

277
17-20

*A COURSE IN
MODERN
LINGUISTICS*

CHARLES F. HOCKETT
PROFESSOR OF LINGUISTICS AND ANTHROPOLOGY
CORNELL UNIVERSITY

THE MACMILLAN COMPANY
NEW YORK

ספריה בית קפון
מסונכסיסה העברית

IMMEDIATE CONSTITUENTS

17.1. In §§14–16 we outlined the essential nature of grammar and its relationship to other aspects of language. In this and the following fourteen sections we shall investigate grammatical systems in greater detail.

Specialists have been working for a long time on the problem of analyzing, describing, and comparing grammatical systems, and the degree of accuracy achieved is much greater than the layman would suspect. At the same time, there remain many points on which precision is still impossible. Some linguists like to believe that grammatical analysis has become a completely objective operation, but this is not true. Phonemic analysis has been brought much nearer such a state: complete precision is not always possible, but we can at least pinpoint the areas of indeterminacy and usually see why they remain indeterminate. But grammatical analysis is still, to a surprising extent, an art: the best and clearest descriptions of languages are achieved not by investigators who follow some rigid set of rules, but by those who through some accident of life-history have developed a flair for it.

Consequently, the reader will find in these sections many an example which the writer has handled in one way, but which might also be handled in some other way. The writer has not sought to be ambiguous or arbitrary, but he refuses to speak definitely in cases where he cannot. Indeed, the reader should be alert for possible instances where conciseness of statement has unintentionally concealed uncertainty.

In grammatical study we are concerned with morphemes and their arrangements, but not, save in an ancillary way, with the phonemic shapes which represent morphemes. Consequently, in the present sections we shall usually cite examples in their traditional orthography, provided the language in question has one and that it involves only the

Latin alphabet. Classical Greek and Chinese examples are given in well-established *transliterations* or *Romanizations*. Genuine phonemic notation will be used only when advisable for some special reason, or for languages like Menomini which have no traditional orthography.

17.2. Hierarchical Structure. The man on the street is inclined to identify language with words, and to think that to study words is to study language. This view incorporates two errors. We obviate one when we realize that morphemes, rather than words, are the elementary building-blocks of language in its grammatical aspect, though this shift of emphasis in no sense implies that words are unimportant. The other error is more subtle: the notion, often unstated, that we need only examine words (or morphemes) as isolated units, longer utterances being simply mechanical combinations of the smaller units.

If this were the case, then all we would have to learn in studying a foreign language would be the individual morphemes and their meanings. The meaning of any whole utterance would be immediately obvious on the basis of the meanings of the ultimate constituents. Anyone who has actually studied a foreign language knows that this is not true. For a striking example of the falsity of the assumption, we turn to Chinese, which is better than French or German or Spanish for this purpose because it differs more drastically from English. Here is a commonplace Chinese sentence: *jèige yóutǔng dàgài dzài wǔfēn jūng yǐnèi néng lyóujǐnglè*. Apart from intonation, this sentence includes seventeen successive segmental morphemes, as follows:

- (1) *j-* 'this, proximal, near the speaker';
- (2) *-èi* 'thing or state';
- (3) *-ge* 'discrete concrete object, animate or inanimate';
- (4) *yóu* 'oil, grease';
- (5) *tǔng* 'cylindrical container';
- (6) *dà* 'large, great, greatly';
- (7) *gài* 'generality, majority';
- (8) *dzài* '(be) at, in, on';
- (9) *wǔ* 'five';
- (10) *fēn* 'division, section';
- (11) *jūng* 'clock, hour';
- (12) *yǐ* marker of modification: indicates that something which precedes modifies something that follows;

- (13) *nèi* 'interior, inside';
- (14) *néng* 'can, physical ability';
- (15) *lyóu* 'flow';
- (16) *jìng* 'clean (not necessarily dry), empty';
- (17) *le* marker of completed action or completed change of state.

As is evident, some of these Chinese morphemes have meanings which are not easy to describe precisely in English. One meets similar trouble in trying to describe the meanings of some English morphemes in Chinese—or, in general, the meanings of morphemes in any one language via any other language (§16.2).

A careful scrutiny of the meanings of the seventeen constituent morphemes of the sentence can at best yield some vague notion of what the

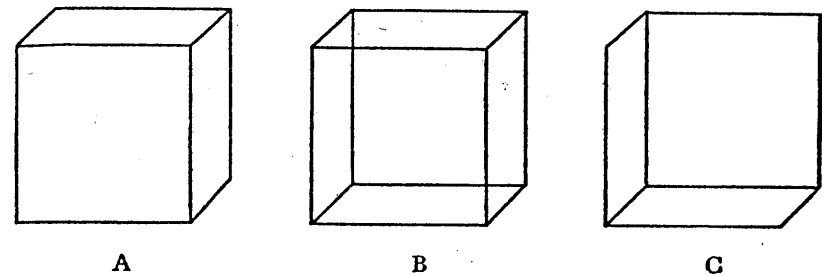


FIGURE 17.1

whole sentence is about. The meaning of the whole sentence happens to be 'This oil drum can be emptied in about five minutes.'

No one—not even a native speaker of Chinese—could know this merely on the basis of the meanings of the ultimate constituent morphemes. Other types of information are also required—types of information which a speaker of Chinese carries around in his head, ready to add to the information carried by what he hears. By virtue of this advance orientation, the native speaker hears the sentence not as a linear string of morphemes, but, as it were, *in depth*, automatically grouping things together in the right way.

An analogy is in order. When we look at the middle line-drawing B of Figure 17.1, we see it either as more like A, to the left, or C, to the right. With a bit of effort, we can make B "jell" in either way. Physically, of course, B is an assemblage of line-segments on a flat surface. The depth that we perceive lies in us, not in the figure. Yet our experi-

ence in visual perception is such that it is hard to see B as a complicated plane figure rather than in three dimensions.

The "depth" which the native speaker of Chinese "reads in" as he hears our Chinese sentence is similar, though with one important difference. All human beings, in all societies, have much the same experiences in visual perception and so would tend to react in the same way to B in Figure 17.1, but the experiences by virtue of which we read "depth" into utterances are specific to the particular language.

Thus the Chinese hearer automatically groups morphemes (6) and (7) together, as depicted in Figure 17.2A. He knows that this particular

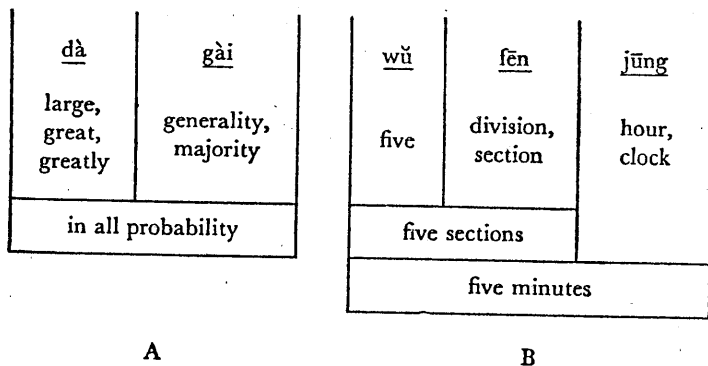


FIGURE 17.2

combination is common, and that it carries the rather special, partly unpredictable, meaning 'probably.' Likewise, he automatically groups (9), (10), and (11) together, as in Figure 17.2B, but in a more complicated way. If we were to ask him what dàgài (morphemes 6 and 7) means, he could tell us, and if we were to ask what wǔfēn jǐng (9, 10, 11) means, he could tell us. But if we were to ask what tǐng dà (5, 6) means, he would be puzzled, for tǐng dà does not mean anything. He would probably be unaware that he had heard this particular morpheme sequence in the sentence, and the speaker of the sentence would scarcely realize that he had said it.

