this behaviour is less clear. Many scholars believe that the chimps' behaviour can be explained as a sophisticated imitation ability rather than as evidence for some form of linguistic processing, and they argue the need for fuller accounts to be provided of chimp behaviour, and of the training methods used, in order to evaluate the claims being made about learning. More systematic data have begun to be collected, but it will be some time before these questions can be resolved.

'Peony nose touch' Peony, one of the 'second generation' of chimps trained in the Premack study, carries out this instruction, which her trainer has placed on the magnetic board. (D. Premack & A. J. Premack, 1983, p. 29.)



WASHOE'S WORDS

Washoe's typical vocabulary can be seen from the signs she used in a study of her responses to 500 questions. The signs were grouped by the authors into 13 general types (it should be noted that an idiosyncratically broad notion of 'noun' is used, including such items as *dirty* and *listen*):

Proper names (her companions)
Don, Dr G, Greg, Roger, Linn, Mrs G, Susan, Washoe

Pronouns
me, we, you

Common nouns

baby	dirty	nut
bath	drink	pants
bed	flower	pencil
berry	food	purse
bird	fruit	ride
blanket	gun	shoe
book	hammer	smoke
brush	hat	spoon
bug	ice	swallow
car	key	sweet
cereal	leaf	tree
chair	listen	water
cheese	lollipop	window
clothes	look	wiper
comb	man	woman
cow	meat	

Possessives
mine, yours

Traits
funny, good, hungry, stupid

Colours
black, white, green, red

Temporal
time

Negative
can't, enough, no

Imperative
gimme, help

Appetitive
please, want

Quantitative
hurry, more

Verbs
bite, catch, cry, go, hug, open, peekaboo, smile, tickle

Locatives
in, out, up, there

Typical sequences

Me Washoe	Food fruit
You me out	Time drink
Susan bite there	Good me

Sarah's symbols (left) Symbols used in communicating with Sarah and the other chimps. (From D. Premack & A. J. Premack, 1983, p. 21.)

SEMIOTICS

Language can also be studied as part of a much wider domain of enquiry: *semiology*, or *semiotics* - a subject which owes a great deal to US philosopher Charles Sanders Peirce (1839–1914), as well as to the work of Ferdinand de Saussure (p. 411). This field investigates the structure of all possible sign systems, and the role these play in the way we create and perceive patterns (or 'meanings') in sociocultural behaviour. The subject is all-inclusive, therefore, dealing with patterned human communication in all its modes (sound, sight, touch, smell, and taste) and in all contexts (e.g. dance, film, politics, eating, clothing). The subject matter of the present book would form but a small section of any proposed encyclopedia of semiotics.

AUDITORY-VOCAL

The diagram below shows the relationship between language, as identified in Parts III–VI, and other aspects of human communication. The structured use of the *auditory–vocal* mode, or channel (p. 404), results in the primary manifestation of language: speech. But non-linguistic uses of the vocal tract are also possible: physiological reflexes, such as coughing and snoring; musical effects, such as whistling; and the communication of identity, in the form of voice quality (§6). The suprasegmental aspects of vocal expression (§29) are usually included within the study of language, though it is difficult to draw a clear-cut boundary line between some of these effects (those placed under the heading of 'paralanguage', such as giggling and whispering) and those that clearly fall outside language.

VISUAL

The visual mode is used for a variety of purposes, some linguistic, some not. The primary way in which visual effects have a linguistic use is in the various deaf sign languages (Part VI). In addition, there is the historically derivative use of the visual mode that resulted in the development of written language. Further writing-based codes, such as semaphore and morse, would also be included here. Non-linguistic forms of visual communication include the systems of facial expression and bodily gesture, which are the subject matter of *kinesics* (p. 406).

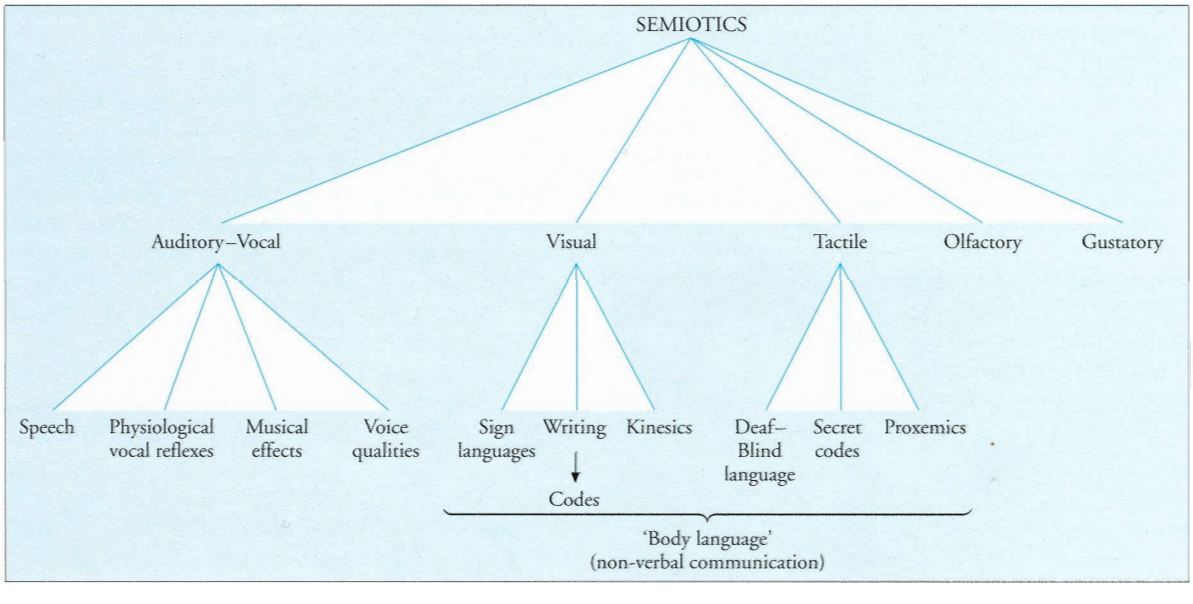
TACTILE

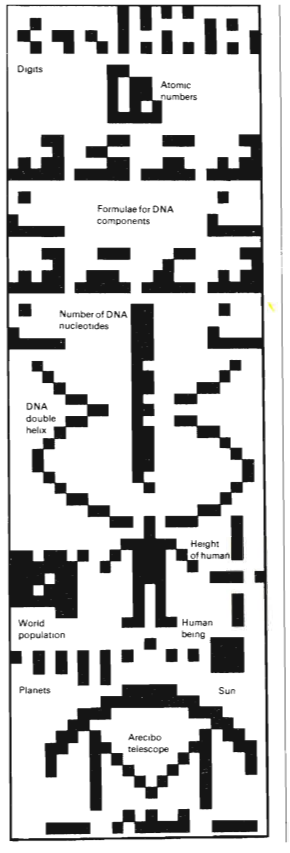
Tactile communication has very limited linguistic function, apart from its use in deaf–blind communication and in various secret codes based on spoken or written language (p. 58). Its main uses are non-linguistic, in the form of the various ways in which bodily contact and physical distance between people can signal contrasts of meaning – the subject matter of *proxemics* (p. 401).

The communicative use of the visual and tactile modes is often referred to as 'nonverbal communication', especially in academic discussion. In everyday terms, it is the area of 'body language'.

OLFACTORY AND GUSTATORY

There seems to be little active role for the olfactory and gustatory modes in human communication (a marked contrast with the important use of these senses for communicative purposes in the animal kingdom). However, they do play an important part in our reception of information about the outside world (e.g. in smelling and tasting food). The communicative use of body odour seems to have a mainly sexual role in human society; but there are several anecdotes of its use in other domains. One linguist even claimed to be able to tell when his informants (p. 414) were under strain (and perhaps therefore were being less reliable) by the different body odour they exuded!





Other modes? This is the pictographic message transmitted into space by the Arecibo radio telescope in Puerto Rico in 1974. The signal was aimed at the cluster of 300,000 stars, known as M13, in the Hercules constellation.

The message consists of a series of radio pulses which can be arranged into a pictogram. It includes data on the chemical basis of life on earth, the human form, and the solar system. It assumes, of course, that the communicative system of the receiving species is capable of responding to the same semiotic contrasts as are displayed in the pictogram (shape, length, etc.). If the entity receiving the signal happens to have a communicative system based on, say, heat, the astronomers will have wasted their time!

The Hercules target is 24,000 light years away – which means that, if any one or thing is there to receive it, and chooses to reply, the response should arrive in about 50,000 years' time.

AUDITORY–VOCAL EFFECTS

The main systems of communication using the auditory–vocal channel have been described elsewhere in this volume (Part IV). However, from time to time linguists have reported types of auditory communication that fall outside the normal use of the vocal apparatus – notably, the whistled speech of several rural populations. This is found in some Central and South American tribes, as well as in the occasional European community (e.g. in Turkey and the Canary Islands, based on Turkish and Spanish respectively).

Whistled speech

Eusebio Martinez was observed one day standing in front of his hut, whistling to a man a considerable distance away. The man was passing on the trail below, going to market to sell a load of corn leaves which he was carrying. The man answered Eusebio's whistle with a whistle. The interchange was repeated several times with different whistles. Finally the man turned around, retraced his steps a short way and came up the footpath to Eusebio's hut. Without saying a word he dumped his load on the ground. Eusebio looked the load over, went into his hut, returned with some money, and paid the man his price. The man turned and left. Not a word had been spoken. They had talked, bargained over the price, and come to an agreement satisfactory to both parties – using only whistles as a medium of communication. (G. M. Cowan, 1948, p. 280.)

This conversation took place between Mazateco speakers, members of a tribe that lives in and around the State of Oaxaca, Mexico. The whistled conversations closely correspond to patterns of spoken language, as has been shown by having the whistlers translate their tunes into speech. It is thus quite unlike the unstructured whistling patterns used as attention signals (e.g. 'wolf-whistling') in Euro-American culture. For example, in the following sequence of whistled utterances (where the tones are classified from 1 (high) to 4 (low), and glides between tones are marked by a dash), quite specific meanings are signalled, as the following transcription of Mazateco shows:

1,1,3,3,2,4 *hme^{1} čʔa^{1} ši^{3} ki^{3}-čʔai^{2}-ve^{4}*
'What did you bring there?'
1,4,1,1 *čʔa^{1} na^{4} hme^{1}-ni^{1}*
'It is a load of corn.'
1,3,3,4,3 *hʁā^{1} ti^{3}-ʔmi^{3} kʊai^{4}-ʔni^{3}*
'Well where are you going with it?'
3,2,4,2,3,4 *te^{3} na^{2}nko^{4} ti^{2}-vhi^{3} koa^{4}*
'I am taking it to Tenango.'
3,3,3,3,2,3,2–4,3 *ʔa^{3}-ti^{3}-ʔmi^{3} ka^{3} te^{2} na^{3}-ni$^{2-4}$ʔni^{3}*
'Are you going to sell it then?'
2,3,3,2,2–3 *ti^{2}-vhi^{3} ka^{3} te^{2} na$^{2-3}$*
'I am going to sell it.'
1,1,3,2,4,4,2,3,1–3,4 *ho^{1} thɻ^{1} čʔai^{3}-ʔni^{2} ʔi^{4}-ta^{4} te^{2} na^{3}-nai$^{1-3}$-vi^{4}*
'How much will you take then? Sell it to me here.'
4–3,4,3,3,3,2,4 *ka$^{4-3}$ tąǫ^{4} kʔoa^{3} nka^{3} hnko^{3} ka^{2} ša^{4}*
'I will take $2.50 a box.'
(G. M. Cowan, 1948, pp. 284–5.)

The whistled tunes are based on the patterns of tone and rhythm used in the spoken language, and can convey precise distinctions. With very few exceptions, each 'syllable' of whistle corresponds to a syllable of speech. Ambiguity is uncommon, because the topic of the conversation is usually something evident in the situation of the speakers. However, it is important for both speakers to use the same musical key, otherwise confusion may arise.

Whistled dialogues tend to contain a small number of exchanges, and the utterances are short. They are most commonly heard when people are at a distance from each other (e.g. when working the land), but they can also be found in a variety of informal settings. Although women are able to understand whistled speech, it is normally used only by and between males.

Nuba (Sudan) musicians prepare for a tribal gathering.

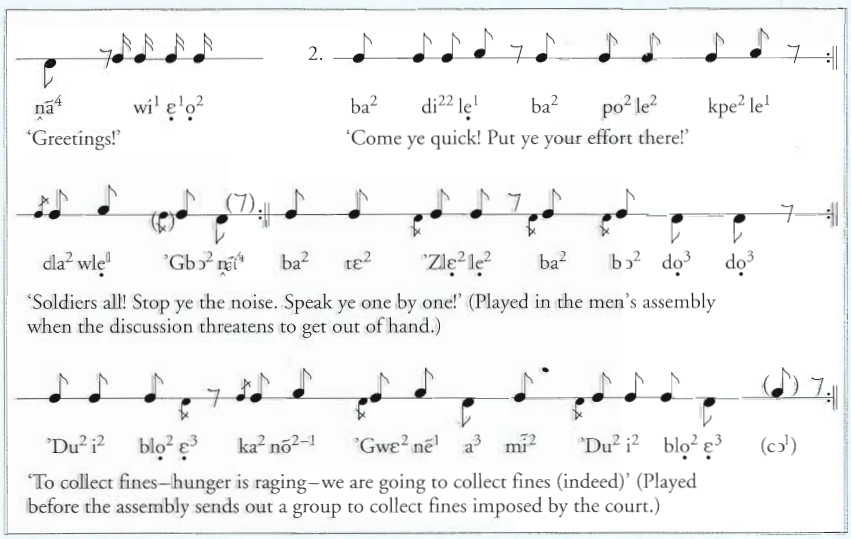
DRUM SIGNALLING

In several parts of the world – notably Africa, the Americas, and the Pacific – drums, gongs, horns, and other musical instruments have been used to simulate selected features of speech (primarily, tones and rhythms). In Africa, drums are the usual instruments involved, and quite elaborate systems of communication have developed.

One system, used among the Jabo tribe of Eastern Liberia, makes use of a wooden signal 'drum' (actually, more like a bell, as it has no skin covering) – a hollowed tree trunk, often over 2 metres in length. This has a longitudinal slit with lips varying in thickness, thus allowing several different tones to be produced. Two straight sticks are used for beating, and further tonal variations can be made by altering the way these sticks hit the drum. Other types of drum are also used for different purposes (such as dancing).

The drummer, an official of the town's law-enforcing authority, controls the way meetings take place, using special signals to do such things as call for order, summon people, and end the meeting. These signals consist mainly of fixed formulae, with a few variations and additions. The Jabo rarely use these drums for communicating with other villages (unlike the drum signalling found in many other parts of Africa).

The words and syllables of Jabo are tonal (§29): there are four basic tones, which are often linked by glides, and these interact with aspects of the vowel and consonant system. There is also considerable variation in the length of these tonal contrasts, which accounts for several of the drum patterns used. Some examples of these signals, with a transcription in Jabo, are given below. (From G. Herzog, 1945.)



TACTILE EFFECTS

The communicative use of touching behaviour, proxemics, has in recent years attracted a great deal of research by psychologists, sociologists, and anthropologists. A very wide range of activities is involved, as is suggested by this small selection of terms expressing bodily contact:

embrace	lay on (hands)	punch
guide	link (arms)	shake (hands)
hold	nudge	slap
kick	pat	spank
kiss	pinch	tickle

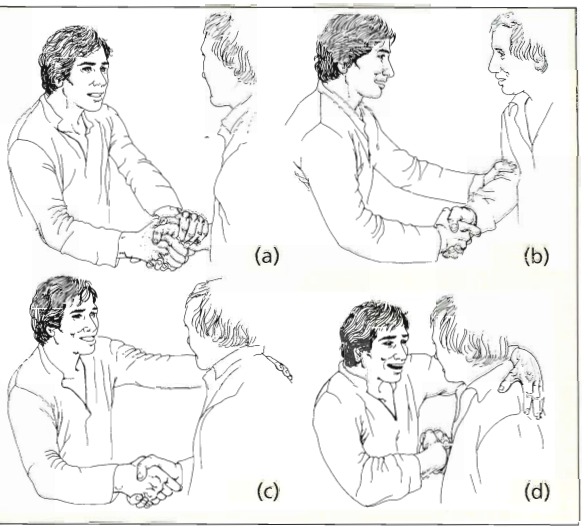
The communicative value of tactile activities is usually fairly clear within a culture, as they comprise some of the most primitive kinds of social interaction (several of the activities are found between animals). They express such 'meanings' as affection, aggression (both real and pretend), sexual attraction, greeting and leave taking, congratulation, gratitude, and the signalling of attention. They operate within a complex system of social constraints: some of the acts tend to be found only in private (notably, sexual touching); some are specialized in function (e.g. the tactile activities carried on by doctors, dentists, hairdressers, or tailors); and some are restricted to certain ceremonies (e.g. weddings, graduation, healing). Everyone has a subjective impression about how these activities take place, and what they mean. But there are many differences in behaviour between individuals and groups, and it is not easy to make accurate generalizations about society as a whole.

It is difficult to study tactile activity in an objective way: a basic problem is how to obtain clear recordings in which the participants are unaware of the observer (especially if the behaviour is being filmed). There are thus few detailed accounts of the range of communicative tactile acts in a society, and of the factors governing their use. It is evident, however, that some societies are much more tolerant of touching than others, so much so that a distinction has been proposed between 'contact' and 'non-contact' societies – those that favour touching (such as Arabs and Latin Americans), and those that avoid it (such as North Europeans and Indians). In one study of couples sitting together in cafés, it was found that in Puerto Rico the people touched each other on average 180 times an hour; in Paris it was 110 times an hour; whereas in London there was no touching at all (S. M. Jourard, 1963).

The distance people stand from each other, and the way they hold their bodies when interacting, are other important facets of proxemic behaviour. There are norms of proximity and orientation within a culture that communicate information about the social relationship between the participants. A common research procedure is to observe the point at which people are made to feel uncomfortable when others invade their 'body space', by moving too close to them (e.g. in a queue, outside a cinema, on a beach). Any cultural variations can easily lead to conflict and misinterpretations. Latin Americans, for example, prefer to stand much closer to each other than North Europeans, so that when the former and the latter converse, there may be a problem. The present author recalls one such conflict during a conversation with a student from Brazil, who came and stood before him at some 45 cm distance – a normal interaction distance for her, but much too close for him. He instinctively moved back to the distance he found most comfortable – nearer 1 metre. However, as he did so, the student moved forward, unconsciously maintaining her own norm. He retreated further, not wishing to be so close to the student. After both had circled the desk several times, he capitulated, and asked her to sit down!

THE AMPLIFIED HAND-SHAKE

In a culture where hand-shaking is a normal formality, extra warmth can be expressed only by extra activity, such as increased firmness, longer duration, and more vigorous vertical movements. The second hand may also be brought into play, as shown in the diagrams, which illustrate increasing warmth: (a) hand clasping, (b) arm clasping, (c) shoulder clasping, and (d) shoulder embracing. (From D. Morris, 1977, p. 93.)



TADOMA COMMUNICATION

Tadoma is a method of tactile speech communication that has evolved between people who are both deaf and blind. Speech is perceived by placing a hand against the face of the speaker and monitoring the articulatory movements involved. Usually, the thumb is used to sense the movements of the lips, and the fingers fan out over the side of the face and neck. Devised in Norway in the 1890s, it got its name from its first use in the U.S. with two deaf–blind children, Tad Chapman and Oma Simpson (R. Vivian, 1966.)

Several other tactile methods of communication are used with the handicapped, such as braille (p. 282). It is also possible to 'translate' such codes as morse and finger spelling (p. 227) into tactile form.

DISTANCE ZONES

An American study suggests that there may be four proximity zones when people interact:

- *Intimate* Less than 45 cm, used for intimate relationships.
- *Personal* Between 45 cm and 1.3 metres, for reasonably close relationships.
- *Social consultative* Between 3 and 4 metres, for more impersonal relationships.
- *Public* Above 4 metres, for public figures and public occasions.

(E. T. Hall, 1959.)

The rules of Indian caste (p. 38) illustrate the point even more precisely. According to tradition in one part of India, members of each caste may not approach each other within the following distances:

Brahmins – Nayars: 2 metres
Nayars – Iravans: 8 metres
Iravans – Cherumans: 10 metres
Cherumans – Nayadis: 20 metres

The rules, which are still followed in some areas, work in an additive way: thus, a Nayadi may not come closer to a Brahmin than 40 metres (M. Argyle, 1975).

VISUAL EFFECTS

The field of non-verbal visual communication, kinesics, can be broken down into several components: facial expression, eye contact, gesture, and body posture. Each component performs a variety of functions. Movements of the face and body can give clues to a person's personality and emotional state. The face, in particular, signals a wide range of emotions, such as fear, happiness, sadness, anger, surprise, interest, and disgust, many of the expressions varying in meaning from culture to culture. In addition, the face and body send signals about the way a social interaction is proceeding: patterns of eye contact show who is talking to whom; facial expression provides feedback to the speaker, expressing such meanings as puzzlement or disbelief; and body posture conveys a person's attitude towards the interaction (e.g. relaxation, interest, boredom). Several kinds of social context are associated with specific facial or body behaviours (e.g. waving while taking leave). Ritual or official occasions are often primarily marked by such factors as kneeling, standing, bowing, or blessing.

Visual effects interact very specifically with speech. Gestures and head movements tend to coincide with points of emphasis. Hand movements in particular can be used to add visual meaning to what has been said ('drawing pictures in the air'). Patterns of gaze distinguish the participants in a conversation: a listener looks at a speaker nearly twice as often as the speaker looks at the listener. They also assist in marking the structure of a conversation (§20): for example, speakers tend to look up towards the ends of their utterances, thus giving their listeners a cue that an opportunity to speak is approaching.

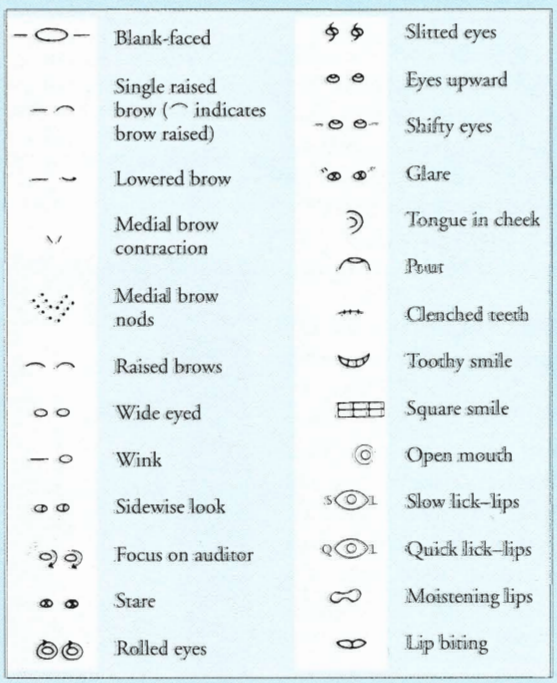
Several visual effects may well be universal, but the focus of interest in recent years has been on the cultural differences that can be observed in face and body movements. Some societies use many gestures and facial expressions (e.g. Italian); others use very few (e.g. Japanese). Moreover, a visual effect may seem to be shared between societies, but in fact convey very different meaning. Thus, in France, using a finger to pull down the eyelid means that the speaker is aware of something going on, whereas in Italy the same gesture means that the listener must *become* aware. Cultural variations in visual effects are among the first things a foreigner notices, but it can be very difficult working out what they mean, and even more difficult deciding whether one is permitted to use them.

EYEBROW FLASHING

When people greet each other at a distance, wishing to show that they are ready to make social contact, they raise their eyebrows with a rapid movement, keeping them raised for about one-sixth of a second. The behaviour has been noted in many parts of the world, and is considered universal (though some cultures suppress it, e.g. the Japanese, who consider it indecent). We are not usually aware that we use this signal, but it evokes a strong response in a greeting situation, and is often reciprocated. To receive an eyebrow flash from someone we do not know is uncomfortable, embarrassing, or even threatening. (After I. Eibl-Eibesfeldt, 1972.) The pictures show an eyebrow flash made by a Samoan (above) and a Yanomami (Waika) Indian (below).

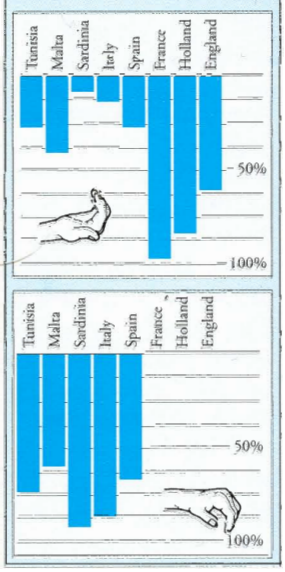
BODY TRANSCRIPTION

Some of the symbols, or *kinegraphs*, which have been used in order to transcribe the various movements of face and body. Different sets of symbols have been devised for different areas of the body: head, face, trunk, shoulder/arm/wrist, hand/fingers, hip/leg/ankle, foot activity, and neck. The symbols below are from the set for facial activities. (From R. L. Birdwhistell, 1952.)



COME HERE?

Beckoning can be carried out with the palm of the hand facing up or down. People used to the former could interpret the latter to mean 'Go away'! The chart shows the preferred pattern in countries between Britain and North Africa. (After D. Morris *et al.*, 1978.)



