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-blooming-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 /piːlit/ 'effort' to produce /pumiːlit/, 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 sub-classes 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
/nosʰ/	my father
/nospʰən/	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:

/piꞏw/	he comes / is coming / came
/piꞏwen/	he is said to be coming / it is said that he came
/piꞏʔ/	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: /ˈnaꞏpeꞏw/ vs /ˈnaꞏpewa/. 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 *didàskomai* '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*).