In Figure 17.3 we portray the organization of the whole Chinese sentence as the native speaker perceives it.

Diagrams of the sort shown in Figures 17.2 and 17.3 are designed to show the *hierarchical structure* or *immediate constituent structure* of composite

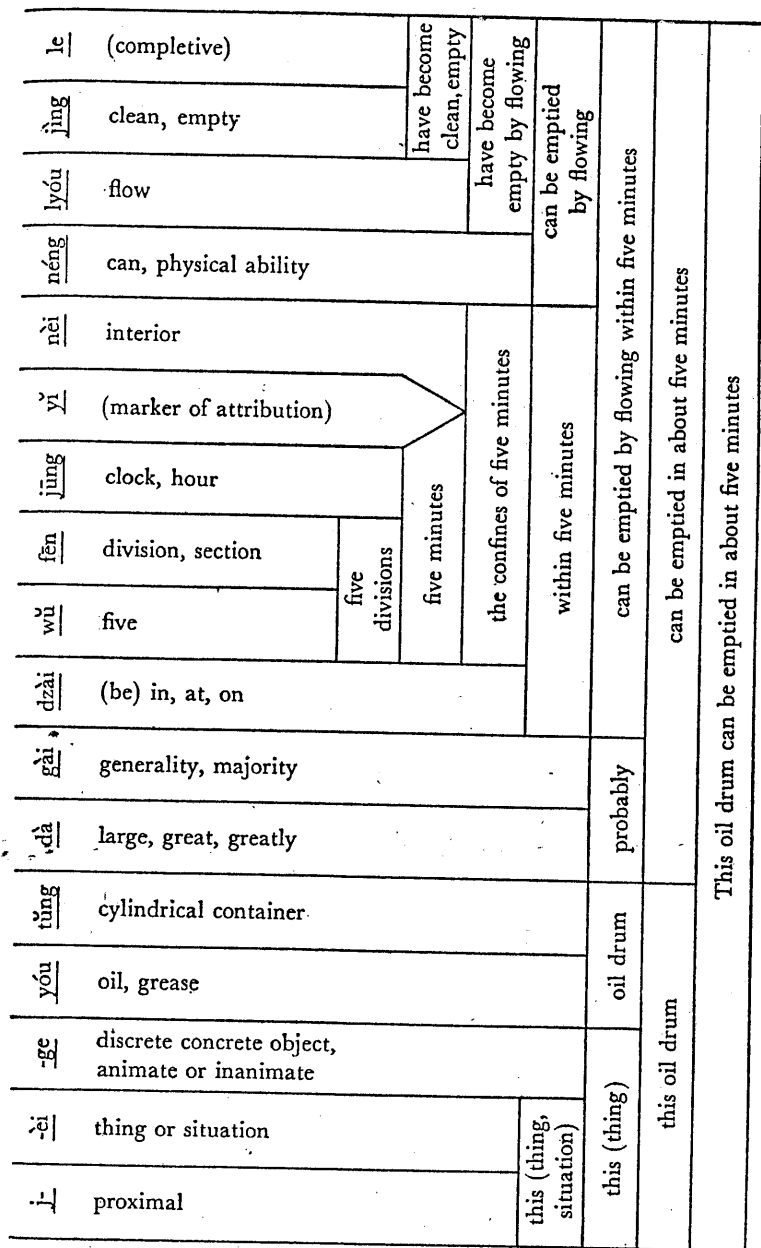


FIGURE 17.3

grammatical forms. Thus the bottom box in Figure 17.2B represents the whole form *wūfēn jūng* 'five minutes'. Working up from the bottom, we see that its *immediate constituents* (for short, *ICs*) are the two smaller forms *wūfēn* 'five sections' and *jūng* 'clock, hour.' The latter is a single morpheme and thus also an *ultimate* constituent of the whole form. The former, however, consists in its turn of the *ICs* *wū* 'five' and *fēn* 'section,' each a single morpheme.

All of the above is applicable also to English or any other language. A meaningless sequence of morphemes like *a man are* can easily be found in normal speech. It occurs in *The sons and daughters of a man are his*

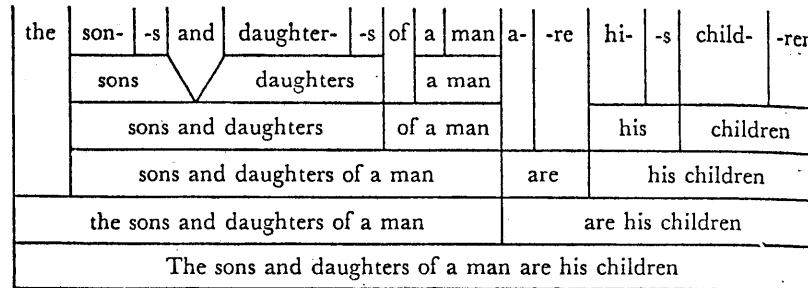


FIGURE 17.4
(Intonation is omitted)

children, diagrammed (omitting intonation) in Figure 17.4. The grammatical forms which occur in this sentence are the morphemes and sequences of morphemes for which boxes are provided: the whole sentence in the lowest box, the two segments *the sons and daughters of a man* and *are his children* in the next to the lowest row of boxes, and so on. Any combination of morphemes, in the sentence for which no box is provided, say *the sons and* or *daughters of*, has the same status as *a man are* or as Chinese *tǔng dà*.

17.3. Ambiguity. It is possible for a single sequence of segmental morphemes to have two alternative hierarchical organizations, usually with a difference of meaning. Sometimes, but not always, the ambiguity is removed by intonation or other context. Ambiguity is not common. In *wild animal house*, for example, the *ICs* are clearly *wild animal* and *house*, rather than *wild* and *animal house*. But in the sentence *He was dancing with the stout major's wife* (with certain of the possible distributions of stress and intonation) we cannot tell whether the man's dancing

partner is stout or not. The ambiguity of its *IC-structure* is shown in Figure 17.5, A and B. Likewise, the expression *old men and women* can have either of two meanings, and either of two corresponding *IC structures*, as shown by Figure 17.6, A and B.