BEING HUMBLE

Points of similarity as well as difference can be seen in the expression of an attitude among various cultures. In one early study, the communication of humility was found to make use of such body postures as the following:

- Join hands over head and bow (China).
- Extend or lower arms (Europe).
- Stretch arms towards person and strike them together (Congo).
- Crouch (Fiji, Tahiti).
- Crawl and shuffle forward; walk on all fours (Dahomey).
- Bend body downward (Samoa).
- Permit someone to place a foot on one's head (Fundah, Tonga).
- Prostrate oneself, face down (Polynesia).
- Bow, extend right arm, then move it down, up to head, and down again (Turkey, Persia).
- Throw oneself on the back, roll from side to side, and slap outside of the thighs (Batokas).

(After M. H. Krout, 1942.)

EURYTHMY

The bodily representation of the sound *a*, expressing the meaning of astonishment and wonder, as recommended in eurythmy (R. Steiner, 1931, p. 40).

This approach, developed by the founder of anthroposophy, Rudolf Steiner (1861–1925), aimed to promote a close harmony between the sounds of speech and patterns of body movement. Eurythmy was seen as 'visible speech', with the body reflecting in its physical shape the forms of sounds as they are articulated. The different sounds are interpreted symbolically (§30), e.g. *u* is seen as the expression of something which chills or stiffens, and this is shown in the body by a pressing together of the arms and legs. According to Steiner, 'The entire universe is expressed when the whole alphabet is repeated from beginning to end.'

Sign 'language'

Many gestural systems have evolved to facilitate communication in particular situations. They are often referred to as 'sign languages', but few have developed any degree of structural complexity or communicative range, and it is therefore important to distinguish them from 'sign language proper' – the natural signing behaviour of the deaf (Part VI). Several might properly be described as 'restricted languages' (p. 56).

In many parts of the world, such as India, Thailand, and Japan, pantomime and dance have come to use complex systems of symbolic hand gestures in association with facial expressions and body movements. The events of a story, its deeper meaning, and the emotional states of the characters may all be conveyed in this way. For example, in the *Bhārata Nāṭya-śāstra* ('principles of dramatic art'), the 6th-century BC manual of Hindu dance, there are over 4,000 picture patterns for the hands (*mudrās*).

- Religious or quasi-religious groups and secret societies often develop ritual signing systems so that members can recognize and communicate with each other. Such signs are used in Freemasonry, practised by some 6 million people mainly in the USA and Britain, and in many of the secret societies of the Far East, such as the Hung Society.
- Several monastic orders developed signing systems of some complexity, especially if their members were vowed to silence, as in the case of the Trappist monks, a development of the medieval Cistercian order.
- Simple signing systems are found in a wide range of professions:

sports players or officials can signal the state of play, or an intention to act in a certain way.

entertainment a group of performers can coordinate their activities, such as acrobats, musicians.

theatres/cinemas ushers can signal the number and location of seats.

casinos officials can report on the state of play, or indicate problems that might affect the participants in a game.

sales/auctions auctioneers can convey the type and amount of selling and buying.

aviation marshalling ground staff can send information about the position of an aircraft, the state of its engines, and its desired position.

radio/television direction producers and directors can signal to performers the amount of time available, instructions about level of loudness or speed of speaking, and information about faults and corrections.

diving divers can communicate depth, direction, time, and the nature of any difficulties they have encountered.

truck driving drivers can exchange courtesy signals, give information about the state of the road, or show they are in trouble.

heavy equipment drivers people controlling cranes, hoists, and other equipment can signal the direction and extent of movement.

fire service fire-officers can send directions about the supply of water, water pressures, and the use of equipment.

bookmaking bookies send signals about the number of a race or horse, and its price (see left).

noisy conditions environmental noise may make verbal communication impossible (e.g. in cotton mills) and a signing system may result.

TICK-TACK TALK

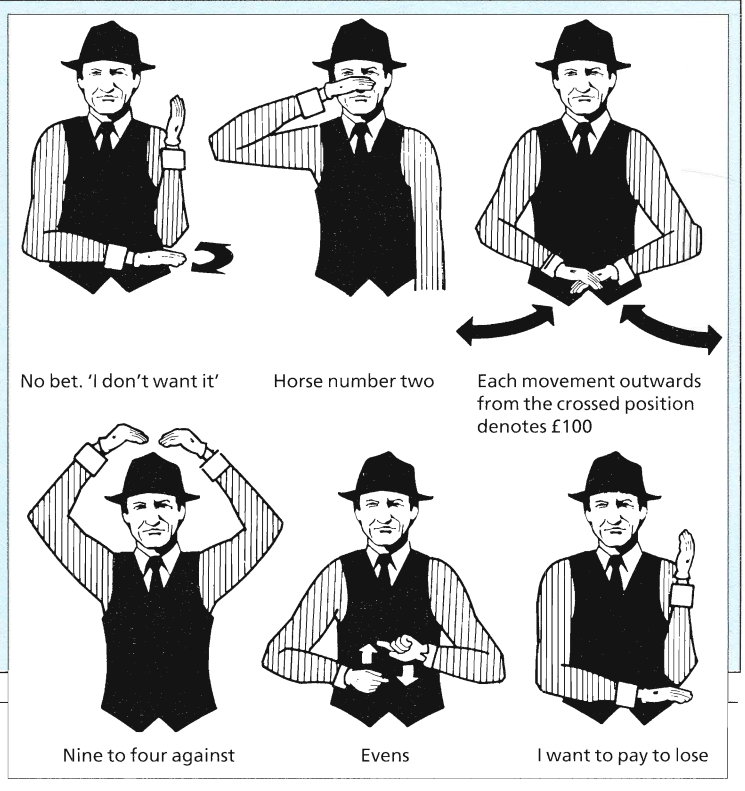
One of the most intriguing sights at dog tracks and racecourses in Britain is the system of tick-tack signing used to circulate information about the way bets are being placed. A signer acts as an agent for a group of bookmakers who have bought his 'twist card', on which the dogs or horses are given different numbers to those on the official race card. The same set of tick-tack signs is used by all signers, identifying the amount of a bet, a horse or dog number, and the number of a race; but only those who have an individual signer's twist card will be able to interpret what a number refers to.

Number signs

1	right hand on top of hat
2	right hand on nose
3	right hand under chin
4	right hand sweeps a curve
5	right hand on shoulder
6	sign 5 then 1
7	sign 5 then 2
8	sign 5 then 3
9	sign 5 then 4
10	clap hands
£5	right hand held up, palm outwards, fingers spread
£10	both arms held up with fingers spread
£50	clenched fists held together
£100	left hand held up with fingers spread
£500	hands outline a circle
£1,000	hands play imaginary piano ('grand piano' = 'a grand' = £1,000)

Some signs for odds

Evens	arms held in front, moving up and down
11/10	hands together, forming a pyramid
6/4	one right finger in the left ear-hole



No bet. 'I don't want it'

Horse number two

Each movement outwards from the crossed position denotes £100

Nine to four against

Evens

I want to pay to lose

OBTAINING LINGUISTIC DATA

Many procedures are available for obtaining data about a language. They range from a carefully planned intensive field investigation in a foreign country to casual introspection about one's mother tongue carried out in an armchair at home.

Informants

In all cases, someone has to act as a source of language data – an *informant,* or *consultant.* Informants are (ideally) native speakers of a language who provide utterances for analysis and other kinds of information about the language (e.g. translations, comments about correctness, or judgments on usage). Often, when studying their mother tongue, linguists act as their own informants, judging the ambiguity, acceptability, or other properties of utterances against their own intuitions. The convenience of this approach makes it widely used, and it is considered a primary datum in the generative approach to linguistics (p. 413). But a linguist's personal judgments are often uncertain, or disagree with the judgments of other linguists, at which point recourse is needed to more objective methods of enquiry, using non-linguists as informants. The latter procedure is unavoidable when working on foreign languages, or in such mother-tongue fields as child speech (§38) or language variation (§§8–12).

Many factors must be considered when selecting informants – whether one is working with single speakers (a common situation when languages have not been described before), two people interacting, small groups, or large-scale samples. Age, sex, social background, and other aspects of identity are important, as these factors are known to influence the kind of language used (Part 11). The topic of the conversation and the characteristics of the social setting (e.g. the level of formality) are also highly relevant, as are the personal qualities of the informants (e.g. their fluency and consistency). For larger studies, scrupulous attention has to be paid to the sampling theory employed. And in all cases decisions have to be made about the best investigative techniques to use.

Recording

Today, data from an informant are often tape recorded. This enables the linguist's claims about the language to be checked, and provides a way of making those claims more accurate ('difficult' pieces of speech can be listened to repeatedly). But obtaining naturalistic, good-quality data is never easy. People talk abnormally when they know they are being recorded, and sound quality can be poor. A variety of tape-recording procedures have thus been devised to minimize the effects of the 'observer's paradox' (how to observe the behaviour of people when they are not being observed). Some recordings are made without the speakers being aware of the fact – a procedure that obtains very natural data, though ethical objections must be anticipated. Alternatively, attempts can be made to make the speaker forget about the recording, such as by keeping the tape recorder out of sight, or using radio microphones. A useful technique is to introduce a topic that quickly involves the speaker, and stimulates a natural language style (e.g. asking older informants to talk about how times have changed in their locality).

An audio tape recording does not solve all the linguist's problems, however. Speech is often unclear or ambiguous. Where possible, therefore, the recording has to be supplemented by the observer's notes about the non-verbal behaviour of the participants, and about the context in general. A facial expression, for example, can dramatically alter the meaning of what is said (p. 406). Video recordings avoid these problems to a large extent, but even they have limitations (the camera can be highly intrusive, and cannot be everywhere), and transcriptions always benefit from any additional commentary provided by an observer (p. 233).

Elicitation

Linguists also make great use of structured sessions, in which they systematically ask their informants for utterances that describe certain actions, objects, or behaviours. With a bilingual informant, or through the use of an interpreter, it is possible to use translation techniques ('How do you say *table* in your language?', 'What does *gua* mean?'). A large number of points can be covered in a short time, using interview worksheets and questionnaires. Often, the researcher wishes to obtain information about just a single variable, in which case a restricted set of questions may be used: a particular feature of pronunciation, for example, can be elicited by asking the informant to say a restricted set of words. There are also several indirect methods of elicitation, such as asking informants to fill the blanks in a substitution frame (e.g. *I – see a car*), or feeding them the wrong stimulus for correction ('Is it possible to say *I no can see?*').

Corpora

A representative sample of language, compiled for the purpose of linguistic analysis, is known as a *corpus.* A corpus enables the linguist to make objective statements about frequency of usage, and it provides accessible data for the use of different researchers. Its range and size are variable. Some corpora attempt to cover the language as a whole, taking extracts from many kinds of text; others are extremely selective, providing a collection of material that deals only with a particular linguistic feature. The size of a corpus depends on practical factors, such as the time available to collect, process, and store the data: it can take up to several hours to provide an accurate transcription of a few minutes of speech (p. 233). Sometimes a small sample of data will be enough to decide a linguistic hypothesis;

A CAUTIONARY TALE

The informant arrived and we started our work. 'How do you say *I run* in your language?' The Indian was quiet for a while. First he looked down; then he looked out. Suddenly his face lit up as if struck by a sudden flash of inspiration. He spoke very rapidly. If I had been able to transcribe what he said, it would have spread across the page several times. I gulped and bravely started to write; but after a few syllables, I was already hopelessly bogged down. 'How did you say that?' With his repetition I added two more syllables, then bogged down again. When I asked for the third repetition, the informant began to waver and finally to change his story, and so I had to give up entirely. To my self-justifying and half self-accusing 'But that surely doesn't all mean just *I run*', he said, 'Why of course not. It means I was sitting here with you; then I looked out of the door and saw a deer, so I quickly grabbed my spear and now I am running after it.' Then, almost philosophically, he added to himself, 'Only a fool would run for nothing.'
(J. A. Loewen, 1964, p. 189.)

A reverse lexicon
An extract from the Brown University Corpus listing words in reverse alphabetical order.

REDEMPTION
EXEMPTION
GUMPTION
RESUMPTION
PRESUMPTION
CONSUMPTION
ASSUMPTION
OPTION
ADOPTION
SORPTION
ABSORPTION
ERUPTION
INTERRUPTION
CORRUPTION
DISRUPTION
DESERTION
INSERTION
ASSERTION
EXERTION
ABORTION
PORTION
PROPORTION

COMPUTER CORPORA

A 'standard' corpus is a large collection of data available for use by many researchers. In English linguistics, there are now three standard computer corpora, all in machine-readable form, and thus, in principle, available anywhere in the world.

The London–Lund Corpus of Spoken English This corpus of educated spoken British English consists of the spoken material collected as part of the Survey of English Usage (see below right). The data consist of 87 texts of 5,000 words each. It was transferred to computer tape in the 1970s at the Survey of Spoken English, University of Lund, and is also partly available in printed form. In addition to the running text, a lexical concordance has been compiled.

The Brown University Corpus of American English This corpus is drawn from U.S. printed sources published in 1961. It comprises 500 samples of about 2,000 words each representing 15 main varieties of the language. It is available via computer tape, printout, and microfiche. Apart from the running text, there are lexical concordances, word frequency lists, and a reverse alphabetical list.

The Lancaster–Oslo/Bergen Corpus of British English This is the British equivalent of the Brown corpus. It was compiled by researchers in the Universities of Lancaster and Oslo, and prepared for computer analysis at the Norwegian Computing Centre for the Humanities in Bergen. Facilities are available similar to those provided by the Brown corpus.

ICAME There is now a clearing centre for storing and distributing information on corpus studies in English: the International Computer Archive of Modern English (ICAME), based at Bergen University. Its aims are to compile an archive of English-language material available for computer processing, and to collect and distribute information on research that uses this material.

by contrast, corpora in major research projects can total millions of running words. An important principle is that all corpora, whatever their size, are inevitably limited in their coverage, and always need to be supplemented by data derived from the intuitions of native speakers of the language, through either introspection or experimentation.

Experiments

Experimental techniques are widely used in linguistics, especially in those fields that have been influenced by the methods of sciences where experimentation is routine. Phonetics (§24) is the subject most involved in this approach, but experimental testing is also common in several other areas, such as child language acquisition (§38) and language pathology (§46). In grammar and semantics, experimental studies usually take the form of controlled methods for eliciting judgments about sentences or the elements they contain. Informants can be asked to identify errors, to rate the acceptability of sentences, to make judgments of perception or comprehension, and to carry out a variety of analytical procedures.

Reconstruction

The limiting case of linguistic study, one might imagine, is when no data are available at all – as in the case of the historical study of language where written records are lacking. But it is possible to break through even this apparent barrier, by using the 'reconstruction' techniques of comparative philology (§50). The forms of Proto-Indo-European and other reconstructed languages may be totally hypothetical in status, but they have nonetheless become a major field of linguistic enquiry.

TAGGING A TEXT

Many of the operations that a computer can perform on a corpus are linguistically trivial, though they save an enormous amount of time (e.g. listing of words in frequency of use or alphabetical order). More interesting is the possibility of automatically analysing the structure of the corpus, from a grammatical, semantic, or phonological point of view (§26). This is the aim of several current research programmes.

A first step is to provide an automatic means of 'tagging' each word in the corpus with a label that indicates its word class (§16). This enables the user to distinguish between such superficially identical items as *bear* (animal) and *bear* (action), or the many different syntactic functions of *that*. Larger constructions (such as different kinds of clause, p. 95) can also be tagged, to facilitate the retrieval of grammatical information.

Two tagged sentences from the London–Lund corpus are given (from J. Svartvik *et al.*, 1982, p. 57). Abbreviations are as follows (other symbols refer to suprasegmental features of pronunciation, §29):

CD	*that* used as subordinator
NP	proper noun
RA	personal pronoun, subject
RN	personal pronoun, object
VA + D	main verb, past tense
VA + G	main verb, *-ing* form
VA + N	main verb, past participle
VB + 5	*was* form of *to be*
*VH + O	contracted form of *have*, present tense
VA + O	main verb, base form

(See further, J. Svartvik & R. Quirk, 1980, from which the classification of Survey of English Usage texts (right) has been taken.)

0101000563 B I⟨RA⟩ II knew ⟨VA + D⟩ that ⟨CD⟩ he ⟨RA⟩ was ⟨VB + 5⟩ c*oming■ ⟨VA + G⟩

0101000564 B I've ⟨RA★VH + O⟩ II heard ⟨VA + N⟩ Stan ⟨NP⟩ !Carter ⟨NP⟩ m*ention ⟨VA + O⟩ him■ ⟨RB⟩

THE SURVEY OF ENGLISH USAGE

This survey, begun in London in 1960 by the British linguist Randolph Quirk (1920–), aims to describe the grammatical repertoire of adult educated native speakers of British English. The corpus comprises 200 texts of spoken or written material, classified as follows (figures refer to the number of texts of each type):

Origin in writing (100)

Printed (46)
Learned arts (6)
Learned sciences (7)
Instructional (6)
Press: general news (4)
Press: specific reporting (4)
Administrative/official (4)
Legal and statutory (3)
Persuasive writing (5)
Prose fiction (7)

Non-printed (36)
Continuous writing: imaginative (5), informative (6)
Social letters: intimate (6), equal (3), distant (4)
Non-social letters: equal (4), distant (4)
Personal journals (4)

As spoken (18)
Drama (4)
Formal scripted oration (3)
Broadcast news (3)
Talks: informative (4), imaginative (2)
Stories (2)

Origin in speech (100)

Monologue (24)
Prepared (but unscripted) oration (6)
Spontaneous oration (10)
Spontaneous commentary: sport (4), non-sport (4)

Dialogue (76)
Surreptitious: intimate (24), distant (10)
Non-surreptitious: intimate (20), distant (6)
Telephone: intimate (10), distant (6)

I GLOSSARY

This glossary contains a brief definition of all the specialized language terms used in the text of this encyclopedia, along with some of the associated linguistic terminology likely to be encountered by the general reader. The glossary excludes four types of term:
(i) words in everyday use that do not raise any particular problem of meaning (such as the names of punctuation marks); (ii) names of different theories and approaches (as in linguistics and language teaching); (iii) the very detailed terminology of grammatical description and particular schools of thought; and (iv) background terms from related disciplines, such as anatomy, acoustics, or medicine. Names of languages, language families, dialects, and scripts are given in Appendix VI. A selection of more specialized dictionaries of linguistic terms is given at the end of Appendix IV.

Glossary conventions

- The alphabetical arrangement of the glossary is letter by letter.
- Each head-word is followed in parentheses by an abbreviated indication of the main sub-field to which it belongs (e.g. *sem* = *semantics*). The abbreviations used are given to the right.
- Within entries, words or phrases that are themselves defined elsewhere in the glossary are preceded by +. Superscript numerals are used when it is important to distinguish a particular sense within cross-references (e.g. grammar[1]).
- Synonymous terms are given in bold type, preceded by the word 'also'.
- Most entries lack exemplification, as this can be found within the body of the encyclopedia; in a few cases, where the main text does not provide sufficient illustration, examples are given in parentheses, without the use of 'e.g.'.
- At the end of each entry, there is a page reference to a section of the encyclopedia where related subject matter may be found.

Abbreviations used

acou	*acoustics*	Lat.	Latin
anat	*anatomy*	*ling*	*general linguistics*
app	*applied linguistics*	*neuro*	*neurolinguistics*
clin	*clinical*	*phonet*	*phonetics*
E.	English	*phonol*	*phonology*
esp.	especially	*phys*	*physiology*
Fr.	French	*poet*	*poetics*
gen	*general application*	*prag*	*pragmatics*
Ger.	German	*psycho*	*psycholinguistics*
gram	*grammar*	*rhet*	*rhetoric*
graph	*graphetics/graphology*	*sem*	*semantics*
hist.	*historical linguistics*	*semiot*	*semiotics*
It.	Italian	*socio*	*sociolinguistics*
J.	Japanese	*styl*	*stylistics*

abessive (*gram*) An +inflection[1] that typically expresses the meaning of 'without'. 92

ablative (*gram*) An +inflection[1] that typically expresses such meanings as 'by/with/from'. 92

ablaut (*hist*) A +vowel change that gives a word a new grammatical function (*drink* → *drank*); also, **gradation**. 299

abstract *see* **concrete**

accent **1** (*phonet*) Features of pronunciation that signal regional or social identity; cf. +dialect. 24 **2** (*phonol*) A type of emphasis given to a spoken word or syllable. 166 **3** (*graph*) A mark above a letter, showing its pronunciation. 196

acceptable (*ling*) Said of any usage that +native speakers feel is possible in a language. 414

accidence (*gram*) Changes in the +form[2] of words signalling different grammatical functions (*walking/walked...*); cf. +morphology. 90

accommodation (*socio*) Adjustments that people make to their speech, influenced by the speech of those they are talking to. 51

accusative (*gram*) An +inflection[1] that typically identifies the +object of a +verb; also, **objective**. 92

acoustic phonetics (*phonet*) The branch of +phonetics that studies the physical properties of speech sounds. 132

acquired (*clin*) Said of any linguistic disorder that results from injury or disease; cf. +developmental. 273

acquisition *see* **language acquisition**

acrolect (*socio*) In +creole studies, the most prestigious +variety of a language, seen in contrast with other varieties. 24

acronym (*gen*) A word made up out of the initial letters of a phrase (*laser*). 90

acrostic (*gen*) A poem or other text in which certain letters in each line make a word. 64

active **1** (*gen*) Said of language that a person actually uses – as opposed to language that is known but not used (**passive** knowledge). 378 **2** (*phonet*) Said of an +articulator that moves (towards an immobile, **passive**, articulator). 130

active voice *see* **voice**

acuity (*phonet*) The ability to detect and discriminate sound. 145

adessive (*gram*) An +inflection[1] that typically expresses the meaning of 'on' a place. 92

adjacency pair (*socio*) A single sequence of stimulus-utterance +response-utterance by two different speakers, e.g. question + answer. 118

adjective (*gram*) A type of word identifying an attribute of a +noun (*a* ***red*** *chair*), in many languages showing +degree contrasts. 91

adjunct (*gram*) A less important or omissible element in a grammatical construction (*She ran* ***quickly***). 95

adnominal (*gram*) Any element in a +noun phrase that is a +modification[1] of the noun. 95

adverb (*gram*) A word whose main function is to specify the kind of action expressed by a +verb (*He spoke* ***angrily***); other functions include acting as +intensifier (***very*** *big*) and as a +sentence connector (***Moreover***, *they laughed*). 91

adverbial (*gram*) Said of +words, +phrases, or +clauses that function as +adverbs. 95

aerometry (*phonet*) The measurement of air flow during speech. 139

affective (*sem*) Said of the emotional or attitudinal meaning of an utterance. 103

affirmative (*gram*) A +sentence or +verb that has no marker of +negation (*He's running*). 95

affix (*gram*) A meaningful form that is attached to another form, to make a more complex +word (***un-*** + *kind* + ***-ness***); cf. +infix, +prefix, +suffix. 90

affixing language (*ling*) A language that uses +affixes as its main way of expressing grammatical relationships. 295

affricate (*phonet*) Said of a +consonant in which a complete +closure of the +vocal tract is gradually released (cf. Ger. *pfennig*). 159

agent(ive) (*sem*) A linguistic form expressing who or what is responsible for an action (***The man*** *laughed*, ***farmer*** 'one who farms'). 93

agglutinative/agglutinating language (*ling*) A type of language in which +words consist of lengthy strings of forms. 295

agnosia (*clin*) Loss of ability to interpret sensory information: **auditory agnosia**, affecting speech sounds. 273

agrammatism (*clin*) A language disorder that produces speech of a typically +telegrammatic quality (*man see ball*). 273

agraphia *see* **dysgraphia**

agreement *see* **concord**

air-stream mechanism (*phonet*) An arrangement of parts of the +vocal tract that acts as a source of energy for speech sound production. 124

alaryngeal (*clin*) Said of speech without the +larynx. 278

alexia *see* **dyslexia**

alienable (*gram*) Applied to relationships where a possessed item is seen as having a temporary or non-essential dependence on a possessor (*the*

man's car); cf. **inalienable**, where the dependence is permanent or necessary (*the man's brain*). 93

allative (*gram*) An ⁺inflection[1] that typically expresses the meaning of 'to' a place. 92

alliteration (*poet*) A sequence of words beginning with the same sound, especially as used in poetry. 74

allo- (*ling*) A variation in the ⁺form[2] of a linguistic unit that does not alter its basic identity, e.g. **allophones** (variants of a ⁺phoneme), **allomorphs** (variants of a ⁺morpheme), **allographs** (variants of a ⁺grapheme). 90, 162, 196

allograph *see* **allo-**

allomorph *see* **allo-**

allonym (*gen*) A name an author assumes that belongs to someone else; cf. ⁺pseudonym. 112

allophone *see* **allo-**

alphabet (*gen*) A writing system in which a set of symbols ('letters') represents the ⁺phonemes of a language; cf. ⁺dual alphabet. 204

alphabetism (*gram*) A word made of initial letters, each being pronounced (*VIP*). 90

alternation (*ling*) The relationship between the different ⁺forms[2] of a linguistic unit, usually symbolized by ~ (*cat ~ cats*). 90

alveolar (*phonet*) Said of a ⁺consonant in which the tongue makes contact with the bony prominence behind the upper teeth (t, n). 157

ambilingual (*gen*) Someone who can speak two languages with equal facility; also, **balanced bilingual**. 364

amelioration (*hist*) A change of meaning in which a word loses an originally unpleasant reference; cf. ⁺deterioration. 332

amplitude (*acou*) The intensity of a sound. 134

anacoluthon (*gram, rhet*) An unexpected break in a ⁺sentence (*John might – Are you listening?*). 52

anacusis (*clin*) Total deafness. 268

anagram (*gen*) A word or phrase formed by changing the order of letters in another word or phrase. 65

analects (*gen*) A selection of passages taken from an author. 66

analogy (*ling*) A change that affects a language when ⁺regular forms begin to influence less regular forms. 236, 332

analytic **1** (*gram*) *see* **isolating** **2** (*sem*) Said of sentences expressing a ⁺tautology (*Bachelors are unmarried*); contrasts with **synthetic**. 107

ananym (*gen*) A name that has been written backwards. 112

anap(a)est (*poet*) A unit of ⁺metre consisting of two light beats followed by a heavy beat. 74

anaphora (*gram*) A feature of grammatical structure referring back to something already expressed; the ⁺pronoun in *When Mary saw John, **she** waved* is 'anaphoric'; cf. ⁺cataphora, ⁺exophoric. 119

anarthria *see* **dysarthria**

animate (*gram*) Said of words (esp. ⁺nouns) that refer to living things, and not to objects or concepts (**inanimates**). 91

anomia (*clin*) A ⁺language[4] disorder in which the primary symptom is difficulty in remembering the names of things. 273

antecedent (*gram*) A part of a ⁺sentence to which some other part grammatically refers (*This is **the cat** that chased the rat*). 119

anthropological linguistics (*ling*) The study of (esp. non-western) languages in relation to social or cultural patterns and beliefs. 418

anthroponomastics (*sem*) The study of personal names. 112

anthropophonics (*phonet*) The study of the human potential for vocal sound. 18

anticipatory *see* **regressive**

antonym (*sem*) A word that is opposite in meaning to another word (*good/bad, single/married*). 105

aorist (*gram*) A form of the ⁺verb in some ⁺inflecting languages, esp. referring to an action without any particular completion, duration, or repetition. 93

aperiodic *see* **periodic**

apex (*phonet*) The tip of the tongue. 131

aphasia (*clin*) A ⁺language[4] disorder resulting from brain damage, which affects a person's ability to produce or understand ⁺grammatical and ⁺semantic structure; also, **dysphasia**. 272

aphasiology (*clin*) The study of ⁺aphasia. 272

aphesis (*hist*) The loss of an ⁺unstressed ⁺vowel from the beginning of a word ('*mongst*). 330

aphonia *see* **dysphonia**

aphorism (*gen*) A succinct statement expressing a general truth (*More haste, less speed*). 53

apico- (*phonet*) Said of a sound using the tip (or ⁺apex) of the tongue, e.g. 'apico-dental'. 157

apocope (*hist*) The omission of a final ⁺syllable, sound, or letter in a word. 330

apostrophe (*rhet*) A ⁺figurative expression in which an idea, inanimate object, or absent person is addressed as if present. 70

appellative (*sem*) A personal name used as an everyday word (*a sandwich*). 112

applied linguistics (*ling*) The application of the theories, methods, or findings of ⁺linguistics to the solution of practical problems. 418

apposition (*gram*) A series of ⁺nouns or ⁺noun phrases with the same meaning and grammatical status (*Mr Jones, the baker*). 95

appropriate (*ling*) Said of any use of language considered to be compatible with a given social situation; cf. ⁺correctness. 2

approximant (*phonet*) A ⁺consonant in which the organs of ⁺articulation approach each other, but without ⁺closure or audible friction (ɻ, j); also, **frictionless continuant**. 159

approximative system *see* **interlanguage**

apraxia (*clin*) Loss of ability to carry out voluntary muscular movements for the production of speech; also, **dyspraxia**. 273

aprosody *see* **dysprosody**

aptitude (*app*) A person's natural ability to learn a language; evaluated using an **aptitude test**; also, **prognostic test**. 375

aptronym (*gen*) A name that fits a person's nature or occupation (*Mr Clever, Mr Smith*). 112

arbitrariness (*ling*) The absence of any physical correspondence between linguistic signals and the entities to which they refer; cf. ⁺iconic. 401

archaism (*gen*) An old word or phrase no longer in general spoken or written use. 332

area (*ling*) A geographical region identified on the basis of its linguistic characteristics. 33

areal linguistics (*ling*) The study of geographical regions which are characterized by shared linguistic properties; cf. **geographical linguistics**. 33

argot (*gen*) Special vocabulary used by a secretive social group, e.g. gypsies. 58

article (*gram*) A word that specifies whether a ⁺noun is ⁺definite or indefinite (*the/a*). 91

articulation (*phonet*) The physiological movements involved in modifying a flow of air to produce speech sounds. 130

articulator (*phonet*) A ⁺vocal organ involved in the production of a speech sound. 130

articulatory phonetics (*phonet*) The branch of ⁺phonetics that studies the way speech sounds are produced by the ⁺vocal organs. 124

artificial language (*gen*) **1** An invented language used to facilitate international communication; also, **auxiliary language**. 354 **2** An invented language used in computer programming, e.g. BASIC. 353

artificial larynx (*clin*) A portable device that provides a source of vibration for speech, for people who have no ⁺larynx. 278

artificial speech (*phonet*) The output of a ⁺speech synthesizer. 149

ascender (*graph*) A part of a letter that extends above the height of the letter *x*. 192

aspect (*gram*) The duration or type of temporal activity denoted by a ⁺verb, e.g. completion or non-completion of an action; cf. ⁺perfective. 93

aspiration (*phonet*) Audible breath that may accompany the ⁺articulation of a sound (E. *pen* pʰɛn). 163

assimilation (*phonol*) The influence exercised by one sound upon the ⁺articulation of another, so that the sounds become more alike. 166

associative meaning (*sem*) The sense associations that are not part of a word's basic meaning (*birthday* → presents, party, etc.). 103

assonance (*poet*) The repeated use of ⁺vowels to achieve a special effect. 74

asterisked form **1** (*ling*) A usage that is not ⁺acceptable or not ⁺grammatical[2] (**do have gone*). 88 **2** (*hist*) A form in linguistic history for which there is no written evidence (Indo-European **penkʷe* 'five'). 294

asyndeton (*rhet*) The omission of ⁺conjunctions to achieve an economical form of expression (*They ran with haste, with fear*). 91

atelic *see* **telic**

attested (*ling*) Said of linguistic forms where there is evidence of present or past usage. 294

attribute **1** (*phonet*) An identifiable feature of sound sensation, e.g. ⁺pitch, ⁺loudness. 144 **2** (*sem*) A defining property of the meaning of a word (*round* is an attribute of *ball*). 107

attributive (*gram*) Said of ⁺adjectives or other forms that are ⁺modifiers of a ⁺noun within the ⁺noun phrase (*the **big** table*); contrasts with predicative uses (*The table is **big***). 95

audiogram (*clin*) A graph used to record a person's ability to hear ⁺pure tones. 268

audiolingual (*app*) Said of a language-teaching method based on the use of drills and dialogues for speaking and listening; also, **aural–oral**. 378

audiology (*clin*) The study of hearing and hearing disorders, esp. their diagnosis, assessment, and treatment. 268

audiometer (*clin*) An electronic instrument that measures the sensitivity of hearing. 268

auditory agnosia *see* **agnosia**

auditory discrimination (*phonet*) The process of distinguishing between (esp. speech) sounds. 145

auditory phonetics (*phonet*) A branch of ⁺phonetics that studies the way people perceive sound. 142

aural-oral *see* **audiolingual**

automatic translation *see* **machine translation**

autonomous speech *see* **idioglossia**

autosegmental (*phonol*) An approach to ⁺phonology that includes the study of features of sound that extend beyond individual ⁺segments. 163

auxiliary language 1 (*socio*) A language adopted by different speech communities for purposes of communication. 354 2 (*gen*) *see* **artificial language**

auxiliary verb (*gram*) A ⁺verb used along with a ⁺lexical verb to make grammatical distinctions (*She **is** going/**might** go*). 91

baby talk (*gen*) 1 A simplified speech style used by adults to children. 237 2 An immature form of speech used by children. 246

back (*phonet*) Said of sounds made in the back part of the mouth ([ɑ]) or with the back part of the tongue ([k], [ɔ]). 131

back-formation (*hist*) A process of ⁺word formation where a new word is formed by removing an imagined ⁺affix from another word (*editor* → *edit*). 332

back slang (*gen*) A secret language in which words are said backwards. 59

balanced bilingual *see* **ambilingual**

basal readers (*app*) The first textbooks used in a graded reading programme. 253

base (*ling*) A component of a ⁺transformational grammar, in which the basic sentence patterns of a language are ⁺generated. 97

basilect (*socio*) In ⁺creole studies, a language ⁺variety furthest away from the one that carries most prestige (the ⁺acrolect). 24

behaviourism (*gen*) The study of observable and measurable behaviour (here, of the linguistic stimuli and responses made by participants in speech situations). 412

bel (*acou*) Unit for the measurement of acoustic intensity; cf. ⁺decibel. 134

bidialectal (*socio*) Applied to someone who is proficient in the use of two ⁺dialects. 24

bidialectism (*socio*) An educational policy that recommends the teaching of a non-standard ⁺dialect along with a ⁺standard one. 26

bilabial (*phonet*) Said of a ⁺consonant made with both lips ([ʬ], [m]). 157

bilingual (*gen*) Said of an individual or a community that regularly uses two languages; cf. ⁺ambilingual. 362

binary (*ling*) Said of any linguistic analysis that sets up an opposition between two alternatives. 79

binary feature (*phonol*) Any ⁺phonetic variable that enables sounds to be classified into two mutually-exclusive possibilities, e.g. ⁺voice[1] ('voiced' vs 'voiceless'). 164

binaural (*phonet*) Using both ears. 142

biolinguistics (*ling*) The study of the biological preconditions for language development and use in human beings, both as individuals and as a race; also, **biological linguistics**. 418

bisyllable (*phonet*) A word with two ⁺syllables. 166

blade (*phonet*) The part of the tongue between the ⁺apex and the ⁺centre; also, **lamina**. 131

blend (*gram*) The result of two elements fusing to form a new word or construction (*breakfast* + *lunch* = *brunch*); cf. ⁺coinage. 90

block (*clin*) In ⁺stuttering, an obstruction experienced by the speaker that prevents the production of speech. 280

body language (*semiot*) Communication using body movement and appearance, as opposed to speaking, writing, or ⁺sign[3]. 403

body size (*graph*) The size of a piece of type. 192

borrow (*hist*) To introduce a word (or some other linguistic feature) from one language or ⁺dialect into another; vocabulary borrowings are usually known as **loan words**. 332

bound form (*gram*) A ⁺morpheme that cannot occur on its own as a ⁺word (E. *de-*, *-tion*). 90

boustrophedon (*graph*) Writing in which lines run in alternate directions. 187

brachygraphy (*graph*) Shorthand writing. 208

bracketing (*ling*) A way of showing the internal structure of a string of elements ((*The girl*) (*ate*) (*a cake*)). 97

breaking *see* **voice mutation**

breath group (*phonet*) A stretch of utterance produced within a single breath expiration. 124

breathy (*phonet*) A ⁺voice quality that involves the use of audible breath. 128

broad (*phonet*) Said of a ⁺transcription of speech that shows only the major ⁺phonetic contrasts; cf. ⁺narrow, ⁺phonemic transcription. 160

Broca's area (*neuro*) An area of the brain that controls the expression of spoken language; cf. ⁺Wernicke's area. 262

buccal (*phonet*) Applied to sounds made in or near the ⁺cavity of the cheek. 127

cacography (*gen*) Bad handwriting or spelling. 276

cacology (*gen*) Unacceptable pronunciation or use of language. 2

cacophony (*gen*) Unpleasant, harsh sounds, esp. of speech. 2

caesura (*poet*) A break in the ⁺rhythm of a line of poetry. 74

calligraphy (*gen*) The art of beautiful handwriting. 190

calque (*hist*) A ⁺borrowed item in which the parts are translated separately into the new language (E. *superman* from Ger. *Übermensch*); also, **loan translation**. 332

cant (*gen*) The special speech of a group with low social standing, e.g. thieves. 58

cardinal number (*gram*) The basic form of a numeral (*one*, etc.); cf. ⁺ordinal. 99

cardinal vowels (*phonet*) A set of reference points, based on auditory and articulatory criteria, used to identify ⁺vowels. 156

caretaker speech (*psycho*) The speech of adults when they talk to children; also, **motherese**. 238

case (*gram*) In an ⁺inflecting language, the form of a ⁺noun, ⁺adjective, or ⁺pronoun, showing its grammatical relationship to other words. 93

catachresis *see* **malapropism**

catalect (*gen*) Any part of an author's literary work seen as separate from the rest. 66

cataphora (*gram*) A feature of grammatical structure that refers forward to another unit; (in *John said **this***, the ⁺pronoun is 'cataphoric'); cf. ⁺anaphora and ⁺exophoric. 119

catenation (*ling*) The linking together of a series of linguistic forms, e.g. sounds or words. 95

catenative (*gram*) A ⁺lexical verb that governs another lexical verb (*try to run*). 91

causative (*gram*) A linguistic element that expresses the notion of 'cause' (the causative verb *kill* = 'cause to die'). 93

cavity (*phonet*) An anatomically defined chamber in the ⁺vocal tract, e.g. oral, nasal. 124

central *see* **centre**

centre (*phonet*) The top part of the tongue, between ⁺front and ⁺back; involved in **central** sounds. 131

centum language (*hist*) An Indo-European language that kept the sound [k] in such words as *centum* ('hundred'); cf. ⁺satem language. 330

channel (*gen*) A medium selected for communication (e.g. speech, writing). 48

character (*graph*) A graphic sign used in a writing system, esp. one that is not part of an ⁺alphabet. 202

chereme (*ling*) The smallest contrastive unit in a ⁺sign language. 223

cherology (*ling*) The study of ⁺sign language. 223

chest pulse (*phonet*) A contraction of the chest muscles that forces air into the ⁺vocal tract. 166

chiasmus (*rhet*) A balanced pattern in which the main elements are reversed. 70

chirography (*graph*) The study of handwriting forms and styles. 188

chrestomathy (*gen*) An anthology of passages usually used for learning a language. 378

chroneme (*phonol*) An abstract unit that accounts for differences in the ⁺duration of speech sounds, e.g. long vs short ⁺consonants. 412

chronogram (*gen*) A phrase or sentence in which letters that are also Roman numerals (e.g. C, X) combine to form a date. 64

chunking (*psycho*) Dividing an utterance into parts, e.g. to make it easier to remember. 173

cipher (*gen*) A secret ⁺code[1] in which letters are transposed or substituted. 58

circumlocution (*gen*) The use of more words than is necessary to express a meaning. 2

class *see* **word class**

classifier (*gram*) A ⁺morpheme which indicates that a word belongs to a particular ⁺semantic class, e.g. animates, large objects. 91

clause (*gram*) A structural unit smaller than the ⁺sentence but larger than ⁺phrases or ⁺words; cf. ⁺dependent, ⁺main clause. 95

clavicular breathing (*clin*) A way of breathing, in which inhalation comes from using the neck muscles to raise the collar bones. 125

cleft palate (*clin*) A congenital fissure in the middle of the ⁺palate, often found along with a split in the upper lip (**cleft lip**, also 'hare lip') and teeth ridge. 279

cleft sentence (*gram*) A sentence in which a single ⁺clause has been split into two sections, each with its own ⁺verb (*It was Mary who arrived*). 95

cliche (*gen*) An expression which has become so overused that it no longer conveys much meaning, and is criticized (*a fate worse than death*). 2

click (*phonet*) A sound produced using the ⁺velaric ⁺air-stream mechanism (E. ʇʇ 'tut'). 126

clinical linguistics (*ling*) The application of linguistics to the analysis of disorders of spoken, written, or ⁺sign language. 418

clipping (*gram*) A process of ⁺word formation in which a new word is produced by shortening (*examination* → *exam*); also, **reduction**. 90

clitic (*gram*) A form that resembles a ⁺word but that cannot stand on its own as a normal utterance because it is structurally dependent on a neighbouring word (Fr. *je*). 91

close (*phonet*) Said of a ⁺vowel made with the tongue in the highest position possible without causing audible friction (e.g. i, u); vowels a degree lower are **half/mid-close**; cf. ⁺open[3]. 153

closed 1 (*gram*) Said of any ⁺word class whose membership is limited to a small number of items, e.g. ⁺pronouns, ⁺conjunctions; cf. ⁺open[1]. 91 2 (*phonol*) Said of a ⁺syllable ending in a ⁺consonant; cf. ⁺open[2]. 166

closure (*phonet*) A contact made between ⁺vocal organs in order to produce a speech sound. 159

cloze procedure (*app*) A technique used in the teaching and testing of reading, in which readers guess words omitted at intervals from a text. 381

cluster (*phonol*) A series of adjacent ⁺consonants occurring at the beginning or end of a ⁺syllable (*stray, books*). 166

cluttering (*clin*) A ⁺speech disorder in which utterances are produced in an excessively rapid and unrhythmical way. 280

coalescence (*hist*) The fusing of originally distinct linguistic units. 330

coarticulation (*phonet*) An ⁺articulation involving the simultaneous or overlapping use of more than one point in the ⁺vocal tract (k͡p, t͡ɕ). 158

cochlea (*anat*) The part of the inner ear that contains the organ of hearing. 143

code 1 (*gen*) Any system of signals used for sending messages, often in secret form. 58 2 (*socio*) A language, or ⁺variety of language. 48

code switching (*socio*) Changing from the use of one language or ⁺variety to another; also, **language mixing**. 365

codify (*app*) To provide a systematic account of a language (esp. its ⁺grammar[1] and vocabulary). 366

cognate (*hist*) A language or linguistic form that is historically derived from the same source as another, e.g. Spanish and French are 'cognate languages', both deriving from Latin. 294

cognitive meaning *see* **denotation**

coherence (*ling*) The underlying logical connectedness of a use of language. 119

cohesion (*ling*) The ⁺formal[1] linkage between the elements of a ⁺discourse or ⁺text (the ⁺pronoun is 'cohesive' in *The man left. **He**...*). 119

coinage (*gen*) The creation of a new word out of existing elements (*postperson*); cf. ⁺blend. 90

collective noun (*gram*) A ⁺noun that denotes a group of entities (*army, government*). 91

collocation (*sem*) The habitual co-occurrence (or mutual **selection**) of ⁺lexical items. 105

coloratura (*gen*) A soprano singer with a high vocal range. 18

comitative (*gram*) An ⁺inflection[1] that typically expresses the meaning 'with'. 92

command (*gen*) A type of ⁺sentence in which someone is told to do (or not do) something. 121

comment (*ling*) Part of a ⁺sentence that says something further about the sentence ⁺topic (*The car **was in the garage***); also, **new** information. 94

comment clause (*gram*) A ⁺clause that adds a parenthetic remark to another clause (*The answer, **you see**, is complicated*). 52

common noun (*gram*) A ⁺noun that refers to a class of objects or concepts (*chair, beauty*); cf. ⁺proper noun. 91

communicative approach (*app*) An approach to language teaching that focuses on language ⁺functions[2] and ⁺communicative competence, and not on ⁺grammatical[1] structure. 378

communicative competence (*ling*) A person's awareness of the ⁺rules[1] governing the ⁺appropriate use of language in social situations. 48

comparative *see* **degree**

comparative linguistics (*ling*) A branch of ⁺linguistics that relates the characteristics of different languages or ⁺varieties. 84

comparative method (*hist*) A technique that compares forms taken from ⁺cognate languages to see if they are historically related. 294

comparative philology (*hist*) The study of the historical relationship between languages. 294

compensation (*phonet*) An alternative ⁺articulation that counteracts the effect of some abnormality in the ⁺vocal organs. 18

competence (*ling*) Unconscious knowledge of the system of ⁺grammatical[1] ⁺rules[1] in a language; cf. ⁺communicative competence, ⁺performance. 413

complement (*gram*) A ⁺clause element that completes what is said about some other element, such as the ⁺subject (*That book looks **nice***). 95

complementarity (*sem*) A type of oppositeness of meaning; two words are **complementaries** if to assert one denies the other (*single/married*). 105

complementary distribution (*phonol*) A property of sounds that cannot appear in the same ⁺phonetic ⁺environment[1] (E. tʰ and t). 163

complex sentence (*gram*) A ⁺sentence consisting of more than one ⁺clause (esp. including a ⁺dependent clause). 95

complex tone (*acou*) A sound wave consisting of two or more ⁺pure tones. 133

component 1 (*ling*) The major sections of a ⁺generative grammar. 82 2 (*sem*) A basic feature of word meaning (*girl* = human, female, etc.). 107

componential analysis (*sem*) The analysis of vocabulary into a finite set of basic elements (⁺components[2]). 107

compound 1 (*ling*) Said of a linguistic unit composed of elements that can function separately elsewhere, e.g. a compound ⁺word/⁺sentence. 90 2 (*socio*) Said of ⁺bilinguals who are thought to have a single meaning system underlying their use of words in both languages; cf. ⁺coordinate[2]. 364

comprehension (*gen*) The ability to understand and interpret language; cf. ⁺production. 263

compressed speech (*phonet*) Speech that has been acoustically altered so that it uses a smaller range of ⁺frequencies than normal. 138

computational linguistics (*ling*) The application of the concepts and techniques of computer science to the analysis of language. 418

computer language *see* **language[2]**

concatenation *see* **catenation**

concord (*gram*) A ⁺grammatical[1] relationship in which the ⁺form[2] of one element requires the corresponding form of another (***She** eats*). 95

concordance (*gen*) An ordered list of the words used in a particular text or ⁺corpus. 415

concrete 1 (*gram*) Said of ⁺nouns that refer to physical entities (*book, train*); contrasts with **abstract**. 91 2 (*phonol*) Said of analyses that emphasize the ⁺phonetic reality of speech sounds; contrasts with **abstract**. 165

conditional (*gram*) 1 Said of a ⁺clause that expresses a hypothesis or condition (***If it rains**, **you'll get wet***). 95 2 Said of a ⁺verb form that expresses hypothetical meaning (Fr. 'conditional tense' *je march**erais*** 'I would walk'). 93

conditioning (*ling*) The influence of linguistic ⁺context[1] on a ⁺form[2] (E. *a* → *an* when followed by a ⁺vowel). 166

conductive (*clin*) Said of a hearing loss where sound fails to reach the ⁺cochlea. 268

conjugation (*gram*) The set of ⁺verbs that occur in the same forms in an ⁺inflecting language. 295

conjunction (*gram*) A word that connects words or other constructions (*cat **and** dog*). 91

connective/connector (*gram*) An item whose function is to link linguistic units, e.g. ⁺conjunctions, certain ⁺adverbs (*however*). 91

connotation (*sem*) The personal associations aroused by words; cf. ⁺denotation. 103

consonance (*poet*) The repetition of sounds in the same position in a sequence of words. 74

consonant (*phonol*) A speech sound that functions at the ⁺margins of ⁺syllables, produced when the ⁺vocal tract is either blocked or so restricted that there is audible friction (ʃ, s, etc.); cf. ⁺vowel, ⁺semi-vowel. 157

constative (*ling*) An utterance that is a descriptive statement, analysable into truth values (*The table is red*); cf. ⁺performative. 121

constituent (*gram*) A linguistic unit that is a component of a larger construction. 96

constituent analysis (*gram*) A process of analysing a construction into its major components (**immediate constituents**), each component being further analysed until a set of irreducible elements is left (**ultimate constituents**). 96

constriction (*phonet*) A narrowing in the ⁺vocal tract, in order to produce a speech sound. 159

contact (*socio*) Said of languages or ⁺dialects in close geographical or social proximity, which thus influence each other. 33

content word (*gram*) A type of word that has an independent, 'dictionary' meaning (*chair, run*); cf. ⁺function word. 91

context (*ling*) **1** The linguistic environment of an element. 82 **2** The non-linguistic situation in which language is used. 100

continuant (*phonet*) A speech sound made with an incomplete ⁺closure of the ⁺vocal tract. 159

continuous *see* **progressive**[1]

contoid (*phonet*) A ⁺consonant defined solely in ⁺phonetic terms. 153

contour (*phonol*) **1** A distinctive sequence of ⁺prosodic features (esp. ⁺tones[1]). 169 **2** Said of a ⁺tone language that uses ⁺gliding tones. 174

contraction 1 (*gram*) A shortened linguistic ⁺form[2] attached to an adjacent form (*I'm*), or a ⁺fusion of forms (Fr. *de le* → *du*). 166 **2** (*poet*) The ⁺elision of ⁺syllables to keep a line's ⁺metre regular. 74

contradictory *see* **complementarity**

contrary *see* **antonym**

contrast (*ling*) Any ⁺formal[1] difference that serves to distinguish meanings in a language; **contrastive** differences are also known as **distinctive, functional, significant**. 162

contrastive *see* **contrast**

contrastive analysis (*app*) The identification of structural differences between languages, seen as points of potential learning difficulty. 376

contrastive stress (*phonol*) Extra emphasis given to a word, in order to draw attention to its meaning (*John bought a* ***red*** *car*). 171

convention (*gen*) The tacit agreement of speakers to use the same ⁺rules[1] in order to communicate. 408

conventionalism (*sem*) The view that there is a relationship of ⁺arbitrariness between words and things; also, **nominalism**; cf. naturalism. 408

convergence (*socio*) A process of linguistic change in which ⁺dialects or ⁺accents[1] become more like each other; contrasts with **divergence**. 51

conversational implicature (*prag*) An implication deduced from an utterance, using the ⁺cooperative principles that govern the efficiency of conversations (*A bus!* = 'We must run'). 117

conversational maxims (*prag*) General principles thought to underlie the efficient use of language, e.g. speakers should be relevant and clear. 117