Ambiguity is often eliminated by context: *The stout major's wife is very thin, The stout major's wife has a very thin husband, The old men and women stayed at home while the young men went to war, The old men and women stayed at home while the young folks went dancing.*

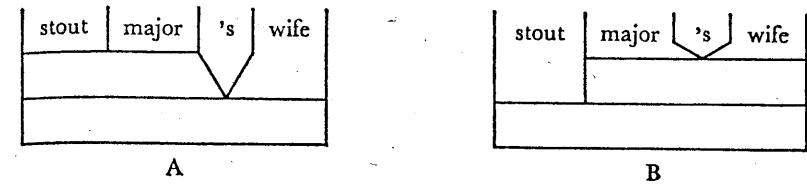


FIGURE 17.5

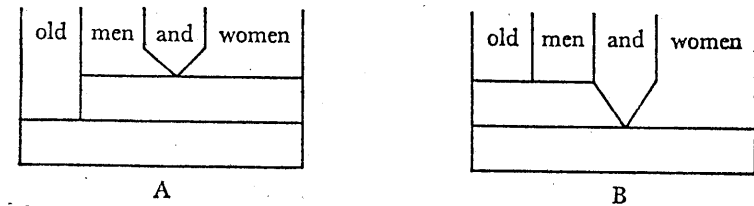


FIGURE 17.6

Such ambiguities remind us again of the analogy with visual perception: Figure 17.1B is ambiguous in that it looks now more like A and now more like C.

17.4. Markers. We must account for the slanting lines appearing in some of the diagrams. In Figure 17.4, for example, the diagram indicates that the *ICs* of *sons and daughters* are the two words *sons* and *daughters*. How about the *and*? How can a form participate in a larger form without being a constituent of it?

Of course, a different interpretation would be possible, but the one we have chosen indicates that *and*, rather than being one of the *ICs* of *sons and daughters*, is what we may call a *structural marker* or *signal*. Some morphemes, that is, serve not directly as carriers of meaning, but only as markers of the structural relationships between other forms. *And* marks the fact that something before it (here *sons*) and something after

it (here *daughters*) are the ICs of a larger grammatical form, and *and* also marks that larger form as being of a certain type. We would choose a similar interpretation for the *or* of *sons* or *daughters*.

17.5. Multiple ICs. In all our diagrams so far, composite forms have been shown as consisting of just two ICs. Bipartite composite forms are extremely common, but there is no universal restriction to two ICs. English has a few cases of composite forms with three ICs; for example, *foot-pound-second* or *centimeter-gram-second*. Figure 17.7 shows the way of diagramming them.

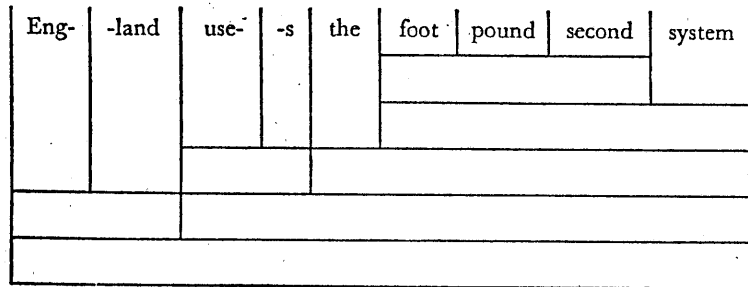
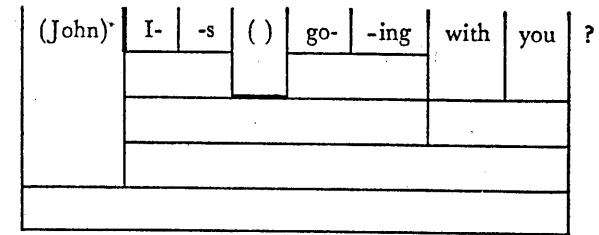


FIGURE 17.7

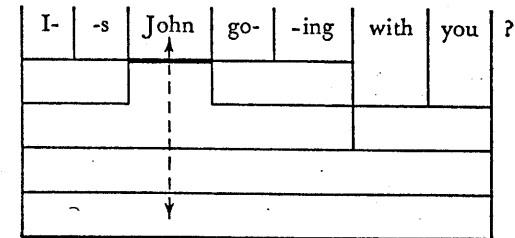
17.6. Discontinuous ICs. Our examples so far have had another property which is common but not universal: forms which belong together as ICs of a larger form have been next to each other in linear sequence. But *discontinuous* constituents are not at all uncommon. For example, in the English sentence *Is John going with you?*, setting intonation aside, one IC is *John* and the other is the discontinuous sequence *Is . . . going with you*.

Figures 17.8A and B show two graphic devices for handling this. In Figure 17.8A, the form *John* is entered at the beginning to render diagramming easy, but is parenthesized to indicate that it is not actually spoken there; the empty parentheses after *is* indicate the position it actually occupies in the sequence. In Figure 17.8B we avoid the duplication, but place a heavy line below the entry *John*, and mark with a dotted arrow the connection between *John* and the larger form of which it is one IC.

17.7. Simultaneous ICs. An intonation morpheme is probably always to be interpreted as one IC of the macrosegment which includes



A



B

FIGURE 17.8

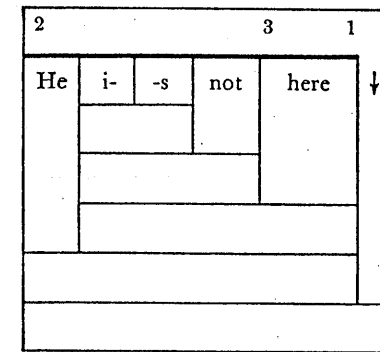


FIGURE 17.9

it, the remainder of the macrosegment, no matter how complex, constituting the other IC. In order to show this diagrammatically we have to introduce another special device, illustrated in Figure 17.9. It is necessary to mark the positions of the PLs and TC correctly, since any alternation in their position might yield a different sentence (e.g., ²*He is* ³*not here*¹↓).

Diagramming is not an end in itself, but a convenient means of revealing hierarchical structure. For this, it is useful to have diagrammatic conventions. But where the structure is unusual, diagramming may become excessively complex. In such instances, we shall avoid diagrams and resort to verbal description.

NOTES

New terms: *hierarchical structure, immediate constituents (= ICs), discontinuous ICs, simultaneous ICs, structural markers.*

For ICs we follow Bloomfield 1933, p. 161, as elaborated by Wells 1947, and with modifications partly expounded in Hockett 1954.

18.

FORM CLASSES and CONSTRUCTIONS

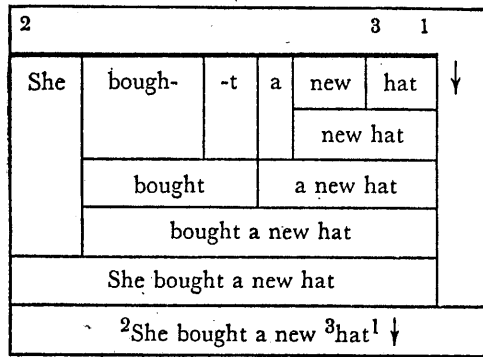
18.1. Recurrent Patterns. The property of language which renders it such a powerful means of communication is that one can say something that has never been said before, and yet be perfectly understood, often without either speaker or audience being aware of the novelty. A novel utterance is built from familiar raw-materials, by familiar patterns of putting raw-materials together. Neither the raw-materials nor the patterns need be new in order for the utterance to be different from any that has occurred previously.

We know from earlier discussion that the raw-materials are morphemes, and that the patterns are hierarchical rather than mere linear juxtaposition. Here we shall pursue the matter of patterns further.