conversation analysis (*ling*) A method of studying the structure of conversations using the techniques of ⁺ethnomethodology. 116

converseness (*sem*) A type of oppositeness of meaning, such that one word presupposes the other (*buy/sell*). 105

conversion (*gram*) A type of ⁺word formation in which an item changes its ⁺word class without the addition of an ⁺affix (*smell* = verb/noun). 90

cooperative principle (*prag*) A tacit agreement between speakers to follow the same set of ⁺conventions ('maxims') when communicating. 117

coordinate **1** (*gram*) Said of ⁺clauses displaying ⁺coordination. 95 **2** (*socio*) Said of ⁺bilinguals who are thought to have different meanings for the corresponding words in their two languages. 364

coordination (*gram*) The linking of linguistic units that have the same grammatical status, e.g. two ⁺noun phrases (***the cat*** *and* ***the dog***). 95

coordinator (*gram*) A ⁺conjunction used in ⁺coordination (*and, but*). 95

coprolalia (*clin*) Uncontrolled use of obscene language. 266

copula (*gram*) A ⁺verb whose main role is to link other elements of the ⁺clause (*It* ***is*** *ready*). 95

coreference (*sem*) The use of elements that can be interpreted only by referring to another element in a text. 119

coronal (*phonet*) Said of sounds where the ⁺blade of the tongue is raised to the hard ⁺palate. 157

corpus (*ling*) A collection of language data brought together for linguistic analysis. 415

correctness (*gen*) An absolute standard of language use deriving from the rules of institutions (e.g. language academies) or respected publications (e.g. dictionaries); cf. ⁺appropriate. 2

correlative (*gram*) Said of constructions using a pair of connecting words (*either/or*). 95

countability *see* **countable**

countable (*gram*) Said of ⁺nouns denoting separable entities, as shown by their use with such forms as *a* (*dog, chair*); **count(able)** nouns contrast with **uncountable/non-count** (⁺mass) nouns. 91

creaky (*phonet*) A ⁺voice quality produced by very slow vibration of the ⁺vocal folds. 128

creativity (*ling*) A characteristic of language that enables speakers to produce and understand an indefinitely large number of sentences. 401

creole (*socio*) A ⁺pidgin that has become the mother tongue of a speech community (through a process of **creolization**). 338

critical period (*psycho*) A period of time in child development during which language is thought to be most easily learned. 265

cross-sectional (*gen*) Said of studies that sample the language of a group of individuals at a single point in time; cf. ⁺longitudinal. 231

cryptanalysis (*gen*) The process of ⁺deciphering or ⁺decoding secret messages (**cryptograms**). 58

cryptograms *see* **cryptanalysis, cryptography**

cryptography (*gen*) The preparation of secret messages (**cryptograms**), using ⁺codes[1] and ⁺ciphers. 58

cryptology (*gen*) The study of ⁺cryptography and ⁺cryptanalysis. 58

cryptophasia *see* **idioglossia**

cued speech (*clin*) A method of ⁺speech-reading in which manual cues help to distinguish sounds. 227

cuneiform (*graph*) An ancient writing system that used wedge-shaped characters. 200

cursive (*gen*) A form of handwriting in which separate characters in a sequence have been joined. 188

cycle (*acou*) A single complete vibration, forming part of a ⁺sound wave. 133

dactyl (*poet*) A unit of rhythm in poetic ⁺metre, consisting of one heavy beat followed by two light beats. 74

dactylology (*clin*) Signing in which each letter of the alphabet is given its own sign; also, **finger spelling**. 227

dative (*gram*) An ⁺inflection[1] that typically expresses an ⁺indirect object relationship (Lat. *Dedi epistolam puellae* 'I gave the letter to the girl'). 93

daughter language *see* **parent language**

decibel (*phonet*) A unit for measuring the relative ⁺intensity of sounds, esp. in the assessment of hearing loss. 134

decipher (*gen*) To work out the meaning of a message in ⁺code[1] (esp. in ⁺cipher). 58

declarative (*gram*) A grammatical construction used in expressing a ⁺statement (*The dog barked*). 121

declension (*gram*) A set of ⁺nouns, ⁺adjectives, or ⁺pro-nouns that show the same ⁺inflections[1] (**decline**). 93

decline *see* **declension**

decode (*gen*) **1** To use the brain to interpret an incoming linguistic signal. 264 **2** To convert a secret message into intelligible language. 58

deconstruction (*styl*) An approach to literary theory that aims to show the contradiction in ⁺structuralist principles of textual analysis. 79

decreolization (*socio*) Change in a ⁺creole that makes it more like the ⁺standard language of an area. 338

deep grammar/structure (*ling*) An underlying level of grammatical organization that specifies how sentences should be interpreted; cf. ⁺surface grammar/structure. 98

defective 1 (*gram*) Applied to words that do not follow all the rules of the class to which they belong (E. ⁺auxiliary verbs, which lack the usual verb ⁺inflections[1]). 91 2 (*graph*) A writing system consisting only of ⁺consonant symbols. 204

defining *see* **restrictive**

defining vocabulary (*app*) A core set of words used to define other words. 111

definite (*gram, sem*) Said of a specific, identifiable entity or class of entities (***the*** *car*); contrasts with **indefinite** (***a*** *car*). 91

degree (*gram*) A contrast of comparison in ⁺adverbs or ⁺adjectives; usually identified as **positive** (*big*), **comparative** (*bigger*), and **superlative** (*biggest*). 92

deixis (*ling*) Features of language that refer directly to the personal, temporal, or locational characteristics of the situation (**deictic forms**) (*you, now, here*). 106

deletion (*ling*) Omitting an element of sentence structure (*that* in *I said he was ready*). 97

demonstrative (*gram*) Applied to forms whose function is to distinguish one item from other members of the same class (*this/that*). 99

denasal 1 (*phonet*) Said of sounds whose ⁺nasality has been reduced or removed. 130 2 (*clin*) Said of a ⁺voice quality with poor nasal ⁺resonance. 278

denotation (*sem*) The objective ('dictionary') relationship between a word and the reality to which it refers; also, **cognitive/referential** meaning; cf. ⁺connotation. 100

dental (*phonet*) Said of a ⁺consonant made by the ⁺apex and rims of the tongue against the teeth. 157

dependent (*gram*) Said of any element whose ⁺form[2] or ⁺function[1] is determined by another part of the sentence (in *the red car*, the ⁺article and ⁺adjective depend on the ⁺noun); also, **subordinate**. 95

derivation 1 (*gram*) A major process of ⁺word formation, esp. using ⁺affixes to produce new words (*act* → *action*); cf. ⁺inflection[1]. 90 2 (*gram*) The set of analytical steps required to ⁺generate a sentence. 97 3 (*hist*) The origins or historical development of a language or form. 294

descender (*graph*) A part of a letter that extends below the depth of the letter *x*. 192

description (*ling*) An objective and systematic account of the patterns and use of a language or ⁺variety; cf. ⁺prescription. 2

deterioration (*hist*) A change of meaning in which a word acquires a negative evaluation; also, **pejoration**; cf. ⁺amelioration. 332

determinative (*graph*) A part of a ⁺logogram that indicates its ⁺semantic content; also, **radical**; cf. ⁺phonetic. 201

determiner (*gram*) An item that co-occurs with a ⁺noun to express such meanings as number or quantity (*the, some, each*). 96

determinism *see* **linguistic relativity**

developmental (*clin*) Said of any linguistic disorder that arises out of an abnormal process of development in the child, e.g. 'developmental ⁺aphasia'; cf. ⁺acquired. 273

developmental (psycho)linguistics (*ling*) The study of the acquisition of language in children. 228

deviance (*ling*) Failure to conform to the ⁺rules[1] of the language. 88

devoiced (*phonet*) Said of a sound in which the normal amount of ⁺vocal fold vibration (⁺voice[1]) has been reduced. 165

diachronic *see* **historical, synchronic**

diacritic (*graph*) A mark added to a symbol to alter its value, e.g. an ⁺accent[3]. 156

diadochokinesis (*clin*) The ability to carry out rapid repetitive movements of the ⁺vocal organs. 273

diagnostic test (*app*) A test to show what a language learner knows and does not know. 381

diagramming *see* **parsing**

dialect (*ling*) A language ⁺variety in which the use of grammar and vocabulary identifies the regional or social background of the user; cf. ⁺accent[1]. 24

dialect continuum (*socio*) A chain of dialects whose end-points are not mutually intelligible. 25

dialectology (*ling*) The study of (esp. regional) dialects; also, **dialect geography**. 26

dialinguistics (*ling*) The study of the range of ⁺dialects and languages in a speech community. 26

dichotic listening (*psycho*) A technique for determining which half of the brain is primarily involved in processing auditory effects. 261

diction (*gen*) The effective choice of words, esp. the vocabulary used by a poet or other writer. 73

diglossia (*socio*) The use of two ⁺varieties of a language throughout a ⁺speech community, each with a distinct set of social functions. 43

digraph (*graph*) 1 A ⁺graphic unit in which two symbols have combined to function as one (*encyclopædia*). 367 2 Any sequence of two letters pronounced as a single sound (***ship***, ***wood***). 367

dimeter (*poet*) A line of verse containing two units of rhythm (⁺foot). 74

diminutive (*gram*) An ⁺affix with the general meaning of 'little' (It. *-ino*). 90

diphthong (*phonet*) A ⁺vowel in which there is a perceptible change in quality during a ⁺syllable (*time, road*); cf. ⁺monophthong, ⁺triphthong. 156

diplomatics (*graph*) The study of legal and administrative documents. 189

directive (*prag*) An utterance intended to get other people to do (or not do) something (*Sit down*); also, **command**. 121

direct method (*app*) A method of language teaching that emphasizes speaking in the ⁺target[2] language and avoids the conscious learning of ⁺grammar[1]. 378

direct object (*gram*) A ⁺clause element immediately affected by the action of the ⁺verb (*She hit* ***him***); contrasts with a less directly affected (indirect) object (*I gave* ***John*** *a letter*). 95

direct speech (*gen*) The actual utterance spoken by a person; cf. ⁺indirect (or reported) speech. 77

discontinuous (*gram*) The splitting of a grammatical construction by the insertion of another unit (***switch*** *the light* ***on***). 95

discourse (*ling*) A continuous stretch of (esp. spoken) language larger than a ⁺sentence. 116

discourse analysis (*ling*) The study of patterns of linguistic organization in ⁺discourses. 116

discovery procedure (*ling*) A set of techniques automatically applicable to a sample of language to produce a correct ⁺grammatical[2] analysis. 412

discrete (*ling*) Said of linguistic elements that have clearly defined boundaries. 401

disjunction (*sem*) An alternative or contrastive relationship between elements in a sentence (*Either we're early or the bus is late*). 107

displacement (*semiot*) The ability of language to refer to contexts removed from the speaker's immediate situation (*I was angry yesterday*). 401

dissimilation (*phonol*) The influence sound segments have on each other, so that they become less alike. 330

dissonance (*gen*) The use of sounds to convey unpleasant effects. 74

distinctive (*phonol*) Said of a feature capable of making a difference of meaning between otherwise identical forms, e.g. ⁺vocal fold vibration; cf. ⁺contrast. 162

distribution (*ling*) The total set of linguistic ⁺environments[1] in which an item can occur. 163

disyllable (*phonol*) A word of two ⁺syllables. 166

ditransitive (*gram*) Said of ⁺verbs that take two ⁺objects (*give, show*). 95

divergence *see* **convergence**

dominant language (*socio*) 1 The most important language in a ⁺multilingual speech community. 362 2 The language a ⁺bilingual knows best. 364

dorsal (*phonet*) Said of sounds made with the ⁺back ('dorsum') of the tongue ([k], [g]). 131

doublet (*gen*) A type of word game in which a series of single-letter substitutions links pairs of words. 65

downdrift (*phonol*) A gradual lowering of ⁺tones throughout an utterance in a ⁺tone language. 174

drift (*hist*) A gradual series of related changes in the historical development of a language. 330

dual (*gram*) A ⁺grammatical[1] contrast of ⁺number in some languages, referring to 'two of'. 92

dual alphabet (*graph*) The use of capital and small letters in a single system. 188

dualism (*sem*) A theory that postulates a direct, two-way relationship between linguistic forms and the entities to which they refer. 100

duality of structure (*ling*) The structural organization of language into two abstract ⁺levels[1]: meaningful units (e.g. words) and meaningless segments (sounds, letters). 401

duration (*phonet*) The length of time involved in the ⁺articulation of a sound or ⁺syllable. 171

dynamic 1 (*gram*) Type of ⁺verb that expresses activities and changes of state, allowing such forms as the ⁺progressive[1] (*He's running*); cf. ⁺stative. 93 2 (*socio*) Said of language analyses that take account of temporal change. 330

dyne (*acou*) A unit of measurement for sound pressure. 134

dysarthria (*clin*) A motor speech disorder that leaves someone unable to articulate speech sounds; in severe form, also, **anarthria**. 273

dysfluency (*clin*) The loss of ability to control the smooth flow of ⁺speech production, resulting in hesitancy, poor ⁺rhythm, ⁺stuttering, etc. 280

dysgraphia (*clin*) A ⁺language[4] disorder that primarily affects the ability to write; also, **agraphia**. 274

dyslalia (*clin*) A disorder of ⁺articulation that has no clear physical cause. 279

dyslexia (*clin*) A ⁺language[4] disorder that affects the ability to read; also, **alexia**, **word blindness**. 274

dysnomia *see* **anomia**

dysphasia *see* **aphasia**

dysphemism (*rhet*) A use of language that emphasizes unpleasantness (*a horrible dirty day*); cf. ⁺euphemism. 61

dysphonia (*clin*) The loss of ability to use the ⁺vocal folds to produce normal ⁺voice[1]; in severe form, **aphonia**. 278

dyspraxia *see* **apraxia**

dysprosody (*clin*) The loss of ability to produce speech with a normal ⁺intonation. 278

dysrhythmia (*clin*) The loss of ability to produce normal ⁺rhythm in ⁺speech production. 280

ear training (*phonet*) A technique in ⁺phonetics to train the ability to identify speech sounds. 160

echolalia (*clin*) The automatic repetition of all or part of what someone has said. 273

economy (*ling*) The use of the smallest possible number of ⁺rules[1] and symbols in carrying out a linguistic analysis. 165

educational linguistics (*ling*) The application of ⁺linguistics to language teaching and learning in schools and other educational settings. 250

egocentric speech (*psycho*) Speech that does not take account of the needs of the listener. 237

egressive (*phonet*) Said of sounds produced using an outwards-moving ⁺air-stream mechanism. 125

ejective (*phonet*) A ⁺consonant produced using the ⁺glottalic ⁺air-stream mechanism. 126

elaborated code (*socio*) A relatively formal, educated language use involving a wide range of linguistic structures; cf. ⁺restricted code. 40

elative (*gram*) An ⁺inflection[1] that typically expresses the meaning 'out of' a place. 92

electroaerometer (*phonet*) An instrument that records air flow during speech. 139

electrokymograph (*phonet*) An instrument that records the changes in the air flow from mouth and nose during speech. 139

electrolaryngograph (*phonet*) An instrument that records ⁺vocal fold vibration. 141

electromyograph (*phonet*) An instrument that records muscular contractions during speech. 139

electropalatograph (*phonet*) An instrument that makes a continuous record of the contacts between tongue and ⁺palate during speech. 140

elicit (*ling*) To obtain utterances or linguistic judgments from ⁺informants. 414

elision (*phonol, poet*) The omission of sounds in connected speech (*bacon 'n' eggs*). 166

ellipsis (*gram, rhet*) The omission of part of a sentence (e.g. for economy, emphasis), where the missing element is understood from the ⁺context[1] (*A: Where's the book? B: On the table*). 94

elocution (*gen*) The art of speech training to produce effective public speaking. 70

embedding (*gram*) Inserting one ⁺grammatical[1] unit within another (*The man **who left** was my uncle*). 95

emic (*phonol*) An approach to speech analysis that sets up a system of abstract ⁺contrastive units, esp. ⁺phonemes; cf. ⁺etic. 412

emotive meaning (*sem*) The emotional content of a use of language. 10

empty word (*gram*) A meaningless word that expresses a grammatical relationship (*It's today he goes*); also, **prop word**; cf. ⁺content word. 91

encipher (*gen*) To write a message using a ⁺cipher. 58

enclitic (*gram*) An ⁺unstressed form attached to a preceding word (*cannot*). 91

encode (*gen*) To give linguistic shape to a meaning, as part of communication. 264 **2** To convert a message from one system of signals into another (esp. for secrecy); cf. ⁺decode[2]. 58

endocentric (*gram*) Said of a construction where there is a ⁺grammatical[1] ⁺head (*the tall men*); cf. ⁺exocentric, which lacks a head (*People left early*). 95

endophoric (*gram*) Said of the relationships of ⁺cohesion that help to define the structure of a ⁺text; cf. ⁺exophoric. 119

enjamb(e)ment (*poet*) The running on of a sentence between two couplets of verse without pause. 74

environment 1 (*ling*) The parts of an ⁺utterance or ⁺text that are adjacent to an item of language. 165 **2** (*socio*) The social or cultural situation in which a particular use of language takes place. 48

epenthesis (*phonol*) The insertion of an extra (**epenthetic**) sound in the middle of a word. 330

epicene (*gram*) A ⁺noun that can refer to either sex without changing its form (*teacher*). 47

epiglottis (*anat*) A structure that closes over the ⁺larynx during swallowing. 124

epigram (*gen*) A short, witty statement, in verse or prose. 53

epigraph (*gen*) **1** An inscription on stone, buildings, coins, etc. 189 **2** A phrase or quotation above a section in a book or on the title page. 53

epigraphy (*gen*) The study of inscriptions, esp. their interpretation in ancient times. 189

epithet (*gen*) Any item that characterizes a ⁺noun and is regularly associated with it (*Ethelred **the Unready***). 105

eponym (*gen*) The name of a person after whom something, e.g. a place, a book title, is named (*Washington, Hamlet*). 112

equative (*gram*) Applied to a ⁺clause which relates two elements that are identical in meaning (*Mr Jones is a butcher*). 95

ergative (*gram*) Applied to a construction in some languages where the ⁺object of a ⁺transitive verb and the ⁺subject of an ⁺intransitive one are in the same ⁺case. 93

error **1** (*neuro*) An inaccuracy in the spontaneous use of language attributable to a malfunctioning of the neuromuscular commands from the brain. 264 **2** (*app*) A language learner's systematic use of a linguistic item that does not conform to the ⁺rules[1] of the target[2] language, because knowledge of these rules is incomplete; contrasts with unsystematic, ⁺performance faults (**mistakes**). 376

error analysis (*app*) The systematic interpretation of the unacceptable forms used by someone learning a language. 376

esophageal *see* **oesophageal**

essive (*gram*) An ⁺inflection[1] that typically expresses the meaning 'at' a place. 92

état de langue (*ling*) The 'state of a language' seen at a particular time, regardless of its antecedents or subsequent history. 411

ethnography of communication (*socio*) The study of language in relation to the social and cultural variables that influence human interaction. 48

ethnolinguistics (*ling*) The study of language in relation to ethnic groups and behaviour. 418

ethnomethodology (*socio*) The detailed study of the techniques used during linguistic interaction. 116

etic (*phonet*) The analysis of the physical patterns of speech without reference to their function within the language; cf. ⁺emic. 412

etymological fallacy (*hist*) The view that an earlier (or the oldest) meaning of a word is the correct one. 332

etymology (*hist*) The study of the origins and history of the ⁺form[1] and meaning of words. 332

etymon (*hist*) The ⁺form[1] from which a later form derives (Lat. *mater* → Fr. *mère*). 332

euphemism (*gen*) The use of a vague or indirect expression in place of one that is unpleasant or offensive (*pass away* for *die*). 61

euphony (*gen*) A pleasing sequence of sounds. 74

exclamation (*gram*) An emotional expression marked by strong ⁺intonation in speech or by an exclamation point in writing (*Good grief!*); cf. ⁺command, ⁺question, ⁺statement. 121

exclusive (*gram*) Said of a first-⁺person ⁺pronoun (*we*) that does not include the person being addressed; cf. ⁺inclusive. 92

excrescent (*ling*) Said of a sound added to a word to make the pronunciation easier. 330

exegesis (*gen*) An interpretation of a text, esp. of a biblical kind. 389

existential (*gram*) A sentence emphasizing the idea of existence (*There is a book on the table*). 95

exocentric *see* **endocentric**

exophoric (*gram*) Said of a linguistic unit that refers directly to the ⁺extralinguistic situation (*there, her*); cf. ⁺endophoric. 119

expansion 1 (*gram*) The process of adding new elements to a construction, without its basic structure being affected. 95 2 (*psycho*) An adult response to a child which adds grammatical elements that the child has omitted. 233

experimental phonetics (*phonet*) The use of instrumentation and experimental techniques to investigate the properties of speech sounds; also, **instrumental phonetics**. 138

expletive (*gen*) An exclamatory word or phrase, usually obscene or profane. 61

expression (*ling*) 1 Any string of elements treated as a unit for analysis, e.g. a ⁺sentence, ⁺idiom. 95 2 All aspects of linguistic ⁺form[1] (as opposed to meaning). 82

expressive 1 (*gen*) Said of a use of language that displays or affects a person's emotions. 10 2 (*clin*) Said of disorders of language ⁺production, e.g. 'expressive aphasia'. 267

extension 1 (*sem*) The class of entities to which a word is correctly applied, e.g. the extension of *flower* is *rose, daffodil*, etc.; cf. ⁺intension. 107 2 (*hist*) Widening the meaning of a word. 332

extralinguistic (*ling*) Said of anything (other than language) to which language can relate. 82

extraposition (*gram*) Moving an element to a position at one end of a ⁺sentence (*Working here is nice → It's nice working here*). 95

extrinsic (*anat*) Said of sets of muscles that control the gross movements of certain ⁺vocal organs, e.g. tongue, ⁺larynx. 131

eye dialect (*gen*) A way of spelling words that suggests a regional or social way of talking (*Thankee koindly, zur*). 182

eye rhyme (*poet*) A pair of words that seem to rhyme from the spelling, but have different pronunciations (*come/home*). 74

false friends (*app*) Words in different languages that resemble each other in ⁺form[1], but express dissimilar meanings (Fr. *demander* = 'request', not 'demand'); also, **faux amis**, **false cognates**. 349

false vocal folds *see* **ventricular folds**

family (*hist*) A set of languages that derive from a common ancestor (⁺parent) language, and are represented as a **family tree**. 294

feature (*ling*) Any typical, noticeable, or ⁺contrastive property of a ⁺level[1] of language. 82

feedback 1 (*prag*) The ongoing reaction speakers receive from their listeners, which helps them to evaluate the efficiency of their communication. 118 2 (*phonet*) The information speakers obtain by monitoring their own speech activity. 264

felicity conditions (*prag*) The criteria that must be satisfied if a ⁺speech act is to achieve its purpose. 121

feminine *see* **gender**

festination (*clin*) Abnormal increase of speed while speaking. 280

field *see* **semantic field**

fieldwork (*ling*) The principles and procedures of obtaining linguistic data from ⁺informants, esp. in their home environment. 414

figurative (*gen*) Said of an expressive use of language when words are used in a non-literal way to suggest illuminating comparisons and resemblances (**figures of speech**). 70

filled pause (*ling*) A vocal hesitation (*erm*). 174

filter (*acou*) A device used to separate the ⁺frequency components of a ⁺sound wave. 133

filtered speech (*phonet*) Speech passed through ⁺filters to alter its acoustic characteristics. 133

finger spelling *see* **dactylology**

finite (*gram*) A form of a ⁺verb that can occur on its own in a ⁺main clause and permits variations in ⁺tense, ⁺number, and ⁺mood (*They ran, She is running*); contrasts with **non-finite**. 93

finite-state grammar (*ling*) A simple kind of ⁺generative device that is able to process only a very limited range of sentences. 97

first language (*gen*) The language first acquired as a child (**mother tongue**, **native language**), or preferred in a ⁺multilingual situation. 372

first person *see* **person**

'fis' phenomenon (*psycho*) A child's refusal to accept an adult imitation of what it has just said. 242

fixation (*graph*) A period of relative stability between rapid eye movements. 210

flap (*phonet*) A ⁺consonant produced by a single rapid contact between two organs of articulation, e.g. the tongue tip movement [ɾ] in *very*. 159

flexion *see* **inflection**

fluency (*gen*) Smooth, rapid, effortless use of language; cf. ⁺dysfluency. 280

flyting (*poet*) An exchange of curses or personal abuse in verse form. 60

focal area (*socio*) A region where ⁺dialect forms are relatively homogeneous and tend to influence the forms used in adjoining areas. 28

focus (*gram*) An element in a sentence to which the speaker wishes to draw special attention (*It was **John** who wrote to me*). 107

folk etymology (*hist*) Altering an unfamiliar word to make it more familiar (*asparagus → sparrow-grass*); also, **popular etymology**. 332

foot (*phonol, poet*) A basic unit of ⁺rhythm, esp. used in describing poetic ⁺metre. 74

foregrounding (*poet*) Any departure from a linguistic or socially accepted norm, esp. in literary language. 71

foreign language (*app*) A non-native language, esp. one that has no official status in a country; cf. ⁺second language. 344

forensic linguistics (*ling*) The use of linguistic techniques to investigate crimes in which language data constitute part of the evidence. 69

form 1 (*ling*) The outward appearance or structure of language, as opposed to its function, meaning, or social use (**formal** vs ⁺notional[1]). 82 2 (*gram*) The variations in which a linguistic unit can appear (the 'forms' *walk, walks*, etc.). 91

formal 1 *see* **form**[1] 2 *see* **formality**

formalist (*styl*) Said of an approach that studies the structural (⁺formal[1]) basis for literary effects in great detail. 78

formality (*socio*) A scale of language use, relating to situations that are socially careful or correct (**formal**) or otherwise (**informal**). 40

formal universal (*ling*) An obligatory feature of ⁺grammar[2] construction; cf. ⁺substantive universal. 85

formant (*acou*) A concentration of acoustic energy, esp. distinctive in ⁺vowels[1] and ⁺voiced sounds. 135

formative (*gram*) An irreducible grammatical element that enters into the construction of larger linguistic units. 90

form class (*gram*) A set of items that display similar or identical grammatical features. 91

formulaic (*ling*) Said of a sentence that does not permit the usual range of grammatical variation (*Many happy returns*); also, **fossilized** or **stereotyped** sentences, or **routines**. 52

form word *see* **function word**

fortis (*phonet*) Said of ⁺consonants made with relatively strong muscular effort and breath force ([f], [p]); cf. ⁺lenis. 159

fossilized (*ling*) Said of any construction that lacks ⁺productivity, e.g. ⁺idioms (*spick and span*), ⁺formulaic utterances (*So be it!*). 52

frame (*gram*) A specific ⁺structural[2] ⁺context within which a class of items can be used. 95

free form (*gram*) A minimal grammatical unit that can be used as a ⁺word without additional elements; also known as a free ⁺morpheme; cf. ⁺bound form. 90

free translation (*gen*) A ⁺translation expressing the meaning rather than the ⁺form[1] of the ⁺source language; contrasts with **literal** (word-for-word) translation. 346

free variation (*phonol*) The substitution of one sound for another without causing any change of meaning. 163

frequency (*acou*) The number of ⁺sound waves per second produced by a source of vibration. 133

fricative (*phonet*) Said of a ⁺consonant made when two ⁺vocal organs come so close together that the air moving between them produces audible friction ([f],[z]); also, **spirant**. 159

frictionless continuant *see* **approximant**

front (*phonet*) Said of sounds made in the front part of the mouth or by the front part (⁺blade) of the tongue ([i], [t]); cf. ⁺back, ⁺centre. 131

fronting 1 (*phonol*) ⁺Articulation of a sound further forward in the mouth than is normal. 157 2 (*gram*) Moving a ⁺constituent from the middle or end of a ⁺sentence to the front (***Smith** his name was*). 95

full verb *see* **lexical verb**

full word *see* **content word**

function (*ling*) 1 The relationship between a linguistic form and the other elements of the system in which it is used, e.g. a ⁺noun as ⁺subject or ⁺object of a ⁺clause. 95 2 The role language plays in communication (e.g. to express ideas, attitudes) or in particular social situations (e.g. religious, legal). 10

functional 1 (*ling*) Said of linguistic approaches that treat the notion of ⁺function as central, esp.