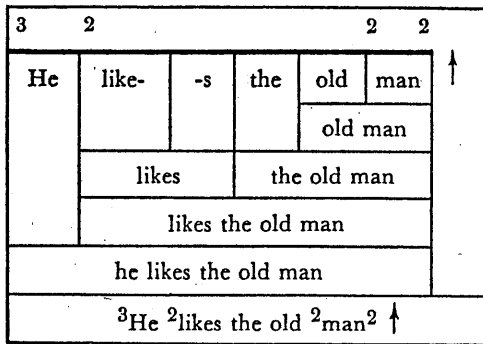
Consider the two English sentences ²*She bought a new* ³*hat*¹↓ and ³*He* ²*likes the old* ²*man*²↑ diagrammed in Figure 18.1. The two are completely distinct in their constituent grammatical forms: no morpheme and no composite form of either occurs in the other. Yet they consist of the same number of constituents, in exactly the same hierarchical arrangement. Furthermore, there is a similarity in meaning beyond that of two sentences selected at random: each sentence asserts something about someone, and each assertion involves some second entity.

We summarize the similarities by saying that the two sentences are of the same pattern, and that the common pattern is responsible for the similarity of meaning. We can portray the common pattern in part without citing any forms, as shown in Figure 18.2. This "empty box" diagram is obtained from either diagram in Figure 18.1 simply by deleting all the entries.

On the other hand, two sentences may involve exactly the same constituents at all hierarchical levels, and yet differ in meaning because of different patterns. Figure 18.3 gives one example. The difference lies not in constituents, but in their arrangement: *John* respectively before



A



B

FIGURE 18.1

or within *is here*. For another example, compare the diagram of ²*She bought a new* ³*hat*¹↓ in Figure 18.1 with that of ²*She bought a* ³*new hat*¹↓ in Figure 18.4. The center of the intonation is differently placed, though the intonation, and all other constituents, are identical. (These examples are not like *old men and women*, discussed in §17.3, for in that example only the ultimate constituents were the same, the hierarchical structure being ambiguous.)

Our technique of diagramming cannot always be counted on to reveal the difference between obviously distinct patterns. Thus compare *She likes fresh milk* and *She likes milk fresh*, partially diagrammed in Figure 18.5. Insofar as the diagrams reveal the pattern, it can be shown as in Figure 18.6, but the two composite forms *fresh milk* and *milk fresh* are

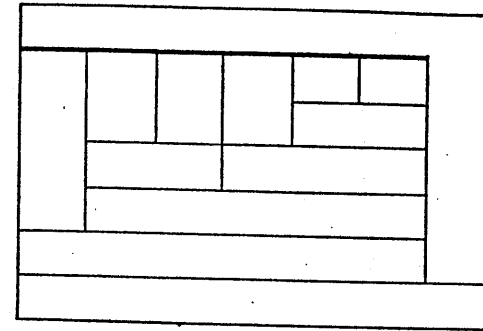
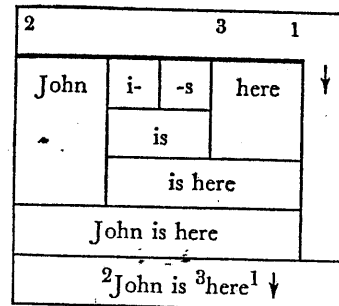
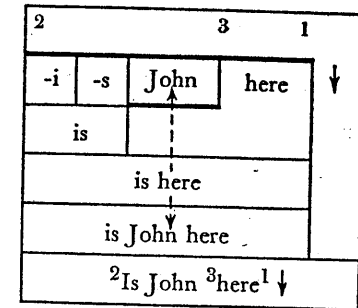


FIGURE 18.2



A



B

FIGURE 18.3

actually quite different. The former can occur in many contexts (*Fresh milk is good for you, Make it with fresh milk*); the latter is quite limited.

Suppose, however, that we had some way to distinguish between patterns that yielded the same empty-box diagram, and that we were to diagram hundreds of English sentences, delete the entries, and list all the different resulting empty-box diagrams. This would obviously give us a very large number of different whole-sentence patterns. But we

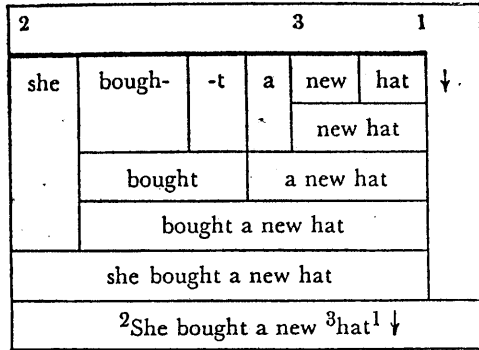


FIGURE 18.4

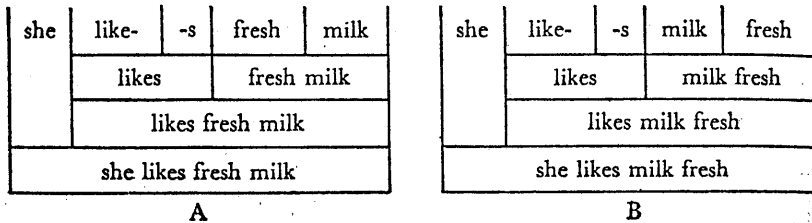


FIGURE 18.5

cannot simply assume that all these patterns are completely distinct. We must investigate a simpler possibility: that the numerous whole-sentence patterns are built up out of a smaller number of simpler ones.

In Figure 18.7 are diagrams of four progressively larger sentences, labelled A through D. The boxes in each are numbered in the lower right-hand corner.

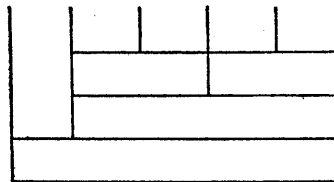


FIGURE 18.6

Sentence A consists of only two ultimate constituents (morphemes), which are therefore also the ICs of the whole sentence: 3 and 2 are the ICs of 1.

Sentence B consists of more than two ultimate constituents, but, once again, of only two immediate constituents: 3 and 2 as in A, are the ICs of 1.

Similar remarks apply to sentences C and D.

Furthermore, the relationship between the two ICs of each whole

sentence is the same. Thus, if we make just one IC-cut in each sentence, ignoring any smaller constituents for the moment, then all four sentences conform to pattern X of Figure 18.8. Box 3 in pattern X can be

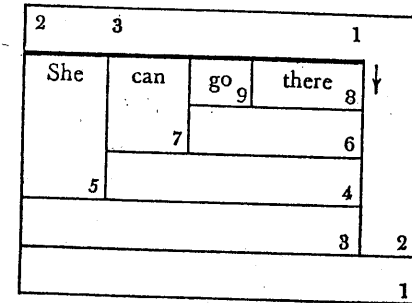
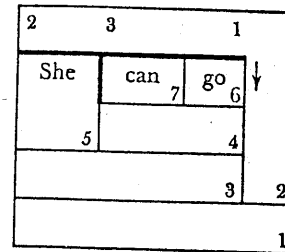
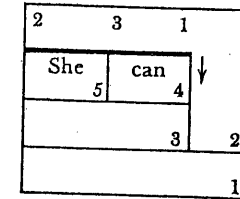
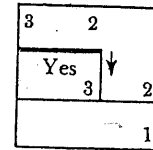