linguistic ⁺form[1] does not directly reflect its communicative purpose (using *It's cold in here* to mean 'Close the window'). 121

inessive (*gram*) An ⁺inflection[1] that typically expresses location or position within a place. 92

infinitive (*gram*) A ⁺non-finite form of the ⁺verb, which in many languages acts as the basic form (E. *run*, Fr. *donner* 'to give'). 93

infix (*gram*) An ⁺affix added within a ⁺root[1]. 90

inflecting/inflected/inflectional language (*ling*) A language in which ⁺words express grammatical relationships by using ⁺inflections[1]; also, **synthetic/fusional** language. 295

inflection/inflexion 1 (*gram*) An ⁺affix that signals a grammatical relationship, e.g. ⁺case, ⁺tense[1] (*girl's, walked*). 90 **2** (*phonet*) Change in voice ⁺pitch during speech. 171

informal *see* **formality**

informant (*ling*) Someone who acts as a source of data for linguistic analysis. 414

information (*ling*) The way a message content is structured, e.g. into **given** and **new**. 120

ingressive (*phonet*) Said of all sounds produced with an inwards-moving air stream. 125

inhalation *see* **inspiration**

initiator (*phonet*) The ⁺vocal organs that are the source of air movement, e.g. lungs. 124

innateness hypothesis (*psycho*) The view that a child is born with a biological predisposition to learn language, involving a knowledge of its ⁺universal structural principles; also, **nativism**. 236

inner ear (*anat*) The part of the ear containing the ⁺cochlea. 143

inspiration (*phys*) The act of drawing air into the lungs; also, **inhalation**. 124

institutional linguistics (*ling*) The study of the problems involved in ⁺language planning. 366

instructive (*gram*) An ⁺inflection[1] that typically expresses the meaning 'by'. 92

instrumental (*gram*) An ⁺inflection[1] that typically expresses the meaning 'by means of'. 92

instrumental phonetics *see* **experimental phonetics**

intensifier (*gram*) A word or phrase that adds force or emphasis (***very** good, **awfully** pretty*). 91

intension (*sem*) The set of defining properties that determines how a term is to be used (*table* → 'legs', 'flat surface', etc.). 107

intensity (*acou*) The power transmitted along a ⁺sound wave. 134

interchangeability (*semiot*) The ability of a signalling system to be mutually transmitted and received by members of the same species. 400

interdental (*phonet*) A ⁺consonant made by the ⁺apex of the tongue between the teeth ([θ], [ð]). 157

interference *see* **transfer**

interjection (*gram*) A class of ⁺words with ⁺emotive meaning, which do not form grammatical relationships with other classes (*Gosh!, Yuk!*). 91

interlanguage (*app*) The language system used at an intermediate stage of foreign language learning. 374

intermediate vowel (*phonet*) A ⁺vowel[1] that falls between two adjacent ⁺cardinal vowels. 156

internal evidence (*hist*) Linguistic features in a text that indicate when the work was written. 189

internal rhyme (*poet*) The rhyming of words within lines of verse. 74

interpersonal function (*ling*) The use of language to establish and maintain social relations; cf. ⁺ideational function, ⁺textual function. 10

interpret (*gen*) To make an oral ⁺translation[1]. 351

interrogative (*gram*) A type of ⁺sentence or ⁺verb form used in the expression of ⁺questions (*Who is he?, Are they there?*); cf. ⁺declarative. 95

interrogative word (*gram*) A word used at the beginning of a ⁺clause to mark it as a ⁺question (***Who** is here?*). 95

intervocalic (*phonet*) A ⁺consonant used between two ⁺vowels[1] (/p/ in *apart*). 166

intonation (*phonol*) The ⁺contrastive use of ⁺pitch in speech. 171

intonation contour *see* **tone unit**

intransitive (*gram*) Said of a ⁺verb or ⁺sentence that cannot take a ⁺direct object (*She's going*); cf. ⁺transitive. 95

intraoral pressure (*phonet*) The build-up of air inside the mouth needed to produce certain speech sounds, e.g. ⁺plosives. 124

intrinsic (*anat*) Said of sets of muscles that control the fine movements of certain ⁺vocal organs, e.g. tongue, ⁺larynx. 131

intrusion (*phonet*) The use of sounds in connected speech that do not appear when the words or ⁺syllables are heard in isolation, e.g. 'intrusive *r*' between ⁺vowels[1] (as in *law(r) and order*). 166

intuition (*gen*) A person's instinctive knowledge of language, which decides whether ⁺sentences are acceptable and how they can be interrelated. 414

invariable word (*gram*) A word that does not undergo any change in structure (*under, but*); cf. ⁺variable word. 91

inversion (*gram*) A reversed sequence of elements (*He is going → Is he going?*). 245

irony (*gen*) Language that expresses a meaning other than that literally conveyed by the words (*That's marvellous*, said of poor work). 70

irregular (*gen*) Said of a linguistic ⁺form[1] that is an exception to a pattern stated in a rule[1]. 408

isochrony/isochronism (*phonet*) A rhythmic pattern in which ⁺stressed ⁺syllables fall at roughly regular intervals throughout an utterance. 171

isogloss (*socio*) A line on a map showing the boundary of an area in which a linguistic feature is used; the lines mark such features as vocabulary (**isolex**), ⁺morphology (**isomorph**), ⁺phonology (**isophone**), ⁺semantics (**isoseme**), or socio-cultural use (**isopleth**). 28

isolating language (*ling*) A language in which ⁺words are ⁺invariable and grammatical relations are shown mainly by ⁺word order, e.g. Chinese; also, **analytic/root** language. 295

iterative (*gram*) A ⁺form[1] that expresses the repetition of an action (*frequently*), esp. as part of the ⁺aspect system. 93

jargon 1 (*gen*) The technical language of a special field. 56 **2** (*gen*) The obscure use of specialized language. 383 **3** (*psycho*) Unintelligible utterance with good ⁺intonation, used by young children when learning to talk. 239 **4** (*clin*) Unintelligible speech in some ⁺language disorders. 273

juncture (*phonol*) ⁺Phonetic boundary features that demarcate units of grammar, e.g. certain features of ⁺pitch, ⁺duration, pause. 166

kana (*graph*) Either of the two Japanese ⁺syllabic[2] writing systems, **hiragana** and **katakana**. 197

katakana *see* **kana**

kernel (*gram*) A basic type of ⁺sentence structure, as used in early ⁺generative grammar. 97

kin(a)esthesis (*phys*) Awareness of the movements and positions of the ⁺vocal organs during speech; also, **kin(a)esthetic feedback**. 124

kineme (*semiot*) The smallest ⁺contrastive unit of body expression. 406

kinesics (*semiot*) The systematic use of facial expression and bodily gestures/movements to communicate meaning. 403

koine (*socio*) The spoken language of a locality that has become a ⁺standard language. 43

kymograph (*phonet*) An early device for recording information about ⁺vocal organ movements. 138

labial (*phonet*) The active use of one or both lips in the ⁺articulation of a sound ([f], [u]). 157

labialization (*phonet*) ⁺Rounding the lips while making a speech sound. 158

labio-dental (*phonet*) Said of a ⁺consonant in which one lip actively contacts the teeth ([f], [v]). 157

labio-velar (*phonet*) A speech sound made at the ⁺velum with simultaneous lip ⁺rounding ([w], [u]). 157

laminal (*phonet*) Said of a ⁺consonant made with the ⁺blade (or lamina) of the tongue in contact with the upper lip, teeth, or ⁺alveolar ridge ([s], [t̻]). 159

langage (*ling*) The human faculty of speech. 411

language (*gen*) **1** The systematic, conventional use of sounds, signs, or written symbols in a human society for communication and self-expression. 400 **2** A specially devised system of symbols for programming and interacting with computers. 400 **3** The means animals use to communicate. 401 **4** (*clin*) The symbolic aspects of language[1], excluding ⁺phonetics (and often ⁺phonology). 267

language acquisition 1 (*psycho*) The process of learning a ⁺first language in children. 228 **2** (*app*) The analogous process of gaining a ⁺foreign or ⁺second language. 370

language acquisition device (*psycho*) The innate capacity that enables children to learn their mother tongue; often, **LAD**. 236

language attitudes (*socio*) The feelings people have about their own language or the language(s) of others. 1

heterographs (*gen*) Words that have the same meaning or pronunciation, but differ in spelling (*bear* vs *bare*). 106

heteronyms (*gen*) Words that differ in meaning but are identical in either pronunciation or spelling (*threw* vs *through*). 106

heterophemy (*gen*) An unintentional error in spoken or written language. 166

heterotopy (*gen*) A misplaced sound during (esp. fast) speech. 280

hexameter (*poet*) A line of verse containing six units of rhythm (⁺foot). 74

hiatus **1** (*phonet*) The use of adjacent ⁺vowels[1] in different ⁺syllables. 166 **2** (*gen*) A break in a ⁺sentence that leaves it incomplete. 52

hierarchy (*ling*) A classification of linguistic units into a series of successively subordinate ⁺levels[3], esp. an analysis of ⁺sentences into ⁺clauses, ⁺phrases, ⁺words, and ⁺morphemes. 82, 95

hieroglyphic (*graph*) A writing system using mainly pictorial symbols; esp. applied to Egyptian. 201

high (*phonet, phonol*) **1** Said of ⁺vowels[1] (and sometimes ⁺consonants) made by raising the tongue towards the roof of the mouth ([i], [k]). 153 **2** Said of ⁺tones[1] that use a relatively high level of ⁺pitch range. 174 **3** (*socio*) Said of the more prestigious ⁺variety in ⁺diglossia. 43

hiragana *see* **kana**

historical linguistics (*ling*) The study of development of language and languages over time; also, **diachronic linguistics** or (with different emphasis) **comparative philology**. 411

hold (*phonet*) To maintain a single position of the ⁺vocal organs for a period of time. 159

holograph (*gen*) A document that is entirely written in the handwriting of its author. 189

holophrase (*psycho*) A grammatically unstructured utterance, usually consisting of a single word, typical of the earliest stage of language learning in children (*dada, allgone*). 244

homographs (*gen*) Words with the same spelling but different meanings (*wind* = 'air' vs *wind* = 'turn'). 106

homonyms (*gen*) Words with the same ⁺form[1] but different meanings (*ear* = 'corn' vs *ear* = 'body part'). 106

homophones (*gen*) Words with the same pronunciation but different meanings (*rode/rowed*). 106

homorganic (*phonet*) Said of sounds made at the same place of ⁺articulation ([p], [b], [m]). 159

honorific (*socio*) A use of language (esp. of grammar[1]) to express levels of politeness or respect. 99

hybrid (*gram*) A ⁺word composed of elements from different languages (*television*, from Greek and Latin). 90

hydronymy (*gen*) The study of the names of rivers, lakes, etc. 114

hyperacusis (*clin*) An extremely acute ability to hear and distinguish sounds. 142

hyperbole (*gen, rhet*) Emphatic exaggeration (*There were millions of people in the cinema*). 70

hypercorrection (*socio*) A linguistic ⁺form[1] that goes beyond the norm of a ⁺target[2] ⁺variety, because of the speaker's desire to be correct; also, **hyper-urbanism**, **overcorrection**. 2

hypercreolization (*socio*) The development of a kind of ⁺creole that is a reaction away from the ⁺standard language. 338

hypernasality (*clin*) Excessive ⁺nasal ⁺resonance in speech. 278

hyper-urbanism *see* **hypercorrection**

hyp(o)acusis (*clin*) An impairment of ability to hear and distinguish sounds. 268

hypocoristic (*gen*) A pet name (*Bill, honey*). 112

hyponasality (*clin*) Lack of normal ⁺nasal ⁺resonance in speech. 278

hyponymy (*sem*) The relationship between specific and general words, where the former is included in the latter (*cat* is a **hyponym** of *animal*). 105

hypostatize (*gen*) To speak of an abstract quality as if it were human. 70

hypotaxis (*gram*) The linking of a ⁺dependent (**hypotactic**) clause to another part of the sentence using ⁺conjunctions (*The boy left when the bell rang*); cf. ⁺parataxis. 95

hysterical (*clin*) Said of disorders of ⁺voice[1] or hearing that are psychological in origin. 278

iamb (*poet*) A unit of ⁺metre consisting of an unstressed ⁺syllable followed by a ⁺stressed syllable (*'To be/ or not/ to be/'*). 74

iconic (*sem*) Said of signals whose physical form corresponds to features of the entities to which they refer (as in ⁺onomatopoeia, e.g. *cuckoo*). 222

ictus (*poet*) The ⁺stressed ⁺syllable in a unit of ⁺metre. 74

idealization (*ling*) The ignoring of certain kinds of variability in linguistic data, in order to reach general conclusions. 413

ideation (*psycho*) The cognitive process of forming ideas and relationships of meaning, prior to their formulation in language. 264

ideational function (*ling*) The use of language to refer to the people, events, etc. in the world; cf. ⁺interpersonal, ⁺textual functions. 10

ideogram (*graph*) A symbol used in a writing system to represent a whole word or concept; also, **ideograph**. 200

ideograph *see* **ideogram**

idioglossia (*gen*) An invented form of speech whose meaning is known only to the inventor, e.g. the language sometimes used by twins; also, **autonomous speech**, **cryptophasia**. 249

idiolect (*ling*) The linguistic system of an individual speaker. 24

idiom (*sem*) A sequence of words that is a unit of meaning (*kick the bucket* = 'die'). 105

illative (*gram*) An ⁺inflection[1] that typically expresses the meaning of 'into' a place. 92

ill formed (*gram*) Said of any ⁺ungrammatical[1] sentence; cf. ⁺well formed. 88

illocutionary act (*prag*) A ⁺speech act involving a ⁺performative verb (baptize, promise, request, etc.); cf. ⁺locutionary/⁺perlocutionary act. 121

imagery **1** (*gen*) The use of ⁺metaphor, ⁺simile, and other ⁺figurative language, esp. in a literary context. 70 **2** (*psycho*) Language that produces clear or vivid mental pictures. 103

imitation (*psycho*) The copying of linguistic behaviour, esp. while learning a language; cf. ⁺comprehension, ⁺production. 236

immediate constituent *see* **constituent**

immersion (*app*) Said of a ⁺bilingual programme in which ⁺monolingual children attend a school where another language is the medium of instruction. 369

imperative (*gram*) A grammatical ⁺mood expressing a ⁺command (*Look!*). 93

imperfect (*gram*) A ⁺tense[1] form expressing such meanings as past duration and continuity (Lat. *amabam* 'I was loving/used to love'). 93

imperfective *see* **perfective**

impersonal (*gram*) Said of constructions or ⁺verbs with an unspecified ⁺agent (*It's raining*). 95

implicational universal (*ling*) A type of ⁺universal statement that takes the form 'If a language has X, then it also has Y.' 85

implicature *see* **conversational implicature**

implosive (*phonet*) A ⁺consonant made using the ⁺glottalic ⁺air-stream mechanism with inwards-flowing air ([ɓ], [ɗ]). 126

inalienable *see* **alienable**

inanimate *see* **animate**

incapsulating language *see* **polysynthetic language**

inceptive (*gram*) Said of a ⁺verb form that specifies the beginning of an action ('be about to'), e.g. Lat. *-escere*; also, **inchoative**. 92

inchoative *see* **inceptive**

inclusive (*gram*) Said of a first-⁺person ⁺pronoun that refers to both the speaker and someone else, as when *we* means 'me and you'; cf. ⁺exclusive. 92

incompatibility (*sem*) A feature of mutually-defining items where the choice of one excludes the use of the others (*The ink is red/blue*). 105

incorporating language *see* **polysynthetic language**

indefinite *see* **definite**

indefinite vowel *see* **schwa**

independent clause *see* **main clause**

indexical **1** (*ling*) Said of features of speech or writing (esp. ⁺voice quality) that reveal the personal characteristics of the user, e.g. age, sex. 173 **2** (*sem*) *see* **deixis**

indicative (*gram*) A grammatical ⁺mood that expresses objective statements (*My car is new*). 93

indirect object *see* **direct object**

indirect question (*gram*) A ⁺question as expressed in ⁺indirect speech (*He asked if she was in*). 77

indirect speech (*gram*) A construction in which the speaker's words are made ⁺subordinate to a ⁺verb of 'saying' (*She replied that she had*); also, **reported speech**; cf. ⁺direct speech. 77

indirect speech act (*prag*) An utterance whose

linguistic ⁺form[1] does not directly reflect its communicative purpose (using *It's cold in here* to mean 'Close the window'). 121

inessive (*gram*) An ⁺inflection[1] that typically expresses location or position within a place. 92

infinitive (*gram*) A ⁺non-finite form of the ⁺verb, which in many languages acts as the basic form (E. *run*, Fr. *donner* 'to give'). 93

infix (*gram*) An ⁺affix added within a ⁺root[1]. 90

inflecting/inflected/inflectional language (*ling*) A language in which ⁺words express grammatical relationships by using ⁺inflections[1]; also, **synthetic/fusional** language. 295

inflection/inflexion 1 (*gram*) An ⁺affix that signals a grammatical relationship, e.g. ⁺case, ⁺tense[1] (*girl's*, *walked*). 90 2 (*phonet*) Change in voice ⁺pitch during speech. 171

informal *see* **formality**

informant (*ling*) Someone who acts as a source of data for linguistic analysis. 414

information (*ling*) The way a message content is structured, e.g. into **given** and **new**. 120

ingressive (*phonet*) Said of all sounds produced with an inwards-moving air stream. 125

inhalation *see* **inspiration**

initiator (*phonet*) The ⁺vocal organs that are the source of air movement, e.g. lungs. 124

innateness hypothesis (*psycho*) The view that a child is born with a biological predisposition to learn language, involving a knowledge of its ⁺universal structural principles; also, **nativism**. 236

inner ear (*anat*) The part of the ear containing the ⁺cochlea. 143

inspiration (*phys*) The act of drawing air into the lungs; also, **inhalation**. 124

institutional linguistics (*ling*) The study of the problems involved in ⁺language planning. 366

instructive (*gram*) An ⁺inflection[1] that typically expresses the meaning 'by'. 92

instrumental (*gram*) An ⁺inflection[1] that typically expresses the meaning 'by means of'. 92

instrumental phonetics *see* **experimental phonetics**

intensifier (*gram*) A word or phrase that adds force or emphasis (***very** good*, ***awfully** pretty*). 91

intension (*sem*) The set of defining properties that determines how a term is to be used (*table* → 'legs', 'flat surface', etc.). 107

intensity (*acou*) The power transmitted along a ⁺sound wave. 134

interchangeability (*semiot*) The ability of a signalling system to be mutually transmitted and received by members of the same species. 400

interdental (*phonet*) A ⁺consonant made by the ⁺apex of the tongue between the teeth ([θ], [ð]). 157

interference *see* **transfer**

interjection (*gram*) A class of ⁺words with ⁺emotive meaning, which do not form grammatical relationships with other classes (*Gosh!*, *Yuk!*). 91

interlanguage (*app*) The language system used at an intermediate stage of foreign language learning. 374

intermediate vowel (*phonet*) A ⁺vowel[1] that falls between two adjacent ⁺cardinal vowels. 156

internal evidence (*hist*) Linguistic features in a text that indicate when the work was written. 189

internal rhyme (*poet*) The rhyming of words within lines of verse. 74

interpersonal function (*ling*) The use of language to establish and maintain social relations; cf. ⁺ideational function, ⁺textual function. 10

interpret (*gen*) To make an oral ⁺translation[1]. 351

interrogative (*gram*) A type of ⁺sentence or ⁺verb form used in the expression of ⁺questions (*Who is he?*, *Are they there?*); cf. ⁺declarative. 95

interrogative word (*gram*) A word used at the beginning of a ⁺clause to mark it as a ⁺question (***Who** is here?*). 95

intervocalic (*phonet*) A ⁺consonant used between two ⁺vowels[1] (/p/ in *apart*). 166

intonation (*phonol*) The ⁺contrastive use of ⁺pitch in speech. 171

intonation contour *see* **tone unit**

intransitive (*gram*) Said of a ⁺verb or ⁺sentence that cannot take a ⁺direct object (*She's going*); cf. ⁺transitive. 95

intraoral pressure (*phonet*) The build-up of air inside the mouth needed to produce certain speech sounds, e.g. ⁺plosives. 124

intrinsic (*anat*) Said of sets of muscles that control the fine movements of certain ⁺vocal organs, e.g. tongue, ⁺larynx. 131

intrusion (*phonet*) The use of sounds in connected speech that do not appear when the words or ⁺syllables are heard in isolation, e.g. 'intrusive *r*' between ⁺vowels[1] (as in *law(r) and order*). 166

intuition (*gen*) A person's instinctive knowledge of language, which decides whether ⁺sentences are acceptable and how they can be interrelated. 414

invariable word (*gram*) A word that does not undergo any change in structure (*under*, *but*); cf. ⁺variable word. 91

inversion (*gram*) A reversed sequence of elements (*He is going* → *Is he going?*). 245

irony (*gen*) Language that expresses a meaning other than that literally conveyed by the words (*That's marvellous*, said of poor work). 70

irregular (*gen*) Said of a linguistic ⁺form[1] that is an exception to a pattern stated in a rule[1]. 408

isochrony/isochronism (*phonet*) A rhythmic pattern in which ⁺stressed ⁺syllables fall at roughly regular intervals throughout an utterance. 171

isogloss (*socio*) A line on a map showing the boundary of an area in which a linguistic feature is used; the lines mark such features as vocabulary (**isolex**), ⁺morphology (**isomorph**), ⁺phonology (**isophone**), ⁺semantics (**isoseme**), or socio-cultural use (**isopleth**). 28

isolating language (*ling*) A language in which ⁺words are ⁺invariable and grammatical relations are shown mainly by ⁺word order, e.g. Chinese; also, **analytic/root** language. 295

iterative (*gram*) A ⁺form[1] that expresses the repetition of an action (*frequently*), esp. as part of the ⁺aspect system. 93

jargon 1 (*gen*) The technical language of a special field. 56 2 (*gen*) The obscure use of specialized language. 383 3 (*psycho*) Unintelligible utterance with good ⁺intonation, used by young children when learning to talk. 239 4 (*clin*) Unintelligible speech in some ⁺language disorders. 273

juncture (*phonol*) ⁺Phonetic boundary features that demarcate units of grammar, e.g. certain features of ⁺pitch, ⁺duration, pause. 166

kana (*graph*) Either of the two Japanese ⁺syllabic[2] writing systems, **hiragana** and **katakana**. 197

katakana *see* **kana**

kernel (*gram*) A basic type of ⁺sentence structure, as used in early ⁺generative grammar. 97

kin(a)esthesis (*phys*) Awareness of the movements and positions of the ⁺vocal organs during speech; also, **kin(a)esthetic feedback**. 124