FIGURE 18.7

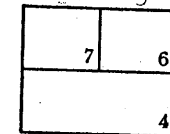
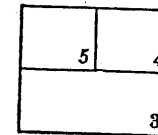
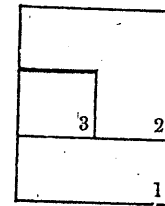


FIGURE 18.8

filled with any of the four intonationless morpheme-sequences *yes, she can, she can go, or she can go there*; Box 2 can be filled with the intonation morpheme /⁽²⁾ 31↓/. *q' b'wen - le no/ma/ni/ni/*
si/zi/ni/ni/ni/

Setting sentence A aside, let us examine boxes 3, 4, and 5 in sentences B, C, and D. In B, the form of box 3 has two ultimate constituents,

which are therefore also its ICs: those in boxes 5 and 4. Now if, in sentences C and D, we cut the form of box 3 only into its *immediate* constituents, ignoring any finer-grained structure, we find the same thing again: the ICs of the form in box 3 of any of the sentences B, C, and D are the forms in boxes 5 and 4. This portion of B, C, and D thus conforms to the pattern labelled Y in Figure 18.8. Box 5 of pattern Y can be filled with the morpheme *she*, and box 4 with any of the morpheme-sequences *can*, *can go*, or *can go there*.

Similarly, setting both A and B aside, we see that the boxes labelled 4 in C and D, providing we make only one IC-cut, are identical and conform to pattern Z in Figure 18.8. Box 7 of pattern Z can be filled with the morpheme *can*, and box 6 with either *go* or *go there*.

Our treatment shows that some composite forms and some single morphemes are alike in that they can participate similarly in larger forms. For example, the single morpheme *yes* and the three composite forms *she can*, *she can go*, and *she can go there*, despite internal differences, share at least one *privilege of occurrence*: each can be conjoined to the intonation /⁽²⁾ 31↓/ to produce an utterance. The morpheme *can* and the composite forms *can go* and *can go there* share at least one privilege of occurrence: each can be used with preceding *she* to build a larger form of a certain kind. Similarly for the morpheme *go* and the composite form *go there*.

18.2. Form-Classes. A class of forms which have similar privileges of occurrence in building larger forms is a *form-class*.

Thus, by virtue of their interchangeability in box 3 of pattern X (Figure 18.8), the forms *yes*, *she can*, *she can go*, and *she can go there*, together with untold thousands of other forms, belong to a single form-class.

Similarly, ability to occur in box 2 of pattern X puts the intonation morpheme /⁽²⁾ 31↓/ in a form-class together with many other intonation morphemes.

Box 5 of pattern Y defines a form-class which includes *she*, *he*, *it*, *John*, *Mary*, *the man on the corner*, *my friend Bill*, and so on endlessly, but which by no means includes all forms, since we can name many which are excluded: *her*, *him*, *them*, *me*, *yes*, *no*, *ripe*, *find her*, *go with us tomorrow*.

Box 4 of pattern Y defines a form-class which includes *can*, *will*, *can go*, *can go there*, *should like potatoes*, *must run faster than that*, but which excludes *quickly*, *yes*, *her*, *him*, *ripe*, *she*, *he*, *John*, *Mary*, and so on.

Every other box in our diagrams similarly defines a form-class. There are not as many form-classes in a language as there are boxes in all the IC-diagrams one could draw, since different boxes often define the same form-class.

One point should be noted about the way we have chosen examples for the classes defined by boxes 4 and 5 of pattern Y. To the first (*she*, *he*, *it*, etc.) one might want to add *I*, *we*, *they*, *the men across the street*, all of which can occur before *can*, *can go*, *can go there*. To the second (*can*, *will*, *can go*, etc.) one might want to add *likes potatoes*, *finds it dull there*, *is trying too hard*, and others which can occur after *she*. The exclusion of these forms was intentional. Form-classes are so constituted that, if *some* member of a given class can occur with *some* member of a second class, then *any* member of the first might occur with *any* member of the second. Had we extended the lists of examples as suggested above, then this would not hold, for no one, speaking standard English, says *I likes potatoes*, *we finds it dull there*, *they is trying too hard*.

18.3. Constructions. If we were to delete the box-numbers from diagrams Y and Z (Figure 18.8), the two empty-box diagrams would be identical. This would conceal an important difference. As defined, Y subsumes composite forms like *she | can*,¹ *she | can go*, *the man on the corner | should like potatoes*, while Z subsumes composite forms like *can | go*, *can | go there*, *should | like potatoes*.

The difference can be underscored by regarding the numbers in the boxes in diagrams X, Y, and Z as names for form-classes. Thus, in Y, "5" means "any member of the form-class which includes *she*, *he*, *it*, *John*, *Mary*, etc." But numbers are mnemonically poor labels, and descriptive terms of some sort would be better. Let us replace "5" by *third person singular subjects*, "4" by *modal predicates*, and "3" by *predications*. Diagram

third person singular subject	modal predicate
predication	

FIGURE 18.9

Y then takes the form shown in Figure 18.9, and can be read off as follows: "any third person singular subject, followed by any modal predicate, builds a predication."

Either a statement of this sort, or a diagram of just two hierarchical

¹ Here and in the following sections, the vertical slash "|" is not a phonemic symbol, but an indication of the placement of the cut of the form into ICs.

levels in which the boxes contain labels for form-classes instead of for specific forms, describes a *construction*.

A construction is thus a pattern for building composite forms of a specific form-class out of ICs of specific form-classes. The description of a specific construction asserts that "any member of such-and-such a form-class, conjoined to any member of a certain other form-class, produces a form which belongs to a certain third form-class." We see immediately why care must be exercised in discriminating between form-classes. If we put *I* along with *she*, *he*, *it*, or *likes potatoes* along with *can*, *can go*, *should like potatoes*, our descriptions will have to be complicated by specification of exceptions.

The ICs of a composite form are commonly said to *stand in* a certain construction with each other, and the composite form built from the ICs by the construction is also called a *constitute*.

All constitutes built by a single construction are necessarily members of the same form-class. The form-class, however, may include also forms built by some other construction, and even single morphemes. For example, all predications are members of a larger form-class which we may momentarily call "sentence-skeletons": forms to which one can add an intonation to yield a sentence. But not all sentence-skeletons are predications: *yes*, *why*, *the more the merrier*, *milk for health* are the former but not the latter. Again: predications are built by conjoining a third person singular subject and a modal predicate (one construction), but are also built by several other constructions, as in *she | likes potatoes*, *I | can go*, *I | like potatoes*.

Two composite forms built by a single construction may have neither IC in common: *she | can go* and *he | should like potatoes*. The common feature of meaning of such a pair is not due to any shared grammatical form; we therefore say that it is the *meaning of the construction*.

Returning to the examples and problems discussed in §18.1, we now see that constructions are the sort of smaller pattern out of which the patterns of whole sentences are built. The two sentences diagrammed in Figure 18.1 resemble each other in that the same nesting of constructions is involved. Thus *new | hat* and *old | man* are built by a single construction. *A | new hat* and *the | old man* are perhaps both built by a second, though possibly we must say instead that they are built by two highly similar constructions. The features of meaning common to the two sentences are the meanings of the recurrent constructions.

Similarly, differences of meaning in pairs of sentences which contain the same ultimate constituents in the same hierarchical arrangement are due to differences in constructions. In *John is here*, the ICs *John* and *is here* stand in one construction; in *Is John here* the same ICs stand in a different construction (Figure 18.3). In *(she likes) fresh milk* the ICs *fresh* and *milk* stand in one construction (that of *new | hat*, *old | man*); in *(she likes) milk fresh* the same ICs stand in another construction (Figure 18.5).

NOTES

New terms: *privilege of occurrence*, *form-class*, *construction*, *meaning* of a construction, *constitute* (= composite form, but only when being discussed as the product of ICs joined by a specific construction); to *stand in* a construction.

The terms "third person singular subject," "modal predicate," and "predication" have technical status in the description of a particular language only when formally introduced in the treatment of that language. They are useful in English; there is no general guarantee that they are of value for any other language.

Some grammarians use "construct" in place of our "constitute." Some use "construction" ambiguously for our "constitute" and "construction."

WORDS

19.1. The everyday use of the English word "word" is not very precise. In general, the layman looks to writing, and classes as a word whatever he finds written between successive spaces. So *matchbox* is one word, *match box* two, and *match-box* two or one depending on whether or not a hyphen is interpreted as a special sort of space. That these three spellings reflect a single combination of morphemes with a single pronunciation—/mætʃ + bɒks/—is ignored.

When we look at language directly rather than via writing, we must seek other criteria for the determination of words. There are several usable criteria, but they do not yield identical results. The criterion that is easiest to apply yields units most like the "words" of the layman, and it is for these that we shall reserve the term. The other criteria yield stocks of units which differ more radically from the layman's "words," and we shall not call them words, despite their wordlike properties. Instead, we shall introduce special terms for them.

19.2. **Determining Words through Pause and Isolability.** As the first step in determining the words in an utterance, we ask speakers to repeat the utterance slowly and carefully. Suppose someone has just said *John treats his older sisters very nicely* in the normal rapid way, as a single macrosegment (§4). If we ask for a slow repetition, he may break the sentence up into as many as seven successive macrosegments, each with its own intonation and with intervening pauses: *John, treats, his, older, sisters, very, nicely*. Or he may not pause quite so often: *his older, or very nicely*, might be kept as a single macrosegment. Thus we may have to elicit more than one slow careful delivery before we can be sure we have obtained the maximum break-up. Only under very artificial con-

ditions, however, would anyone pause at additional points, say between *old* and *-er*.

A word is thus any segment of a sentence bounded by successive points at which pausing is possible. The example contains seven words. It contains this number whether actually delivered as one macrosegment or as several, since words are defined in terms of potential pauses, not the actual pauses in any one delivery.

The pausing habits of a literate speaker of English are doubtless conditioned by his literacy, for he may pause more freely where writing habits leave a space than where they do not. But there are exceptions: many a speaker will break *matchbox* or *blackbird* into two macrosegments. Furthermore, our habits of leaving spaces in writing have not developed by mere chance. They reflect, with some distortion, speech habits, including habits of where one pauses in slow, careful speech. And the pause procedure for determining words also yields consistent results when applied to a language for which there is no commonly used writing system.

When we suspect that some factor is obscuring the results of the pause procedure, there is a supplementary procedure to which we can turn. We look for other contexts for a form, in which it will indeed occur as a whole macrosegment. Thus someone might consistently fail to break up the sentence *I'm going outside* into more than four macrosegments, *I, am, going, outside*. Yet we might hear the same person playing role *B* in the following conversation (a conversation the writer has actually heard):

A: *Where're you going?*

B: *Out.*

A: *Out where?*

B: *Side.*

From it we conclude that *outside* is two words, not one.

Words thus defined are not always identical with the layman's words. *Outside* and *don't* are single words for the layman because of the spelling, but two each for us. Yet our words have many of the properties of the layman's. They are, on the average, larger than a single morpheme: *boy, girl, man* one morpheme each, but *boyish, manly* two each, *muddily* three, and so on. At the same time, they are on the average smaller than

utterances: one-word utterances like *Hi!* and *Yes?* are normal, but in the minority. We sometimes unthinkingly assume that a sentence is composed *solely* of words. This is false under either the lay definition of "word" or our own, since it leaves intonation out of account. Thus our initial example involves the intonation morpheme /³ ² ²²↑/ as well as its seven constituent words.

All morphemes can be classed according to their status relative to words. A morpheme which occurs only as part or all of some word is a *segmental* morpheme: *John*, *treat*, *-s*. A morpheme which is not part of a word is *suprasegmental*: all intonations, and some other morphemes, such as the secondary stress /' / on the second segmental constituent of /mæč + baks/, which is not part of the word *box*.

A form consisting of two or more words is a *phrase*.

19.3. Minimum Free Forms. Another lay expectation about words is that they are invariably grammatical units, or, in our terms, grammatical forms. This is not true under our definition. *Twenty* and *eighth* are words, so that *twenty-eighth* is two. But the ICs of *twenty-eighth* are not the words *twenty* and *eighth*; they are the form *twenty-eight*, itself two words, and the form *-th*, less than a word.

There are two types of wordlike unit which fulfil lay expectations on the present count, at the expense of ease of determination and of some of the other properties one should like wordlike units to have. One of these is the *minimum free form*.

Some forms of one or more segmental morphemes, like English *act*, *John*, *hat*, *actor*, *actors*, *John's*, *John's hat*, *John's hat is on the table*, have the property that on occasion they may occur as whole utterances, requiring only the addition of a suitable intonation. Thus in answer to *Whose hat is that?* one might simply say /³janz¹↓/. This property is *freedom*; forms that have it are *free*. Other segmental forms are not free, and are therefore called *bound*: the *-or* of *actor*, the *-dom* of *kingdom*, the *-s* of *sisters*, the *-ation* of *condemnation*.

Some free forms consist of ICs which are all also free. *John's hat* is an example, since both *John's* and *hat* are free. Many, however, do not consist wholly of free ICs; they are therefore *minimum free forms*. *Actor*, *John's*, *regain* qualify: in each, one IC (*act*, *John*, *gain*) is free, but the other (*-or*, *-s*, *re-*) is bound. *Confer* also qualifies: if it is more than one morpheme (investigators disagree) then its ICs are *con-* and *-fer*, both bound. *Act* is a minimum free form because it has no ICs.

The wordlike properties of minimum free forms are clear. They are always grammatical forms, because we pay strict attention to IC structure in finding and defining them. They are on the average larger than morphemes and smaller than whole utterances. Their unwordlike property is that sometimes they are larger than single words. *Twenty-eighth* is two words, but only one minimum free form, since one of its ICs (*-th*) is bound. Even the four-word sequence *the Mayor of Boston's* (as in *the Mayor of Boston's hat*) is a single minimum free form, since the bound form *-s* is one of its ICs.

19.4. Lexemes. The other variety of wordlike unit which fulfils the expectation that "words" should always be grammatical forms is the *lexeme*.