kineme (*semiot*) The smallest ⁺contrastive unit of body expression. 406

kinesics (*semiot*) The systematic use of facial expression and bodily gestures/movements to communicate meaning. 403

koine (*socio*) The spoken language of a locality that has become a ⁺standard language. 43

kymograph (*phonet*) An early device for recording information about ⁺vocal organ movements. 138

labial (*phonet*) The active use of one or both lips in the ⁺articulation of a sound ([f], [u]). 157

labialization (*phonet*) ⁺Rounding the lips while making a speech sound. 158

labio-dental (*phonet*) Said of a ⁺consonant in which one lip actively contacts the teeth ([f], [v]). 157

labio-velar (*phonet*) A speech sound made at the ⁺velum with simultaneous lip ⁺rounding ([w], [u]). 157

laminal (*phonet*) Said of a ⁺consonant made with the ⁺blade (or lamina) of the tongue in contact with the upper lip, teeth, or ⁺alveolar ridge ([s], [t̻]). 159

langage (*ling*) The human faculty of speech. 411

language (*gen*) 1 The systematic, conventional use of sounds, signs, or written symbols in a human society for communication and self-expression. 400 2 A specially devised system of symbols for programming and interacting with computers. 400 3 The means animals use to communicate. 401 4 (*clin*) The symbolic aspects of language[1], excluding ⁺phonetics (and often ⁺phonology). 267

language acquisition 1 (*psycho*) The process of learning a ⁺first language in children. 228 2 (*app*) The analogous process of gaining a ⁺foreign or ⁺second language. 370

language acquisition device (*psycho*) The innate capacity that enables children to learn their mother tongue; often, **LAD**. 236

language attitudes (*socio*) The feelings people have about their own language or the language(s) of others. 1

language centre/center (*neuro*) A brain area controlling ⁺production or ⁺comprehension. 262

language change (*hist*) Change within a language over a period of time; cf. ⁺language shift. 330

language contact (*socio*) A situation of prolonged association between the speakers of different languages. 362

language disorder (*clin*) A serious abnormality in the system underlying the use of language. 266

language laboratory (*app*) A classroom that uses tape-recorder booths to enable students to listen and respond to foreign utterances. 381

language loss **1** (*socio*) The gradual loss of ability to use a language, e.g. in immigrant situations. 362 **2** (*clin*) The sudden loss of language as a result of brain damage. 272

language loyalty (*socio*) The personal attachment to a language that leads to its continued use in a country where other languages are ⁺dominant. 369

language maintenance (*socio*) The continued use of and support for a language in a ⁺bilingual or ⁺multilingual community. 362

language pathologist/pathology *see* **speech pathologist/pathology**

language planning (*socio*) Official intentions and policies affecting language use in a country. 366

language shift (*socio*) A permanent change in a person's choice of language for everyday purposes (esp. as a result of immigrant movement). 362

language therapist *see* **speech pathologist**

langue (*ling*) The language system shared by a ⁺speech community; cf. ⁺*parole*. 411

laryngeal (*phonet*) A speech sound made in the ⁺larynx. 128

laryngectomee (*clin*) Someone who has had a ⁺laryngectomy. 278

laryngectomy (*clin*) The surgical removal of some or all of the ⁺larynx. 278

laryngology (*clin*) The study of the anatomy, physiology, and diseases of the ⁺larynx. 128

laryngopharynx (*anat*) The lower part of the ⁺pharynx, between ⁺larynx and ⁺oropharynx. 130

laryngoscope (*clin*) A device inserted into the mouth to enable the ⁺larynx to be seen. 129

larynx (*anat*) The part of the ⁺trachea containing the ⁺vocal folds. 128

lateral (*phonet*) Said of a ⁺consonant in which air escapes around one or both sides of a ⁺closure made in the mouth, as in the various kinds of *l* sound. 159

lateralization/laterality (*neuro*) The primary involvement of one hemisphere of the brain in the exercise of a bodily function, e.g. language. 260

latinate (*gram*) Applied to any ⁺grammar[1] that is based on the terms and categories used in classical Latin grammar. 2

law (*hist*) A statement of the predictable relationships (esp. in the use of sounds) between different languages or states of a language. 330

lax *see* **tension**

leading (*graph*) The spacing between lines of type. 192

lect (*socio*) A collection of linguistic phenomena that has a functional identity within a speech community, e.g. a regional or social ⁺variety. 24

length *see* **duration**

lenis (*phonet*) Said of ⁺consonants made with a relatively weak degree of muscular effort and breath force ([b], [v]); cf. ⁺fortis. 159

lenition (*phonet*) A relaxation of muscular effort during ⁺articulation. 159

lento (*phonet*) Said of speech produced slowly or with careful ⁺articulation. 171

lesion (*clin*) An abnormal change in body tissue due to injury or disease. 261

level **1** (*ling*) A major dimension of the structural organization of language, capable of independent study, e.g. ⁺phonology, ⁺syntax, 82 **2** (*gram*) A kind of representation recognized within the ⁺derivation[2] of a sentence, e.g. ⁺deep vs ⁺surface grammar. 413 **3** (*gram*) One of a series of structural layers within a ⁺sentence (⁺clause, ⁺phrase, ⁺word, etc.); also, **rank**. 95 **4** (*phonol*) A degree of ⁺pitch height or ⁺loudness during speech. 172 **5** (*socio*) A mode of expression felt to suit a type of social situation (formal, intimate, etc.). 40

lexeme (*sem*) The smallest ⁺contrastive unit in a ⁺semantic system (*run, cat, switch on*); also, **lexical item**. 104

lexical diffusion (*socio*) The gradual spread of a linguistic change through a language. 334

lexical item *see* **lexeme**

lexical field *see* **semantic field**

lexical verb (*gram*) A ⁺verb expressing an action, event, or state; also, **full-/main** verb; cf. ⁺auxiliary verb. 91

lexical word *see* **content word**

lexicography (*gen*) The art and science of dictionary-making (by **lexicographers**). 108

lexicology (*sem*) The study of the history and present state of a language's vocabulary. 108

lexicon (*sem*) The vocabulary of a language, esp. in dictionary form; also, **lexis**. 108

lexicostatistics (*hist*) A method for comparing the rates of change in sets of words in hypothetically related languages; cf. ⁺glottochronology. 333

lexis *see* **lexicon**

liaison (*phonol*) The pronunciation of a ⁺consonant at the end of a word when the next word begins with a ⁺vowel (Fr. *C'est un* ... 'It is a ...'); cf. ⁺linking. 166

ligature (*graph*) A character in which two or more letters have been joined together (*æ, œ*). 196

linear (*graph*) Said of ⁺scripts using simply drawn characters instead of pictorial writing. 185

lingua franca (*gen*) A medium of communication for people who speak different ⁺first languages. 359

lingual/linguo- (*phonet*) Said of any sound made with the tongue. 131

linguist **1** (*gen*) Someone who is proficient in several languages. 418 **2** (*ling*) A student or practitioner of the subject of ⁺linguistics; also, **linguistician**. 418

linguistic **1** (*gen*) Pertaining to ⁺language[1]. **2** (*ling*) Pertaining to ⁺linguistics. 418

linguistic atlas (*ling*) A set of maps showing the geographical distribution of linguistic items; also, **dialect atlas**. 30

linguistic change *see* **language change**

linguistic geography *see* **geographical linguistics**

linguistician *see* **linguist**[2]

linguistic relativity/determinism (*ling*) The hypothesis that a language's structure governs the way in which its speakers view the world. 15

linguistics (*ling*) The science of language. 408

linguistic science(s) *see* **linguistics**

linking (*phonol*) A sound introduced between two ⁺syllables or ⁺words, for ease of pronunciation (E. 'linking /r/' of *car and* ...); cf. ⁺liaison. 166

linking verb *see* **copula**

lipogram (*gen*) A text from which a specific letter has been omitted throughout. 65

lip reading *see* **speech reading**

liquid (*phonet*) [l]- or [r]-type ⁺consonants. 168

lisp (*clin*) An abnormal ⁺articulation of a ⁺sibilant ⁺consonant, esp. [s]. 279

literal (*gen*) The usual meaning of a word or phrase; cf. ⁺figurative. 70

literal translation *see* **free translation**

loan translation *see* **calque**

loan word *see* **borrow**

localization (*neuro*) The control of a specific kind of behaviour, e.g. speech, by a specific area of the brain. 262

locative (*gram*) A form that expresses location (*at the corner*). 93

locutionary act (*prag*) The ⁺speech act of making a meaningful utterance; cf. ⁺illocutionary act. 121

logocentrism (*styl*) A language- or word-centred view of literature or other behaviour. 79

logogram (*graph*) A symbol that represents a ⁺word (as in Chinese); also, **logograph**. 202

logograph *see* **logogram**

logogriph (*gen*) A word puzzle using ⁺anagrams. 65

logop(a)edics *see* **speech pathology**

logop(a)edist *see* **speech pathologist**

logorrhoea (*gen, clin*) Excessive, uncontrolled, incoherent speech. 273

long (*phonol*) Said of a ⁺phoneme that ⁺contrasts because of its greater ⁺duration (the ⁺vowel[1] of *beat* compared with *bit*). 153

longitudinal (*gen*) Said of studies that follow ⁺language acquisition over a period of time; cf. ⁺cross-sectional. 231

look-and-say (*app*) A method of teaching reading that focuses on the recognition of whole words; also, **whole word**; cf. ⁺phonics. 253

loudness (*phonet*) The auditory sensation that primarily relates to a sound's intensity; also, **volume**. 44

low (*phonet, phonol*) **1** Said of ⁺vowels[1] made with the tongue in the bottom area of the mouth ([a], [ɑ]). 153 **2** Said of ⁺tones[1] that use a

relatively low level of †pitch range. 174 **3** (*socio*) Said of the less prestigious †variety in †diglossia. 43

machine translation (*gen*) The use of a computer to carry out the task of †translation; also, **automatic translation**. 352

macrolinguistics (*ling*) A broad conception of linguistic enquiry, including psychological, cultural, etc. factors. 408

main clause (*gram*) A †clause that does not depend on any other part of a †sentence (***The man arrived after the bus left***); also, **independent clause**; cf. †subordination. 95

maintenance *see* **language maintenance**

main verb *see* **lexical verb**

majuscule (*graph*) A form of writing consisting of capital letters; cf. †minuscule. 188

malapropism (*gen*) An inappropriate word, used because of its similarity in sound to the intended word (*a paradigm of virtue*). 77

malformation (*gen*) An unacceptable †word formation due to a wrong †analogy (*gooses* for *geese*). 332

manner **1** (*phonet*) The specific process of †articulation used in a sound's †production (†plosive, etc.). 159 **2** (*gram*) An †adverbial answering the question 'how?' (*quickly*). 91

manual alphabet *see* **dactylology**

manualism (*clin*) The teaching of †sign[3] to the deaf, to the exclusion of speech; cf. †oralism. 269

margins (*phonet*) Sound †segments that form the boundaries of a †syllable ([k], [p] in *cup*). 166

marking/markedness (*ling*) The presence/absence of a particular †contrastive feature in a language or languages. 85

masculine *see* **gender**

mass (*gram*) Said of †nouns that typically express general concepts and lack an indefinite †article or †plural (*information*); cf. †countable. 91

matched guise (*socio*) Recording two languages or †dialects by the same speaker, in order to elicit listener †language attitudes. 23

mathematical linguistics (*ling*) The study of the mathematical properties of language, esp. using statistical or algebraic concepts. 418

matronymic (*sem*) A name derived from that of a person's mother (*Marjorison*); also, **metronymic**. 112

maxims *see* **conversational maxims**

measure *see* **metre**

mechanical translation *see* **automatic translation**

medium (*gen*) A dimension of message transmission, esp. speech, writing, †sign[3]. 123

mel (*acou*) A unit of measurement for †pitch. 144

mentalistic (*ling*) Said of the study of language through introspection rather than through the description of behaviour; cf. †behaviourism. 413

merger (*hist*) The coming together of linguistic units that were originally distinguishable. 330

mesolect (*socio*) In †creole studies, a †variety between †acrolect and †basilect. 338

metalanguage (*ling*) A language used for talking about language. 250

metanalysis (*hist*) A word deriving from a word-boundary error (E. *a naddre* → *an adder*). 330

metaphor (*gen*) A †figurative expression in which one notion is described in terms usually associated with another (*launch an idea*). 70

metathesis (*ling*) Alteration in a normal sequence of elements, esp. sounds (*aks* for *ask*). 330

metonymy (*hist*) A †semantic change where an attribute is used for the whole (*crown* = *king*). 70

metre/meter (*poet*) A rhythmical verse pattern; also, **measure**. 74

metrics (*poet*) The study of metrical structure. 74

metronymic *see* **matronymic**

microlinguistic (*ling*) Said of highly detailed studies of language data. 408

mid (*phonet*) Said of a †vowel[1] †articulated between †high[1] and †low[1] tongue positions ([e], [ʌ]); cf. †close, †open[3]. 156

middle ear (*anat*) Part of the ear between the ear drum and the †inner ear. 142

minim (*graph*) A single downstroke of the pen. 189

minimal pair (*phonol*) Words that differ in meaning when only one sound is changed. 162

minuscule (*graph*) A form of writing consisting of small letters; cf. †majuscule. 188

miscue (*app*) An error made by someone learning to read; studied by **miscue analysis**. 252

mismatch (*psycho*) A child's †semantically inappropriate use of a word, where there is no apparent basis for the error. 246

mistake *see* **error[2]**

mixing *see* **code-switching**

modal (*gram*) A †verb that signals contrasts in speaker attitude (†mood), e.g. *may, can*. 93

modality **1**(*semiot*) A †medium of communication. 400 **2** (*gram*) The system of †modal expression. 93

mode (*semiot*) A †medium of communication. 400

modelling (*app*) Providing language examples for a learner to follow. 372

modification **1** (*gram*) The structural dependence of one element (a **modifier**) upon another. 95 **2** (*phonet*) Movement that affects the air flow in the †vocal tract. 130 **3** (*hist*) Any of several kinds of †formal[1] change in a word (*man* → *men*). 330

modifier *see* **modification**

monaural (*phonet*) Using one ear; cf. †binaural. 142

monitoring (*app*) Critical self-listening. 374

monogenesis (*hist*) The view that all languages come from an original language; cf. †polygenesis. 293

monoglot *see* **monolingual**

monolingual (*gen*) Said of a person/community with only one language; also, **unilingual**; cf. †bilingual, †multilingual. 362

monologue (*gen*) Speech by an individual person. 48

monometer (*poet*) A line of verse containing a single unit of rhythm (†foot). 74

monomorphemic (*gram*) Said of a †word consisting of a single †morpheme. 90

monophthong (*phonet*) A †vowel[1] with no detectable change in quality during a †syllable (*car*). 156

monosyllabic (*phonol*) Said of a †word consisting of a single †syllable. 86

mood (*gram*) Attitudes of fact, wish, possibility, etc. conveyed by a †verb (a †modal) or †clause, e.g. †indicative, †subjunctive. 93

mora (*phonol*) A minimal unit of rhythmical time equivalent to a short †syllable. 74

morph (*gram*) The physical form of a †morpheme. 90

morpheme (*gram*) The smallest †contrastive unit of †grammar (*man, de-, -tion, -s*, etc.); cf. †bound form, †free form. 90

morphemics (*gram*) The study of †morphemes. 90

morphology (*gram*) The study of †word structure, esp. in terms of †morphemes. 90

morphophonemics *see* **morphophonology**

morphophonology (*gram*) The study of the relations between †morphology and †phonology. 90

morphosyntactic (*gram*) Said of a category whose definition involves both †morphology and †syntax, e.g. †number. 90

morphotactics (*gram*) The arrangement of †morphemes in a linear sequence. 90

motherese *see* **caretaker speech**

mother tongue *see* **first language**

motor phonetics *see* **articulation**

move (*prag*) A unit of speech in a †discourse. 116

multilingual (*gen*) Said of a person/community with several languages; cf. †monolingual. 362

mutation (*gram, hist*) A sound change in a word due to the influence of adjacent †morphemes or †words (Welsh ***p****en* 'head' → *fy* ***mh****en* 'my head'). 90

mutism (*clin*) Involuntary inability to speak. 266

mytheme (*styl*) The smallest contrastive unit of structure found in mythical narratives. 79

narrow (*phonet*) Said of a †transcription that shows many †phonetic details; cf. †broad. 160

nasal *see* **nasality**

nasality (*phonet*) Sound made with the soft †palate lowered, thus allowing air to resonate in the nose (**nasals**), e.g. [m], [n], or **nasalized** sounds, e.g. [ã]. 130

nasopharynx (*anat*) The part of the †pharynx adjoining the nasal †cavity. 130

native language *see* **first language**

native speaker (*gen*) A person whose language is a †first language or 'mother tongue'. 372

nativism *see* **innateness hypothesis**

naturalism (*sem*) The view that there is a close, 'natural' connection between words and things; cf. †conventionalism. 408

natural language (*gen*) A language with ⁺native speakers; cf. ⁺auxiliary language[2], ⁺language[2]. 354

negation (*gram*) A process expressing the denial or contradiction of some or all of the meaning of a sentence; **negative** forms (**negators**) include *not*, *un-*, etc.; cf. ⁺affirmative. 245

negative, negator *see* **negation**

neologism (*gen*) A new or invented word or expression (*linguistified*). 73

neurolinguistics (*ling*) The study of brain structure and function in relation to language use, acquisition, and disorder. 263

neuter *see* **gender**

neutralization (*phonol*) The loss of a ⁺contrast between two ⁺phonemes in a particular ⁺environment[1] (/t/ vs /d/ is 'neutralized' in *stop*). 163

neutral vowel *see* **schwa**

new *see* **comment**

node *see* **nodule**

nodule (*clin*) A small localized swelling ('node'), esp. on the ⁺vocal folds. 278

noise (*acou*) A complex ⁺sound wave with irregular vibrations. 137

nomenclature (*gen*) A system of terms used in a specialized field. 384

nominal (*gram*) A ⁺noun or noun-like item. 91

nominalism *see* **conventionalism**

nominalization (*gram*) Forming a ⁺noun from some other ⁺word class (*redness, my answering ...*). 91

nominative (*gram*) An ⁺inflection[1] that typically identifies the ⁺subject of a ⁺verb (Ger. ***Der Mann*** *seht den Mann* 'The man sees the man'). 93

nonce formation (*ling*) An invented or accidental linguistic form, used once only (*brillig*). 90

non-count *see* **countable**

non-defining *see* **restrictive**

non-finite *see* **finite**

non-restrictive *see* **restrictive**

non-standard *see* **standard**

non-verbal (*semiot*) Said of communication that does not use words, e.g. gestural. 403

normative *see* **prescription**

notation *see* **transcription**

notional **1** (*gram*) Said of a grammar whose terms rely on ⁺extralinguistic notions, e.g. action, duration, time; cf. ⁺formal[1]. 91 **2** (*app*) Said of a syllabus based on an analysis of sentence meanings and functions; cf. ⁺communicative approach. 378

noun (*gram*) A ⁺word class with a naming function, typically showing contrasts of ⁺countability and ⁺number, and capable of acting as ⁺subject or ⁺object of a ⁺clause. 91

noun phrase (*gram*) A ⁺phrase with a ⁺noun as ⁺head (*the tall man in a hat*). 95

nuclear *see* **nucleus**

nucleus (*phonol*) The ⁺syllable in a ⁺tone group that carries maximum ⁺pitch prominence (**nuclear tone, tonic**) (*She went to* ***London***). 172

number (*gram*) The grammatical category that expresses such contrasts as '⁺singular/⁺plural/⁺dual' (*cat/cats, she is/they are*). 93

object (*gram*) A ⁺clause element that expresses the result of an action (cf. ⁺direct/⁺indirect object). 95

objective *see* **accusative**

object language (*ling*) A language that is the object of analysis (using a ⁺metalanguage). 82

oblique (*gram*) Said of any ⁺case form of a ⁺word except the ⁺nominative. 93

obsolescent (*gen*) Said of a word or sense no longer used. 332

obstruent (*phonet*) Sounds made with a constriction (⁺plosives, ⁺fricatives, ⁺affricates). 159

obviative (*gram*) A fourth-⁺person form used in some languages, typically contrasting with the third person to mean 'someone/something else'. 92

occlusion (*phonet*) The length of the ⁺closure during the ⁺articulation of a ⁺stop ⁺consonant. 159

oesophageal/esophageal (*phonet*) Said of sounds or ⁺voice[1] produced in the upper part of the oesophagus, esp. after ⁺laryngectomy. 278

off-glide, on-glide *see* **glide**

onomasiology (*sem*) The study of sets of associated concepts in relation to their linguistic forms. 100

onomastics (*sem*) The study of the ⁺etymology and use of ⁺proper names. 112

onomatology *see* **onomastics**

onomatopoeia (*sem, poet*) Words that imitate the sounds of the world (*splash, murmur*). 176

ontogeny (*ling*) Growth and decay (here, of language) in the individual; cf. ⁺phylogeny. 230

open **1** (*gram*) Said of a ⁺word class with unlimited membership (⁺noun, ⁺adjective, ⁺adverb, ⁺verb); cf. ⁺closed[1]. 91 **2** (*phonol*) Said of a ⁺syllable that ends in a ⁺vowel[1]; cf. ⁺closed[2]. 166 **3** (*phonet*) Said of ⁺vowels[1] made with the tongue in the lowest possible position ([a], [ɑ]); ⁺vowels a degree higher are **half-/mid-open**. 153

opposition (*phonol*) A linguistically important contrast between sounds. 162

optative (*gram*) A ⁺mood of the ⁺verb, in some languages expressing desire or wish. 93

oracy (*app*) Ability in speaking and listening. 250

oral (*phonet*) Said of sounds made in the mouth (as opposed to the nose, ⁺**nasal**). 152

oralism (*clin*) The teaching of speech to the deaf, to the exclusion of ⁺sign[3]; cf. ⁺manualism. 269

ordinal (*gram*) A class of numerals (*first*, etc.); cf. ⁺cardinal number. 99

oropharynx (*anat*) The part of the ⁺pharynx adjacent to the oral cavity. 130

orthoepy (*gen*) The study of correct pronunciation, esp. as practised in the 17th/18th centuries. 331

orthography (*gen*) The study of the use of letters and the rules of spelling in a language. 196

orthophonist *see* **speech pathologist**

oscillograph (*acou*) An instrument that provides a graphic representation of ⁺sound waves (an **oscillogram**). 138

oscilloscope (*acou*) An instrument for the visual display of ⁺sound waves. 138

ossicles (*anat*) The bones of the ⁺middle ear. 143

otology (*clin*) The study of diseases of the ear. 268

oto(rhino)laryngology (*clin*) The study of diseases of the ear, nose, and throat. 268

overcorrection *see* **hypercorrection**

overextension *see* **overgeneralization**

overgeneralization (*psycho*) A learner's extension of a word meaning or grammatical ⁺rule[1] beyond its normal use (*men* → *mens*); also, **overextension**. 246

overtone *see* **harmonic**

oxymoron (*rhet*) A ⁺figurative combination of incongruous or contradictory words. 70

oxytone (*gen*) A word with heavy ⁺stress on the final ⁺syllable (*represent*). 171

paedography (*graph*) A writing system devised to help children to read. 196

palaeography (*graph*) The study of ancient writings and inscriptions. 189

palatal (*phonet*) Said of sounds made in the area of the hard ⁺palate ([ç], [j]). 157

palatalization (*phonet*) An ⁺articulation in which the tongue moves towards the hard ⁺palate while another sound is being made. 158

palate (*anat*) The arched bony structure that forms the roof of the mouth; divided into the **hard palate** and **soft palate** (**velum**). 124

palato-alveolar (*phonet*) Said of a ⁺consonant made between the ⁺alveolar ridge and the hard ⁺palate ([ʃ]). 157

palatography (*phonet*) The instrumental study of tongue contact with the ⁺palate, displayed as a **palatogram**. 140

palilalia (*clin*) Involuntary repetition of words or phrases. 272

palilology (*rhet*) Word repetition for emphasis. 70

palindrome (*gen*) Words or expressions that read the same backwards or forwards. 65

pangram (*gen*) A sentence that contains every letter of the alphabet. 65

paracusis (*clin*) Any hearing abnormality. 268

paradigm (*gram*) The set of ⁺inflectional[1] ⁺forms[1] of a word (Lat. *amo/amas/amat...*). 90

paradigmatic (*ling*) The relationship of ⁺substitution between a linguistic unit and other units at a particular place in a ⁺structure[2]. 411

paradox (*gen*) An apparent contradiction that contains a truth. 70

paragram (*gen*) A play on words by altering a letter, esp. in humour. 63

paralanguage (*ling*) Features of speech or ⁺body language considered to be marginal to language; studied by **paralinguistics**. 171

paralinguistics *see* **paralanguage**

parallelism (*styl*) The use of paired sounds, words, or constructions. 60

paraphasia (*clin*) An involuntary error in the production of words or phrases. 272

paraphrase (*gen*) An alternative version of a sentence that does not change its meaning. 107

pararhyme (*poet*) The repetition of the same initial and final consonants in different words (*tail/tall*). 74

parataxis (*gram*) Constructions joined without the use of †conjunctions (*I had tea, eggs...*); cf. †hypotaxis. 95

parent language (*hist*) A language from which other languages descend, e.g. Latin is the parent of **daughter** languages French, Spanish, etc., which are thus **sister** languages to each other. 294