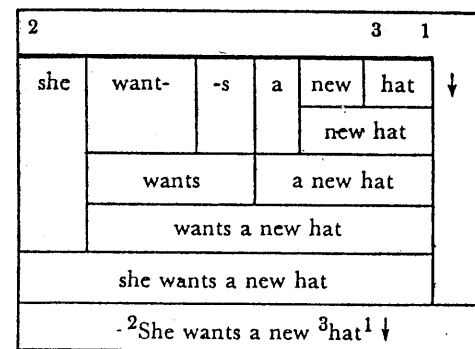


FIGURE 19.1

The sentence ²*She wants a new* ³*hat*¹↓ is diagrammed in Figure 19.1. The sentence includes only seven ultimate constituents (morphemes), but there are thirteen different grammatical forms in all, at various hierarchical levels. The diagram includes a box for each.

The morpheme-sequence *wants a new* occurs in this sentence, but only as an "accidental" concatenation (like *a man are* in *The sons and daughters of a man are his children*, §17.2). The morpheme-sequence *a new hat*, on the other hand, occurs in the sentence as a grammatical form. However, we can easily find some other sentence in which the same sequence, *a new hat*, occurs only accidentally: *She wants a new hatrack* will do (Figure 19.2).

Another morpheme-sequence in the sentence is *wants*. *Wants* is a grammatical form in this sentence just as *is a new hat*. But *wants* is differ-

ent from a *new hat* in that we can find no English utterance in which *wants* occurs only as an accident. It is a grammatical form wherever it occurs. This property is manifested by every single morpheme of a language, but only by some of the occurrent sequences of morphemes.

This yields a threefold classification of all the morphemes and morpheme-combinations of a language: (1) those which occur only as accidents: *a man are, wants a new*; (2) those which occur sometimes as accidents, sometimes as grammatical forms: *a new hat*; (3) those that occur only as grammatical forms: *wants, hat*.

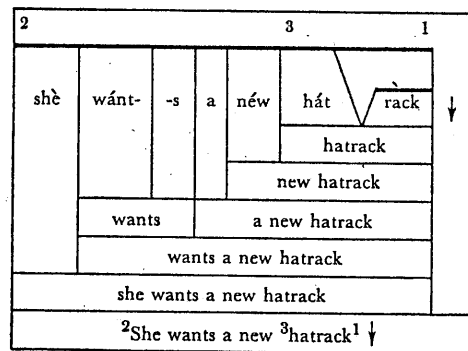


FIGURE 19.2

Let us momentarily call any morpheme or morpheme-sequence of this third variety an X. Although an X is by definition always a grammatical form, in some environments it is an IC of a larger X, whereas in other environments it is not. For example, *want* is an X; in *She wants a new hat*, *want* is an IC of *wants*, which is also an X; but in *I want a new hat* or *They don't want any help*, *want* is an IC of larger forms which are not X's.

Any X, in any context in which it is *not* an IC of a larger X, is a *lexeme*. *Want* is a lexeme in *I want a new hat*, but not in *She wants a new hat*. The constituent lexemes of the latter sentence are *she, wants, a, new, hat*, and the intonation morpheme /² ³1 ↓/.

Any grammatical form larger than a lexeme is a *nonce-form*. The term suggests, not inappropriately, that a speaker coins such larger grammatical forms when he needs them; if on another occasion the same larger grammatical form is again needed, he coins it afresh. The nonce-

forms in our sample sentence (Figure 19.1) are *new hat, a new hat, wants a new hat*, the intonationless *she wants a new hat*, and the whole sentence.

The wordlike properties of lexemes are clear. A lexeme is always a grammatical form, by definition. Lexemes are on the average larger than single morphemes but smaller than whole utterances. Many words are lexemes in many occurrences, and many lexemes are words. A minor deviation is that intonations count as lexemes: when speaking of words, or of minimum free forms, intonations have to be taken into account separately.

But lexemes are in many cases smaller than what we usually call words, even by the formal definition of §19.2. The lexemes in the two-word sequence *twenty-eighth* are *twenty, eight*, and *-th*. Those in *the Mayor of Boston's* are *the, Mayor, of, Boston*, and *'s*. *Red-haired* is two words; but it is a single minimum free form, since the ICs are *red hair* (free) and *-ed* (bound); and it is three lexemes, *red, hair*, and *-ed*.

19.5. Idioms. A final lay assumption about "words," which does not actually hold either for the layman's words or for ours, is that they should always have some sort of meaning of their own, predictable in terms of their structure if they are larger than morphemes, and reasonably constant from one occurrence to another.

Minimum free forms and lexemes also do not meet this requirement. The units which do are the least wordlike of any of the types we shall discuss. The best approach to these units, which we shall call *idioms*, is via examples in some other language.

The Chinese form *yóutǔng* has as ICs the two morphemes *yóu* 'oil, grease' and *tǔng* 'large cylindrical container'; the first IC modifies the second, as *black* modifies *cat* in *black cat* or *grease* modifies *rack* in *grease rack*. Given this information, but knowing nothing else about Chinese or the culture of China, we can venture a reasonable guess as to the meaning of *yóutǔng*: 'oil container,' 'oil drum,' or the like. This guess is correct.

The ICs of Chinese *mǎshàng* are *mǎ* 'horse' and *shàng* 'space on or above, top, ascend.' As in the preceding example, the first constituent modifies the second. Reasonable guesses at the meaning of *mǎshàng* would be 'horse's back,' 'horseback,' or possibly 'on horseback.' These are wrong. The meaning is 'quickly, right away.'

This meaning is not surprising when we remember that until recently the most rapid mode of travel was by horse. But it is one thing to con-

sider a meaning reasonable after we know it, and quite different to deduce the meaning of a form from its structure. A native speaker of Chinese is no better off than we, for he can know *mǎ* and *shàng* and still not understand *mǎshàng* unless he has learned the meaning of the latter as a separate fact about his language.

Let us momentarily use the term "Y" for any grammatical form the meaning of which is not deducible from its structure. Any Y, in any occurrence in which it is not a constituent of a larger Y, is an *idiom*. A vast number of composite forms in any language are idioms. If we are to be consistent in our use of the definition, we are forced also to grant every morpheme idiomatic status, save when it is occurring as a constituent of a larger idiom, since a morpheme has no structure from which its meaning could be deduced.

Thus *new* is an idiom in *She wants a new hat*, but not in *I'm going to New York*, because here it is part of the larger idiom *New York*. *New York*, in turn, is an idiom in the preceding sentence but not in *The New York Times* or *The New Yorker*, since in the latter expressions *New York* occurs as part of larger idioms. The advantage of this feature of our definition, and of the inclusion of morphemes as idioms when they are not parts of larger idioms, is that we can now assert that any utterance consists wholly of an integral number of idioms. Any composite form which is not itself idiomatic consists of smaller forms which are.

A composite form in another language cannot be called an idiom merely because its meaning seems queer to us. The test must be applied within the language. French *Elle est garde-malade* 'She is a nurse' may seem peculiar to us because it contains no equivalent for English *a*, but this is the regular habit in French, and the sentence is no idiom. On the other hand, though French *mariage de convenance* finds its exact counterpart in English *marriage of convenience*, both the French and the English phrases are idioms.