parole (*ling*) The concrete utterances of a speaker; cf. †*langue*. 411

paronomasia (*gen*) A play on words, or pun. 63

paronym (*hist*) A word that comes from the same †root[2] as another (*wise/wisdom*). 90

paroxytone (*phonol*) A word with heavy †stress on the penultimate syllable (*telegraphic*). 171

parsing (*gram*) Analysing and labelling the grammatical elements of a †sentence; also, **diagramming**. 251

participle (*gram*) A word derived from a †verb and used as an †adjective (*a smiling face*). 91

particle (*gram*) An †invariable word with a †grammatical[1] function (***to go, not***). 91

partitive (*gram*) A form that refers to a part or quantity (*some, piece, ounce*). 92

part of speech *see* **word class**

pasigraphy (*gen*) The use of a system of symbols understood between languages (*1, 2, +, £*). 202

passive *see* **active, voice**[2]

patient *see* **goal**

patois (*gen*) A provincial †dialect. 24

patronymic (*gen*) A name derived from that of a person's father (*Peterson*). 112

pejoration *see* **deterioration**

pejorative (*gen*) Said of a linguistic form that expresses a disparaging meaning (*goodish*). 332

pentameter/pentametre (*poet*) A line of verse containing five units of rhythm (†foot). 74

perfect (*gram*) A †tense[1] form typically referring to a past action that has present relevance (*I have asked*); cf. †pluperfect. 93

perfective (*gram*) A †verb †aspect typically stressing the completion of an action; contrasts with **imperfective**. 93

performance (*ling*) The language actually used by people in speaking or writing; cf. †competence. 413

performative (*prag*) An †utterance or †verb that performs an action (*promise, baptise*). 121

periodic (*acou*) Said of a †waveform that involves a repeated pattern of vibration; contrasts with **aperiodic** (random) vibration. 133

periphrasis 1 (*gram*) The use of separate †words instead of †inflections[1] to express a †grammatical[1] relationship (**periphrastic**) (*more happy* for *happier*). 92 2 *see* **circumlocution**

perlocutionary act (*prag*) A †speech act that achieves a particular effect on a listener (frightens, persuades); cf. †locutionary act. 121

perseveration (*clin*) Involuntary continued use of a linguistic form. 273

person (*gram*) A grammatical form (esp. a †pronoun or †verb) referring directly to the speaker ('first person'), addressee ('second person'), or others involved in an interaction (esp. 'third person'). 93

personal pronoun *see* **person**

personification (*poet, rhet*) The †figurative attribution of human qualities to non-human notions. 70

petroglyph (*gen*) An ancient stone inscription; also, **petrogram**. 198

petrogram *see* **petroglyph**

pharyngeal (*phonet*) Said of sounds made in the †pharynx ([ħ], [ʕ]). 157

pharyngealization (*phonet*) Narrowing of the †pharynx while another speech sound is being made. 158

pharynx (*anat*) The part of the throat above the †larynx. 130

phatic (*ling*) Said of language used to establish atmosphere or maintain social contact. 10

philology *see* **comparative philology**

philosophical linguistics (*ling*) The study of language in relation to philosophical concepts. 418

phon (*acou*) Unit of measurement for the †loudness level of sound. 144

phon(a)esthenia (*clin*) An abnormally weak †voice quality[1]. 278

phon(a)esthetics (*phonet*) The study of the aesthetic or symbolic properties of sound. 176

phonation (*phonet*) The production of †voice[1] through the use of the †vocal folds. 128

phone (*phonet*) The smallest perceptible †discrete †segment of speech sound. 152

phoneme (*phonol*) The smallest †contrastive unit in the sound system of a language. 162

phonemics (*phonol*) The analysis of †phonemes. 162

phonemic transcription (*phonol*) A †transcription of the †phonemes in an utterance. 162

phonetic 1 (*phonet*) Pertaining to phonetics. 152 2 (*graph*) Part of a †logogram that indicates its pronunciation; cf. †determinative. 202

phonetic alphabet *see* **phonetic transcription**

phonetician (*phonet*) A †phonetics specialist. 152

phonetics (*phonet*) The science of speech sounds, esp. of their production, transmission, and reception ('†acoustic/†articulatory/†auditory phonetics'). 152

phonetic spelling (*gen*) A spelling system that represents speech sounds in a one-to-one way. 215

phonetic transcription (*phonet*) A †transcription of all distinguishable phones in an utterance, using special symbols (a **phonetic alphabet**). 160

phoniatrics (*clin*) Study of pathologies affecting †voice quality[1] and pronunciation. 266

phonics (*app*) A method of teaching reading that trains recognition of the sound values of individual letters; cf. †look-and-say. 253

phonogram (*graph*) A symbol representing a speech sound; cf. †logogram. 201

phonography (*graph*) A writing system that represents individual speech sounds. 199

phonologist (*phonol*) A †phonology specialist. 162

phonology (*phonol*) The study of the sound systems of languages. 162

phonostylistics (*poet*) The study of the expressive use of sound, esp. in poetry. 74

phonotactics (*phonol*) The specific sequences of sounds that occur in a language. 162

phrasal verb (*gram*) A †verb consisting of a lexical element and †particle(s) (*get up*). 91

phrase (*gram*) A group of words smaller than a †clause, forming a †grammatical[1] unit (*in a box*). 95

phrase marker (*gram*) A structural representation of a sentence in †generative grammar, usually in the form of a †tree diagram. 96

phrase-structure grammar (*gram*) A †generative grammar that provides an analysis of sentences into †constituent elements. 96

phylogeny (*hist*) Historical development (here, of language) in communities or in the human race as a whole; cf. †ontogeny. 330

physiological phonetics *see* **articulatory phonetics**

pictogram/pictograph (*graph*) A symbol used in picture writing. 199

pidgin (*socio*) A language with a reduced range of structure and use, with no †native speakers. 336

pidginize (*socio*) To develop into a †pidgin. 336

pitch (*phonet*) The auditory sensation of the height of a sound. 133

place of articulation (*phonet*) The anatomical point in the †vocal tract where a speech sound is produced (†labial, †dental, etc.). 157

pleonasm (*gen*) The unnecessary use of words (*in this present day and age*). 2

plethysmograph (*phonet*) An instrument that records changes in air volume during speech. 125

plosive (*phonet*) Said of a †consonant made by the sudden release of a complete †closure in the †vocal tract ([p], [k]). 159

pluperfect (*gram*) A †verb form that typically expresses completion of an action before a specific past time (*I had jumped*); also, **past perfect**. 93

plural (*gram*) A †word form typically expressing 'more than one' in †number (*cats, them*). 93

plurilingualism *see* **multilingualism**

plurisegmental *see* **suprasegmental**

pneumograph (*phonet*) An instrument that measures chest movements during breathing. 139

pneumotachograph (*phonet*) An instrument that measures air flow from nose and mouth. 139

poetics (*poet*) The linguistic analysis of poetry (and sometimes of other creative language use). 73

point size (*graph*) A system for measuring the size of pieces of type. 192

polarity (*gram*) The system of contrast between ⁺affirmative and ⁺negative in a language. 93

polyalphabetic (*gen*) Said of a ⁺cipher that makes use of many letter transformations. 58

polygenesis (*hist*) The view that languages come from several original sources; cf. ⁺monogenesis. 293

polyglot/polylingual *see* **multilingual**

polysemia/polysemy (*sem*) Several meanings of a word (*plain* = 'dull/obvious/...'). 106

polysemic/polysemous (*sem*) Showing ⁺polysemy. 106

polysyllabic (*phonet*) Having more than one ⁺syllable. 87

polysynthetic (*ling*) Said of a language that uses long ⁺word forms with a complex ⁺morphology; also, **incorporating**, **incapsulating**. 295

polysystemic (*ling*) Said of an analysis that sets up different linguistic systems at different places in ⁺structure[2]. 412

popular etymology *see* **folk etymology**

portmanteau (*gram*) A ⁺morph that can be analysed into more than one ⁺morpheme (Fr. *au* = *à le*). 90

positive *see* **affirmative**, **degree**, **polarity**

possessive (*gram*) A linguistic form that indicates possession (*my*, *mine*, *Mary's*). 93

postalveolar (*phonet*) Said of a ⁺consonant made at the rear of the ⁺alveolar ridge. 157

postcreole continuum (*socio*) A related series of ⁺varieties that develops when ⁺creole speakers are taught in the ⁺standard language. 338

postmodification (*gram*) Items that occur within a ⁺phrase after the ⁺head (*the man* ***in a suit***). 95

postposition (*gram*) A ⁺particle that follows the ⁺noun it ⁺governs (Jap. *X kara Y made* 'from X to Y'); cf. ⁺preposition. 92

post-structuralism (*styl*) A reaction to the ⁺structuralist analysis of literary texts. 79

postvocalic (*phonet*) Following a ⁺vowel[1]. 166

pragmatics (*prag*) The study of the factors influencing a person's choice of language[1]. 120

predicate (*gram*) The ⁺clause element that gives information about the ⁺subject (*She* ***saw a dog***). 94

predicative *see* **attributive**

prefix (*gram*) An ⁺affix added initially to a ⁺root[1] (***un****happy*). 90

prelinguistic (*psycho*) Said of child utterance before the emergence of language. 230

preliterate (*hist*) Said of a language before a writing system has developed. 198

premodification (*gram*) Items that occur within a ⁺phrase before the ⁺head (***the funny*** *clown*). 95

preposition (*gram*) Items that ⁺govern and typically precede ⁺nouns, ⁺pronouns, and certain other forms (***in*** *the box*, ***to*** *me*, ***by*** *running*). 91

presby(a)cusis (*clin*) Gradual loss of the ability to hear and distinguish sounds as a result of old age. 268

prescription (*gen*) An authoritarian (**prescriptive** or **normative**) statement about the correctness of a particular use of language; cf. ⁺description. 2

prescriptive *see* **prescription**

presupposition (*sem*) The information that a speaker assumes to be already known; cf. ⁺focus. 120

preterite (*gram*) The ⁺simple[1] past ⁺tense form of a ⁺verb (*I* ***saw***). 93

prevocalic (*phonet*) Preceding a ⁺vowel[1]. 166

principal parts (*gram*) The ⁺forms[2] of a verb required to determine its ⁺conjugation (Lat. *amo/amare/amavi/amatum*). 91

proclitic (*gram*) An unstressed word that depends on and is pronounced with a following word (*an*). cf. ⁺enclitic. 91

production (*ling*) The active use of language; cf. ⁺comprehension. 263

productivity (*ling*) The creative capacity of language users to produce and understand an indefinitely large number of sentences. 401

proficiency test (*app*) A test that measures how much of a language someone knows. 381

pro-form (*gram*) An item that substitutes for another item or construction (***so does*** *John*). 119

prognostic test *see* **aptitude test**

progressive **1** (*gram*) A ⁺verb form that typically expresses duration or incompleteness (*He* ***is running***); also, **continuous**; cf. ⁺simple. 93 **2** (*phonol*) Said of an ⁺assimilation when one sound causes a change in the following sound ([ʃ] → [ʧ] in *did she*). 166

prolongation (*clin*) The abnormal or controlled lengthening of a sound in ⁺stuttering. 280

prominence (*phonet*) The degree to which an element stands out from others in its ⁺environment[1]. 171

pronominal (*gram*) An item that functions as a ⁺pronoun. 91

pronoun (*gram*) An item that can substitute for a ⁺noun or ⁺noun phrase (*it*, *who*, *himself*). 91

proper name/noun (*gram*) A ⁺noun that labels a unique place, person, animal, etc. and lacks the grammatical forms of a ⁺common noun. 112

proposition (*sem*) A unit of meaning in ⁺statement form that is asserted to be true or false (*The cat is asleep*). 107

prop word *see* **empty word**

proscriptive (*ling*) Said of ⁺prescriptive ⁺rules[2] that forbid a usage, e.g. criticism of *very unique*. 2

prosodic features *see* **prosody**[1]

prosody **1** (*phonol*) The linguistic use of ⁺pitch, ⁺loudness, ⁺tempo, and ⁺rhythm. 171 **2** (*poet*) The study of versification. 74

pro(s)thesis (*phonol*) The insertion of an extra sound at the beginning of a word. 330

proto-language **1** (*hist*) A hypothetical ancestor language or form ('Proto-Indo-European'). 294 **2** (*psycho*) A stage before the emergence of a recognized linguistic form (**proto-word**). 239

proto-word *see* **proto-language**

proverb (*gen*) A short, pithy, rhythmical saying expressing a general belief. 53

proxemics (*semiot*) The study of the communicative function of body distance, posture, etc. 403

pseudepigraphy (*gen*) The false ascription of an author's name to a written work. 189

pseudolinguistic (*gen*) Said of vocal behaviour with a superficial resemblance to language. 11

pseudonym (*gen*) A fictitious name, esp. of an author. 112

psittacism (*gen*) Meaningless repetitive ('parrot-like') speech. 272

psycholinguistics (*psycho*) The study of language in relation to psychological processes. 418

pulmonic (*phonet*) Pertaining to the lungs. 125

pure tone (*acou*) A ⁺sound wave of a single ⁺frequency; cf. ⁺complex tone. 132

pure vowel (*phonet*) A ⁺vowel[1] that does not change in quality during a ⁺syllable; cf. ⁺diphthong. 156

purism (*gen*) The view that a language needs to preserve traditional standards of correctness and be protected from foreign influence. 2

qualifier (*gram*) A word or phrase that limits the meaning of another element (***red*** *car*). 95

quality (*phonet*) The characteristic ⁺resonance, or ⁺timbre, of a sound. 133

quantifier (*sem*) An item expressing amount (*all*, *some*, *each*). 91

quantitative linguistics *see* **mathematical linguistics**

quantity (*phonol*) The relative ⁺duration of ⁺contrastive sounds and syllables. 171

question (*gram*) A sentence that asks for information or a response. 121

radical *see* **determinative**

rank *see* **level**[3]

readability formula (*app*) A measure of the ease with which a written text can be read. 254

realization (*phonol*) The physical expression of an abstract linguistic unit. 82

rebus (*gen*) A combination of letters, pictures, and pictograms to make words and sentences. 65

received pronunciation (*phonol*) The regionally neutral, prestige accent of British English. 39

receptive aphasia (*clin*) A disorder of language[4] ⁺comprehension caused by brain damage; cf. ⁺expressive[2]. 272

recipient *see* **goal**

reciprocal **1** (*gram*) An item that expresses the meaning of mutual relationship (*each other*). 91 **2** (*phonol*) A type of ⁺assimilation in which sounds influence each other. 166

reconstruction (*hist*) The ⁺comparative linguistic analysis of extant texts to work out an earlier, non-extant state of a language. 294

recursive (*gram*) Said of a ⁺grammatical[1] ⁺rule[1] that is capable of repeated application. 97

reduction **1** (*gram*) The lack of one or more of the normal ⁺constituents in a construction (*gone to town*); cf. ⁺ellipsis. 95 **2** (*phonol*) A ⁺vowel[1] that becomes ⁺central when a word is ⁺unstressed

([a] → [ə] as in *she can* → *she c'n go*). 166
3 (*hist*) A narrowing of meaning. 332
4 *see* **clipping**.

redundant (*ling*) Said of a feature that is unnecessary for the identification or maintenance of a linguistic ⁺contrast. 146

reduplication (*gram*) **1** A ⁺form² involving a repeated element (Lat. *curro* 'run' → *cucurri* 'ran'). 177 **2** A type of ⁺compound¹ word using repeated elements (*helter-skelter*). 90

reference (*sem*) The relationship between linguistic forms and entities in the world (**referents**). 102

referent *see* **reference**

referential *see* **denotation, reference**

reflexive (*gram*) A construction or ⁺verb in which ⁺subject and ⁺object relate to the same entity (*She washed herself*). 93

reflexiveness (*semiot*) The capability of language to 'talk about' itself; cf. ⁺metalanguage. 401

regional dialect *see* **dialect**

register **1** (*phonet*) A physiologically determined range of the human ⁺voice¹, e.g. falsetto. 18 **2** (*socio*) A socially defined ⁺variety of language, e.g. scientific, legal, etc. 52 **3** (*phonol*) Said of a ⁺tone language that does not use ⁺gliding tones. 174

regression (*psycho*) A backward eye movement while reading a line of print. 210

regressive (*phonol*) Said of an ⁺assimilation when one sound causes a change in the preceding sound ([t] → [p] in *hot pig*); also, **anticipatory**. 166

regular (*ling*) Said of a linguistic form that conforms to the ⁺rules¹ of the language. 408

related (*hist*) Said of languages or forms that share a common origin. 294

relative clause *see* **relative pronoun**

relative pronoun (*gram*) The item that introduces a ⁺dependent ⁺clause (**relative clause**) in a ⁺noun phrase, referring back to the ⁺noun (*the car* ***which*** *was sold...*). 95

relativity *see* **linguistic relativity**

release (*phonet*) ⁺Vocal organ movement away from a point of ⁺articulation, esp. in ⁺plosives. 159

relexification (*socio*) A process in the development of ⁺pidgins in which original Portuguese vocabulary is replaced by native language words. 339

relic area (*socio*) A ⁺dialect area that preserves linguistic features from an earlier period. 28

repair (*prag*) The correction of a misunderstanding or error made during a conversation. 116

repertoire (*socio*) The range of languages or ⁺varieties that a speaker has available. 48

reported speech *see* **indirect speech**

resonance (*phonet*) Air vibrations in the ⁺vocal tract that are set in motion by ⁺phonation. 130

respiration (*phys*) The act of breathing. 124

restricted code (*socio*) An informal ⁺variety of language thought to display a reduced range of structures; cf. ⁺elaborated code. 40

restricted language (*socio*) A highly reduced linguistic system found in narrowly defined settings, e.g. heraldry, weather reporting. 56

restrictive (*gram*) Said of a ⁺modifier that is an essential part of the identity of another element (*my brother* ***who's abroad***); also, **defining**; contrasts with **non-restrictive** or **non-defining**, where the modification is not essential (*my brother,* ***who's abroad***). 95

retracted (*phonet*) Said of the backwards movement of an ⁺articulator, e.g. the ⁺apex of the tongue. 157

retroflex (*phonet*) Said of sounds made when the ⁺apex of the tongue is curled back in the direction of the hard ⁺palate ([ʈ], [ɽ]. 157

rewrite rule (*gram*) A ⁺rule¹ in ⁺generative grammar of the form 'X → Y' (= 'replace X by Y'). 97

rheme (*ling*) The new information conveyed in a sentence; cf. ⁺theme. 120

rhetoric (*rhet*) The study of effective speaking and writing. 70

rhetorical question (*gram*) A ⁺question to which no answer is expected. 121

rhinolalia/rhinophonia (*clin*) ⁺Nasal resonance. 278

rhotacism (*clin*) A defective use of [r]. 279

rhotic area (*socio*) A ⁺dialect area in which /r/ is pronounced following a ⁺vowel (*car*). 28

rhoticization (*phonet*) The ⁺articulation of ⁺vowels¹ with *r*-colouring. 153

rhyme (*poet*) A correspondence of ⁺syllables, esp. at the ends of poetic lines. 74

rhythm (*phonol*) The perceived regularity of prominent units in speech. 171

roll *see* **trill**

romanization (*graph*) The use of the Latin alphabet to transcribe non-Latin writing systems. 315

root **1** (*gram*) The basic form of a word, from which other words derive (***meaningfulness***); cf. ⁺stem. 90 **2** (*hist*) The earliest form of a word. 332 **3** (*phonet*) The furthest-back part of the tongue. 131 **4** *see* **isolating language**

rounded *see* **rounding**

rounding (*phonet*) The visual appearance of the lips, permitting ⁺contrasts of **rounded** ([u]) and **unrounded** ([i]). 152

routine *see* **formulaic**

rule (*gram*) **1** A generalization about linguistic structure. 97 **2** A ⁺prescriptive recommendation about correct usage. 3

saccades (*psycho*) Rapid eye movements used when searching for an object. 210

salience (*phonet, psycho*) The perceptual prominence of a sound. 145

sandhi (*gram*) A sound change affecting a word used in a specific grammatical ⁺context¹ (*do* → *don't*). 409

satem language (*hist*) An Indo-European language that replaced [k] by [s] in such words as *centum* ('hundred'); cf. ⁺centum language. 330

scansion (*poet*) The analysis of ⁺metre. 74

scheme (*rhet*) A ⁺figurative effect, e.g. ⁺rhyme, that changes the structure of language without affecting its meaning; cf. ⁺trope. 70

schwa/shwa (*phonet*) An ⁺unstressed ⁺vowel¹ [ə] made in the centre of the mouth, heard at the end of such words as *after* and *the*. 153

script (*graph*) Any system of written signs. 196

secondary articulation (*phonet*) The lesser point of ⁺stricture in a sound involving two points of ⁺articulation, e.g. lip ⁺rounding. 158

second language (*app*) A non-native language, esp. one that has an official role in a country. 372

second person *see* **person**

segment (*phonet*) A ⁺discrete unit that can be identified in the stream of speech. 163

segmental phonology (*phonol*) The analysis of speech into ⁺phones or ⁺phonemes; cf. ⁺suprasegmental phonology. 162

segmentation (*phonet, gram*) The process of analysing speech into ⁺segments. 96, 162

selection(al) features *see* **collocation**

semantic *see* **semantics**

semantic component (*sem*) An element of a word's meaning (*girl* → young, female, human). 107

semantic differential (*psycho*) A technique for measuring the emotional associations of words. 103

semantic feature *see* **semantic component**

semantic field (*sem*) An area of meaning identified by a set of mutually defining words (colour, furniture, etc.). 104

semantic relations (*sem*) The ⁺sense relations that exist between words, e.g. ⁺hyponymy. 105

semantics (*sem*) The study of linguistic meaning; also, semasiology, sematology, semology. 100

semasiology/sematology *see* **semantics**

semi-consonant *see* **semi-vowel**

semiology *see* **semiotics**

semiotics (*semiot*) The study of the properties of signs and signalling systems, esp. as found in all forms of human communication; also, **semiology**, **significs**. 403

semi-vowel (*phonet*) A sound that displays certain properties of both ⁺consonants and ⁺vowels¹ ([l], [j]); also, **semi-consonant**. 153

semology *see* **semantics**

sense relations (*sem*) The meaning relations between words, as identified by the use of ⁺synonyms, ⁺antonyms, etc.; cf. ⁺reference. 102

sensorineural (*clin*) Said of hearing loss due to damage to the ⁺inner ear. 268

sentence (*gram*) The largest structural unit that displays stateable ⁺grammatical¹ relationships, not ⁺dependent on any other ⁺structure². 94

sequencing **1** (*psycho*) Psychological processing of a series of linguistic elements. 277 **2** (*app*) The order in which a graded series of items is given to a learner. 378 **3** (*prag*) The rule-governed succession of utterances in a ⁺discourse. 120

shwa *see* **schwa**

sibilant (*phonet*) A †fricative made with a groove-like †stricture in the front part of the tongue, to produce a hissing sound ([s], [ʃ]). 159

sight vocabulary (*app*) Words that can be recognized as wholes by someone learning to read. 252

sigmatism **1** (*clin*) Abnormal pronunciation of [s], esp. as a †lisp. 279 **2** (*poet*) The repetitive use of [s] for effect. 74

sign **1** (*semiot*) A feature of language or behaviour that conveys meaning, esp. as used conventionally in a system; also, **symbol**. 411 **2** (*graph*) A mark used as an element in a writing system; also, **symbol**. 196 3 (*ling*) Deaf †sign language. 222

signifiant (*sem*) That which signifies; contrasts with ***signifié***, that which is signified. 411

significant *see* **contrastive**

signification (*sem*) The relationship between signs and the things or concepts to which they refer. 100

significs *see* **semiotics**

signifié *see* ***signifiant***

sign language (*ling*) A system of manual communication, esp. one used by the deaf. 222

simile (*rhet*) A †figurative expression that makes an explicit comparison (*as tall as a tower*). 70

simple (*gram*) **1** Said of a †tense[1] form that has no †auxiliary verb ('simple present' *He runs*, etc.); cf. †progressive[1]. 93 **2** Said of a †sentence containing one †clause; cf. †complex sentence, †compound[1]. 95

sine wave (*acou*) A simple †waveform that produces a †pure tone. 132

singular (*gram*) A form that typically expresses 'one of' in †number (***dog***, ***It is***). 93

sister language *see* **parent language**

situation (*ling*) The †extralinguistic setting in which a use of language takes place. 48

slang (*gen*) **1** Informal, †non-standard vocabulary. 53 **2** The †jargon[1] of a special group. 56

slot (*gram*) A place in a construction where a class of items can be inserted (*the – car*). 95

social dialect *see* **dialect**

sociolect (*socio*) A social †dialect. 38

sociolinguistics (*socio*) The study of the relationship between language and society. 418

soft palate *see* **palate**

solecism (*gen*) A minor deviation from what is considered to be linguistically correct. 2

sonagram/sonagraph (*phonet*) *see* **spectrograph**

sonant (*phonet*) A †voiced sound. 128

sone (*acou*) Unit of measurement of †loudness. 144

sonorant (*phonet*) A †voiced sound made with a relatively free passage of air ([a], [l], [n]). 159

sonority (*phonet*) The relative prominence or 'carrying power' of a sound. 166

sound change (*hist*) A change in the sound system of a language, over a period of time. 330

sound law (*hist*) A regular, predictable series of †sound changes. 330

sound pressure level (*acou*) The level of a sound as measured in †decibels. 134

sound shift (*hist*) A series of related †sound changes. 330

sound symbolism (*phonet*) A direct association between the sounds of language and the properties of the external world. 176

sound system (*phonol*) The network of †phonetic †contrasts comprising a language's †phonology. 167

sound wave (*acou*) A wave-like air disturbance from a vibrating body, which transmits sound. 132

source language (*ling*) A language from which a word or text is taken. 346

spectrograph (*phonet*) An instrument that gives a visual representation of the acoustic features of speech sounds, in the form of a **spectrogram**; also, **sonagraph/sonagram**. 136

spectrum (*acou*) The range of †frequencies that comprise a †sound wave. 135

speculative grammar (*ling*) A type of grammatical treatise written in the middle ages. 410

speech **1** (*gen*) The oral medium of transmission for language (**spoken language**). 123 **2** (*clin*) The †phonetic †level[1] of communication (where disorder can occur); cf. †language[4]. 267

speech act (*ling*) An †utterance defined in terms of the intentions of the speaker and the effect it has on the listener, e.g. a †directive. 121

speech and language therapist/therapy *see* **speech pathologist/pathology**

speech community (*socio*) A group of people, identified regionally or socially, who share at least one language or †variety. 48

speech defect (*clin*) A regular, involuntary deviation from the norms of speech. 266

speech disorder (*clin*) A serious abnormality in the system underlying the use of spoken language. 266

speech event (*prag*) A specific act or exchange of speech (greeting, sermon, conversation, etc.) 48

speech impairment *see* **speech defect**

speech pathologist (*clin*) A person trained to diagnose, assess, and treat †speech disorders; also, **language pathologist/therapist**, **logop(a)edist**, **orthophonist**, **speech and language therapist**. 266

speech pathology (*clin*) The study of all forms of involuntary, abnormal linguistic behaviour; also, **language pathology**, **logop(a)edics**, **speech and language therapy**. 266

speech perception (*psycho*) The reception and recognition of speech by the brain. 145

speech processing (*psycho*) The stages involved in the perception and production of speech. 264

speech production (*psycho*) The planning and execution of acts of speaking. 264

speech reading (*gen*) A method of interpreting a speaker who cannot be heard by following the movement of the mouth; also, **lip reading**. 227

speech reception *see* **speech recognition**

speech recognition (*psycho*) The initial stage of the †decoding[1] process in †speech perception. 149

speech science(s) (*ling*) The study of all factors involved in †speech production and reception. 123

speech stretcher (*phonet*) A device that presents a slowed but undistorted recording of speech. 138

speech surrogate (*ling*) A communication system that replaces the use of speech (as in drum- or whistle-languages). 404

speech synthesizer (*phonet*) A device that simulates the †speech-production process. 146

spelling pronunciation (*gen*) The pronunciation of a word based on its spelling (*says* as /seɪz/). 182

spelling reform (*gen*) A movement to make spelling more regular in its relation to speech. 217

spirant *see* **fricative**

spirometer (*phys*) An instrument for measuring the air capacity of the lungs. 125

split infinitive (*gram*) The insertion of a word between *to* and the †infinitive form of the †verb in English (*to boldly go*). 2

splitting (*hist*) One †phoneme becoming two as a result of †sound change. 330

spondee (*poet*) A unit of †rhythm in poetic †metre, consisting of two †stressed †syllables. 74

spoonerism (*gen*) The transposition of sounds between words, which gives a new meaning (*queer old dean* for *dear old queen*). 264

spread (*phonet*) Said of sounds made with lips stretched sideways ([i]). 152

stammering *see* **stuttering**

standard (*socio*) A prestige †variety, used as an institutionalized norm in a community; forms or varieties not conforming to this norm are **non-standard** or (pejoratively) **sub-standard**. 24

standardization (*socio*) Making a †form[2] or †usage conform to the †standard language. 366

starred form *see* **asterisked form**

state *see* **stative**

statement (*gram*) A sentence that asserts or reports information (*The dog saw the cat*). 121

static *see* **stative**

statistical linguistics (*ling*) The study of the statistical properties of language(s). 86

stative (*gram*) Said of †verbs that express states of affairs rather than actions (*know*, *seem*); also, **static/state** verbs; cf. †dynamic[1]. 93

steganography (*gen*) The use of techniques to conceal the existence of a message. 58

stem (*gram*) The element in a word to which †affixes are attached; cf. †root[1]. 90

stenography (*graph*) Shorthand writing. 208

stereotyped *see* **formulaic**

stop (*phonet*) A †consonant made by a complete †closure in the †vocal tract ([p], [ɓ]). 159

stratification (*ling*) A model of language as a system of related layers, or **strata**. 83

stress (*phonet*) The degree of force with which a †syllable is uttered; syllables may be **stressed** or **unstressed** in various degrees (heavy, weak, etc.). 171

stressed *see* **stress**

stress-timing *see* **isochrony**

stricture (*phonet*) An ⁺articulation in which the air stream is restricted to some degree. 159

string (*ling*) A linear sequence of linguistic elements. 95

strong form (*phonol*) A ⁺stressed ⁺word form. 171

strong verb (*gram*) A ⁺verb that changes its ⁺root[1] ⁺vowel when changing its ⁺tense (*sing/sang*). 91

structural *see* **structure**

structuralism (*ling*) An approach that analyses language (or any human institution or behaviour) into a set of ⁺structures[1]. 79

structural(ist) linguistics (*ling*) The study of a language's system of ⁺formal[1] patterning (esp. in ⁺grammar and ⁺phonology), rather than of the meaning the patterns convey. 412

structural semantics (*sem*) The study of the ⁺sense relations between words. 105

structural word *see* **function word**

structure (*ling*) **1** A system of interrelated elements, which derive their (**structural**) meaning from the relations that hold between them. 96 **2** A sequential pattern of linguistic elements, at some analytical ⁺level[2,3]; cf. ⁺deep/surface structure. 98

stuttering (*clin*) A disorder of speech ⁺fluency marked by hesitancy, ⁺blocks, sound repetitions, etc.; also, **stammering**. 280

stylistics (*ling*) The study of systematic variation in language use (**style**) characteristic of individuals or groups; also, **stylolinguistics**. 66

stylolinguistics *see* **stylistics**

stylometrics *see* **stylostatistics**

stylostatistics (*ling*) The quantification of ⁺stylistic patterns; also, **stylometrics**. 67

subject (*gram*) The ⁺clause ⁺constituent about which something is stated (in the ⁺predicate) (***The books** are on the table*). 94

subjective *see* **nominative**

subjunctive (*gram*) A grammatical ⁺mood used in some ⁺dependent ⁺clauses to express doubt, tentativeness, etc. (***Were** he here...*); cf. ⁺imperative, ⁺indicative. 93

subordinate *see* **dependent**

subordination (*gram*) The dependence of one grammatical unit upon another, as in **subordinate clauses** (*They left **after the show ended***). 95

subordinator (*gram*) A ⁺conjunction used in ⁺subordination (*since, because*). 95

sub-standard *see* **standard**

substantive (*gram*) A ⁺noun or noun-like item. 91

substantive universal (*ling*) Basic elements that a ⁺grammar[2] requires to analyse language data. 85

substitution (*ling*) The replacement of one element by another at a specific place in a ⁺structure[2]. 119

substitution frame (*gram*) A specific ⁺structure[2] in which a ⁺substitution takes place (*a – cat*). 95

substrate/substratum (*hist, socio*) A ⁺variety that has influenced the structure or use of a more dominant variety or language (the **superstratum**) in a community. 335

suffix (*gram*) An ⁺affix that follows a ⁺stem. 90

superfix (*phonol*) A vocal effect that extends over more than one ⁺segment, e.g. ⁺stress. 171

superlative *see* **degree**

superstratum *see* **substratum**

suppletion (*gram*) The use of an unrelated form to complete a ⁺paradigm (*go/goes/going/gone/**went***). 90

suprasegmental (*phonol*) A vocal effect extending over more than one ⁺segment, e.g. ⁺pitch; also, **plurisegmental**. 171

surface grammar/structure (*ling*) A ⁺syntactic representation of a ⁺sentence that comes closest to how the sentence is actually pronounced. 98

switching *see* **code switching**

syllabary (*graph*) A writing system in which the symbols represent ⁺syllables. 203

syllabic **1** (*phonol*) Said of a ⁺consonant that can be used alone as a syllable (/l/ in *bottle*). 166 **2** (*graph*) Said of a writing system in which the symbols represent ⁺syllables. 203

syllabification (*phonol*) The division of a ⁺word into ⁺syllables. 166

syllable (*phonol*) An element of speech that acts as a unit of ⁺rhythm, consisting of a ⁺vowel, ⁺syllabic[1], or vowel/⁺consonant combination. 164

syllable-timed (*phonol*) Said of languages in which the ⁺syllables occur at regular time intervals; cf. ⁺isochrony. 171

symbol *see* **sign**[1,2]

syn(a)esthesia (*sem*) A direct association between ⁺form[1] and meaning (*sl-* in *slimy, slug*, etc.). 176

synchronic (*ling*) Said of an approach that studies language at a theoretical 'point' in time; contrasts with **diachronic**. 411

syncope (*hist*) The loss of sounds or letters from the middle of a word (*bo'sun*). 330

syncretism (*hist*) The merging of ⁺forms[2] originally distinguished by ⁺inflection[1]. 330

syndeton (*gram*) The use of ⁺conjunctions to link constructions. 95

synonym (*sem*) A word that has the same meaning (in a particular ⁺context[1]) as another word (*a nice **range/selection** of flowers*). 105

syntactic (*gram*) Pertaining to ⁺syntax. 94

syntactics *see* **syntax**

syntagm(a) (*gram*) A string of elements forming a unit in ⁺syntax. 94

syntagmatic (*ling*) Said of the linear relationship between elements in a word or construction. 411

syntax (*gram*) **1** The study of ⁺word combinations; also, **syntactics**; cf. ⁺morphology. **2** The study of ⁺sentence structure (including word structure). 94

synthesis *see* **speech synthesis**

synthetic **1** *see* **inflecting language** **2** *see* **analytic**[2]

systematic phonology (*phonol*) An approach that represents the speaker's knowledge of the ⁺phonological relations between words (*telegraph/telegraphy*, etc.). 162

systemic (*ling*) Said of an approach that analyses language into systems of ⁺contrasts, and studies their functional use in social communication. 411

T (*socio*) Said of a linguistic form (esp. a ⁺pronoun) used to express social closeness or familiarity; cf. ⁺V. 45

taboo (*gen*) Said of a linguistic form whose use is avoided in a society. 61

tachistoscope (*psycho*) A device used in reading research that gives a very brief exposure to a visual image, e.g. a letter. 210

tachygraphy (*graph*) Shorthand writing. 208

tactics (*ling*) The systematic arrangements of linguistic units in linear sequence. 82

tag (*gram*) An element attached to the end of an utterance, esp. a **tag question** (*...isn't it?*). 173

tagmeme (*gram*) A grammatical unit that relates an item's ⁺form[1] and ⁺syntactic ⁺function[1]; the central notion in **tagmemic** analysis. 412

tambre, tamber *see* **timbre**

tap (*phonet*) A ⁺consonant made by a single rapid tongue contact against the roof of the mouth (as sometimes heard in the /t/ of *writer*). 168

target **1** (*phonet*) The theoretical position adopted by the ⁺vocal organs during the ⁺articulation of a sound. 137 **2** (*app*) The language or ⁺variety that is the goal of an activity, e.g. into which a ⁺translation is being made. 346

tautology (*gen*) An unnecessary repetition of a word or idea. 390

taxonomic (*ling*) Said of a linguistic approach that is mainly concerned with classification. 412

technography (*graph*) A writing system devised for a specialized field. 196

teknonymic (*sem*) A parent's name that derives from that of a child. 112

telegrammatic/telegraphic (*psycho*) Said of speech that omits ⁺function words and ⁺dependent ⁺content words (*Man kick ball*). 245

telescoped word *see* **blend**

telestich (*gen*) An ⁺acrostic based on the last letters of words or lines. 64

teletex(t) (*gen*) The transmission of ⁺graphic data from a central source to a television screen. 195

telic (*gram*) Said of a ⁺verb when the activity has a clear terminal point (*kick*); contrasts with **atelic** verbs (*play*). 93

tempo (*phonol*) Relative rate of speech. 171

tense **1** (*gram*) A change in the ⁺form[2] of a ⁺verb to mark the time at which an action takes place (past, present, etc.). 93 **2** *see* **tension**

tension (*phonet*) The muscular force used in making a sound, analysed as strong (**tense**), weak (**lax**), etc. 159

tetrameter/tetrametre (*poet*) A line of verse containing four units of rhythm (⁺foot). 74

text (*ling*) A stretch of spoken or written language with a definable communicative function (news report, poem, road sign, etc.). 116

textlinguistics (*ling*) The study of the linguistic ⁺structure[1] of ⁺texts. 116

textual function (*ling*) The use of language to identify ⁺texts. 119

thematization (*ling*) Moving an element to the front of a sentence, to act as the ⁺theme (*Smith his name is*); also, **topicalization**. 120

theme (*ling*) The element at the beginning of a sentence that expresses what is being talked about (***The cat** was in the garden*); cf. ⁺rheme. 120

theography (*styl*) The study of the language people use to talk about God. 51

thesaurus (*gen*) A book of words grouped on the basis of their meaning. 104

third person *see* **person**

timbre (*phonet*) A sound's tonal quality, or 'colour', which differentiates sounds of the same ⁺pitch, ⁺loudness, and ⁺duration. 133

tip *see* **apex**

tmesis (*rhet*) The insertion of a word or phrase within another (*absobloominglutely*). 70

tone **1** (*phonol*) The distinctive ⁺pitch level of a ⁺syllable. 171 **2** *see* **pure tone**

tone group/unit (*phonol*) A distinctive sequence (or ⁺contour[1]) of ⁺tones[1] in an utterance. 171

tone language (*ling*) A language in which word meanings or ⁺grammatical[1] ⁺contrasts are conveyed by variations in ⁺tone. 174

toneme (*phonol*) A ⁺contrastive ⁺tone[1]. 174

tonetics (*phonet*) The study of the ⁺phonetic properties of ⁺tones[1]. 172

tone unit *see* **tone group**

tonic *see* **nucleus**

tonicity (*phonol*) The placement of ⁺nuclear syllables in an utterance. 173

topic (*ling*) The subject about which something is said (***The pen** is red*); also, **given** information; cf. ⁺comment. 94

topicalization *see* **thematization**

toponomasiology, toponomastics, toponomatology *see* **toponymy**

toponymy (*gen*) The study of place names. 112

trachea (*anat*) The passage between lungs and ⁺larynx. 124

trade language (*socio*) A ⁺pidgin used to facilitate communication while trading. 336

traditional (*gram*) Said of the attitudes and analyses found in language studies that antedate ⁺linguistic science. 3

transcription (*phonet*) A method of writing speech sounds in a systematic and consistent way, from a particular point of view (⁺phonetic/⁺phonemic transcription, ⁺narrow/⁺broad); also, **notation**, **script**. 160

transfer (*app*) The influence of a foreign learner's ⁺mother tongue upon the ⁺target[2] language; **positive transfer** facilitates learning, whereas **negative transfer** (**interference**) hinders it. 374

transform(ation) (*ling*) A formal[1] linguistic operation (a **transformational rule**) that shows a correspondence between two structures, e.g. active and passive ⁺voice[2] sentences. 97

tranformational grammar (*ling*) A ⁺grammar[2] that uses ⁺transformations. 413

transition **1** (*phonol*) The way adjacent sounds are linked (⁺glide, ⁺liaison, etc.). 166 **2** (*acou*) An acoustic change reflecting the movement of the ⁺vocal organs towards or away from a ⁺consonant (esp. ⁺plosive) ⁺articulation. 137 **3** (*socio*) Said of a geographical region (a **transition area**) where there is no clear boundary between adjacent ⁺dialects. 28

transitive (*gram*) Said of a ⁺verb taking a ⁺direct object (*She saw a dog*); cf. ⁺intransitive. 93

translation (*gen*) **1** Conversion from one language into another. **2** Conversion of written texts from one language into another; cf. ⁺interpret. 346

translative (*gram*) An ⁺inflection[1] that typically expresses the meaning of change from one place to another. 93

transliteration (*gen*) Conversion of one writing system into another. 348

tree diagram (*gram*) A diagram used in ⁺generative grammar to show the hierarchical ⁺structure[1] of a ⁺sentence. 96

tremor (*clin*) Involuntary shaking of the voice. 19

trial (*gram*) A grammatical contrast of ⁺number in some languages, referring to 'three of'. 92

trigraph (*graph*) Three written symbols representing one speech sound (*man**oeuvre***). 215

trill (*phonet*) A ⁺consonant made by the rapid tapping of one ⁺vocal organ against another (**trilled** /r/); also, **roll**. 159

trimeter (*poet*) A line of verse containing three units of rhythm (⁺foot). 74

triphthong (*phonet*) A ⁺vowel[1] containing three distinct qualities (*tower* /taʊə/). 156

trisyllable (*phonol*) A word containing three ⁺syllables. 166

trivium (*gen*) The medieval study of grammar, rhetoric, and logic. 410

trochee (*poet*) A unit of rhythm in poetic ⁺metre, consisting of a ⁺stressed followed by an unstressed ⁺syllable. 74

trope (*rhet*) A ⁺figurative effect, e.g. ⁺metaphor, that changes the meaning of language; cf. ⁺scheme. 70

turn (*prag*) A single contribution of a speaker to a conversation (a **conversational turn**). 118

typography (*graph*) The study of the graphic features of the printed page. 192

typological linguistics (*ling*) The study of the structural similarities among languages, regardless of their history. 84

ultimate constituent *see* **constituent**

umlaut (*hist*) A ⁺sound change in which a ⁺vowel[1] is influenced by the vowel in the following ⁺syllable (**gosi* → *geese*). 330

uncial (*graph*) A form of writing consisting of large, rounded letters. 188

uncountable *see* **countable**

underextension (*psycho*) The use of a word to refer to only part of its normal meaning, e.g. a child's use of *shoe* to mean only 'own shoe'. 246

underlying structure *see* **deep grammar/structure**

ungrammatical *see* **grammatical**

unilingual *see* **monolingual**

universal (*ling*) A property found in the analysis of all languages; cf. ⁺formal/⁺substantive universal. 84

universal grammar (*ling*) A ⁺grammar[2] specifying the possible form a language's grammar can take. 84

univocalic (*gen*) A written composition that uses only one ⁺vowel[2]. 65

unmarked *see* **marked**

unproductive (*ling*) Said of a linguistic feature that is no longer used in the creation of new forms (the *-th* of *length*, *width*, etc.). 90

unrounded *see* **rounding**

unstressed *see* **stress**

unvoiced *see* **voiceless**

urban dialectology (*socio*) The study of the speech patterns used within a modern city community. 32

usage (*gen*) The speech and writing habits of a community, esp. when there is a choice between alternative forms (**divided usage**). 2

utterance (*ling*) A physically identifiable stretch of speech lacking any grammatical definition; cf. ⁺sentence. 94

uvula (*anat*) The small lobe hanging from the bottom of the soft ⁺palate. 130

uvular (*phonet*) Said of a ⁺consonant made by the ⁺back of the tongue against the uvula ([ʀ]). 157

V (*socio*) Said of a linguistic ⁺form[2] (esp. a ⁺pronoun) used to express politeness or distance; cf. ⁺T. 45

valency (*gram*) The number and type of bonds that ⁺syntactic elements may form with each other. 412

variable rule (*socio*) A ⁺rule[1] that specifies the ⁺extralinguistic conditions governing the use of a linguistic feature (or variable). 32, 334

variable word (*gram*) A ⁺word that expresses ⁺grammatical[1] relationships by changing its ⁺form[2] (*walk/walks/walking*); cf. ⁺invariable word. 91

variant (*ling*) A linguistic ⁺form[2] that is one of a set of alternatives in a given ⁺context[1] (E. plural /s/, /z/, /ɪz/). 90

variety (*socio*) A situationally distinctive system of linguistic expression (legal, formal, etc.). 48

velar (*phonet*) Said of ⁺consonants made by the ⁺back of the tongue against the soft ⁺palate, or **velum** ([k]). 157

velaric (*phonet*) Said of sounds, e.g. ⁺clicks, when the air has been set in motion by a ⁺closure at the soft ⁺palate. 126

velarization (*phonet*) An ⁺articulation in which the tongue moves towards the soft ⁺palate while another sound is being made. 158

velopharyngeal (*anat*) Said of the area between the soft ⁺palate and the back wall of the ⁺pharynx, which separates oral and nasal ⁺cavities. 130

velum *see* **palate**

ventricular folds (*anat*) Bands of tissue that lie above the ⁺vocal folds. 128

verb (*gram*) A ⁺word class displaying such contrasts as ⁺tense[1], ⁺aspect, ⁺voice[2], ⁺mood, and typically used to express an action, event, or state (*run, know, want*). 91

verbal group *see* **verb phrase**

verb phrase (*gram*) **1** A group of words that have the same grammatical function as a single ⁺verb (*has been running*); also, **verbal group**. 95 **2** In ⁺generative grammar, the whole of a sentence apart from the first ⁺noun phrase. 96

verbless (*gram*) A construction that omits a ⁺verb (***Although angry**, they...*). 95

vernacular (*socio*) The indigenous language or ⁺dialect of a community. 35

viewdata (*gen*) The interactive transmission of data between a central source and a local television set. 195

vocal abuse (*clin*) Overuse of the voice, resulting in a ⁺voice disorder. 278

vocal folds (*phonet*) Two muscular folds in the ⁺larynx that vibrate as a source of sound; also known as **vocal cords/lips/bands**. 128

vocalic (*phonet*) Pertaining to a ⁺vowel[1]. 153

vocalization (*phonet*) Any sound or utterance produced by the ⁺vocal organs. 124

vocal nodules *see* **nodules**

vocal organs (*phonet*) The parts of the body involved in the production of speech sounds. 124

vocal tract (*phonet*) The whole of the air passage above the ⁺larynx. 124

vocative (*gram*) A form (esp. a ⁺noun) used to address a person, animal, etc. (*Excuse me, sir*); in some languages identified by an ⁺inflection[1]. 93

vocoid (*phonet*) A speech sound lacking ⁺closure or audible friction; includes ⁺vowels[1] and vowel-like sounds ([l], [j]). 153

voice **1** (*phonet*) The auditory result of ⁺vocal fold vibration (**voiced** sounds, [b], [z], [e]); cf. ⁺voiceless, ⁺devoiced. 128 **2** (*gram*) A grammatical system varying the relationship between ⁺subject and ⁺object of the ⁺verb, esp. contrasting **active** and **passive** voices (*The cat saw the dog* vs *The dog was seen by the cat*). 93

voiced *see* **voice**[1]

voice disorder (*clin*) An involuntary, abnormal ⁺voice quality[1] that interferes with communication; cf. ⁺dysphonia. 278

voiceless (*phonet*) Said of sounds made without ⁺vocal fold vibration ([f]), [p]); also, **unvoiced**. 152

voice mutation (*phonet*) The development of an adult ⁺voice quality[1] after puberty; also, **breaking**. 19

voice onset time (*phonet*) The point when ⁺vocal fold vibration starts relative to the release of a ⁺closure. 137

voiceprint (*phonet*) A ⁺spectrographic display of the acoustic structure of a person's voice. 20

voice quality (*phonet*) **1** The permanent, background, person-identifying feature of speech. 129 **2** A specific tone of voice. 171

volume *see* **loudness**

vowel (*phonet, phonol*) **1** A sound made without ⁺closure or audible friction, which can function as the centre of a ⁺syllable ([e], [i]). 153 **2** (*graph*) The analogous sign in a writing system. 204

wave *see* **sound wave**

waveform (*acou*) A graph of the movement of air particles in a ⁺sound wave. 132

wavelength (*acou*) The distance travelled by a ⁺sound wave during a single ⁺cycle of vibration. 133

weak form (*phonol*) The ⁺unstressed form of a ⁺word in connected speech (*of* → [ə] in *cup of tea*). 166

weak verb (*gram*) A ⁺verb that forms its past ⁺tense[1] by adding an ⁺inflection[1] (*walk* → *walked*); cf. ⁺strong verb. 90

well formed (*ling*) Said of a sentence that can be ⁺generated by the ⁺rules[1] of a ⁺grammar[2]; cf. ⁺ill formed. 88

Wernicke's area (*anat*) An area of the brain that controls language ⁺comprehension; cf. ⁺Broca's area. 262

whisper (*phonet*) Speech produced without ⁺vocal fold vibration. 128

whistled speech (*ling*) A form of communication in which whistling substitutes for the ⁺tones[1] of normal speech. 404

whole word *see* **look-and-say**

widening (*phonet*) Enlarging the ⁺pharynx to produce a different ⁺vowel[1] quality. 153

word (*gram*) The smallest unit of ⁺grammar that can stand alone as a complete utterance, separated by spaces in written language and potentially by pauses in speech. 91

word blindness *see* **dyslexia**

word class (*gram*) A set of words that display the same ⁺formal[1] properties, esp. their ⁺inflections[1] and ⁺distribution (⁺verb, ⁺noun, etc.); also known as **part of speech**. 91

word ending (*gram*) An ⁺inflection[1] used at the end of a word (*horses, walking*). 90

word-finding problem (*clin*) Inability to retrieve a desired word, symptomatic of ⁺aphasia. 273

word formation (*gram*) The process of creating words out of sequences of ⁺morphemes. 90

word order (*gram*) The sequential arrangement of ⁺words in a language. 98

***x* height** (*graph*) The height of the small letter *x*. 192

zero (*ling*) An abstract unit used in an analysis that has no physical realization in speech. 90

zoösemiotics (*semiot*) The study of the properties of animal communication. 402