An idiomatic composite form may coincide in morphemic shape with a form that is not idiomatic. *White paper* is an idiom when it refers to a certain sort of governmental document, but not when it refers merely to paper that is white.

A single form can be two or more idioms. *Statue of Liberty* is one idiom as the designation of an object in New York Bay; it is another in its reference to a certain play in football. *Bear* is presumably the same

morpheme in *women bear children* and in *I can't bear the pain*, but it is different idioms in these two environments.

Idioms are unwordlike especially in that they can be much larger than single words: *Now is the time for all good men to come to the aid of the party*. Yet some idioms are smaller than words. *Bought, went, paid, sold, sang, rang* consist of two morphemes each. One is, respectively, *buy, go, pay, sell, sing, ring*; the other, in all of them, is the "past tense" morpheme. In most occurrences, however, the meanings of the whole forms are predictable from the meanings of the constituents, so that the whole words are not idioms.

In theory, and largely in practice, idioms are the stuff of which dictionaries are made. The reason is obvious: a dictionary-maker need not include a non-idiomatic nonce-form, since a speaker of the language would never look up such a form. He would look up the component parts, if he needed to, and automatically know the meaning of the whole. In practice, of course, no dictionary is ever complete. There are far too many idioms in any language, and more come into existence every day.

19.6. Idioms and Morphemes. The recognition of idioms larger than single morphemes requires a modification of what has heretofore been said about morphemes as the raw-materials from which we build utterances. An idiomatic composite form, like any single morpheme, has to be learned as a whole. Thus it is equally legitimate to say that the raw-materials from which we build utterances are idioms.

Furthermore, we can often be sure that a small form is an idiom, even when it is difficult to decide whether it is one morpheme or more than one. For example, English has many words of the type *remôte, demôte, promôte, réduce, deduce, produce*, each apparently built of two smaller parts, a prefix *re-, de-, pro-*, or the like, and a second part *-môte, -duce*, or the like. But the relationships of meaning are tenuous. Grammarians are not in agreement. Some brush aside the semantic difficulties and take each word as two morphemes, following the phonemic shapes; others regard the parallelisms of phonemic shape as unconvincing and take each word as a single morpheme. Similar problems appear in the analysis of almost every language. An obvious practical step is to set the morphemic problem aside, recognizing that each form is an idiom whether it is one or more morphemes.

NOTES

New terms are the following: *word, segmental and suprasegmental morphemes; minimum free form, free form, bound form, phrase; lexeme, nonce-form; idiom.*

The word criterion stems from current field practice. Freedom and bondage were first developed by Bloomfield (especially 1933, chapter 10). The definition of lexeme follows unpublished work of Bernard Bloch. The term "lexeme" has also been used in the sense of our "idiom"; e.g., by Swadesh 1946a.

Problems. The first and second series of problems below are in what may be called *pure distributional analysis.*

In the first series, each problem presents a set of "utterances," each represented by a sequence of one or more capital letters followed by a period. It is to be assumed that, within a single problem, any recurrent letter or sequence of letters has exactly the same sound *and meaning* in all occurrences. Each problem is *closed*: that is, the "language" in question included only those "utterances" which are listed. None of the "languages" involve suprasegmental morphemes. The data in each problem are to be analyzed in terms of morphemes, minimum free forms, phrases, and bound forms. The first problem is worked out to show the procedure.

(1a) A. AC. AD. ADC. B. BC. BD. BDC. C.

Solution: morphemes: A, B, C, and D (each occurs in more than one environment; each, by definition, is unsusceptible of further division).

minimum free forms: A, B, C, AD, and BD (each occurs as a whole utterance, but constituent D in the last two does not).

phrases: A|C, B|C, AD|C, BD|C (the vertical line marks the boundary between ICs; in each case, both ICs are free forms).

bound forms: D (does not occur as a whole utterance).

(1b) A. AE. AEC. AED. ACF. ADF. B. BE. BCF. BDF. BEC. BED. C. CF. D. DF.

- (1c) A. ACH. ADH. AG. AGC. AGD. B. BCH. BDH. BG. BGC. BGD. C. CH. D. DH. EF. EFCH. EFDH. EFG. EFGC. EFGD.
 (1d) A. AC. ACF. ACG. ACH. ACHFI. ACHGI. AD. ADF. ADG. ADH. ADHFI. ADHGI. AE. AEF. AEG. AEH. AEHFI. AEHGI. AH. AHFI. AHGI. B. BC. BCF. BCG. BCH. BCHFI. BCHGI. BD. BDF. BDG. BDH. BDHFI. BDHGI. BE. BEF. BEG. BEH. BEHFI. BEHGI. C. D. E. F. FI. G. GI.

In the second series, the data are the same as for the first series. In each case, the additional analysis required is to determine the form-classes, list the members of each, and describe all the constructions. Form-classes can be named arbitrarily with numbers, as can constructions, and the latter can be described in terms of the labels for the former. We illustrate with the first problem.

(2a) Data as for 1a above.

Solution: form-classes: 1 A, B. 2 C. 3 D.

constructions: 1|3→4. 1|2→5. 4|2→5.

(2b), (2c), (2d) Data as for 1b, 1c, and 1d above.

The final series of problems has to do with the recognition of idioms in English. Remember that complete agreement between different people can hardly be expected—idiomaticity is a matter of degree.

(3a) List all the idioms larger than single morphemes in the following sentences:

- (a) He took off his hat.
- (b) The schoolhouse burned down last night.
- (c) I'm going to town; want to come along?
- (d) They were dancing with abandon.
- (e) He's a dirty four-flusher!
- (f) We should put a padlock on that door.
- (g) Are you afraid of ghosts?
- (h) I'll find out about that if I can.
- (i) If you can't be good, be careful.
- (j) Have you read *The Egg and I*?
- (k) Sticks and stones will break my bones but names can never harm me.

- (l) For lunch we had our choice of toasted English muffins, Scotch broth with barley, French fried potatoes, Welsh rabbit, a Spanish omelet, Italian spaghetti, a western sandwich, or southern fried chicken, with ice cream for dessert. We went Dutch.
- (m) I'm afraid your boy has two strikes against him.
- (n) That new hat is extremely becoming to you.
- (o) That hat becomes you.
- (3b) Find ten examples of idioms larger than a single "word" (in the ordinary sense of the latter term) which involve and are used as verbs. One example: He *ran out of* money.

MORPHOLOGY and SYNTAX

20.1. In many languages, words play an important grammatical role, in that they are built out of smaller elements by certain patterns, but are put together into sentences by rather different patterns. Accordingly, it is customary to regard the grammatical system of a language as composed of two subsystems. *Morphology* includes the stock of segmental morphemes, and the ways in which words are built out of them. *Syntax* includes the ways in which words, and suprasegmental morphemes, are arranged relative to each other in utterances.

To illustrate, we again use the example of §14.1:

³John ²treats his older sisters very ²nicely²↑

The ultimate *syntactical* constituents of this sentence are the intonation morpheme /³ ² ²²↑/ and the seven words *John, treats, . . . , nicely*. The syntactical structure is shown in Figure 20.1; this differs from earlier diagrams only in that the breakdown stops with whole words.

Our sentence as a whole has no morphology: only the individual words in it do. *John* and *very* have the simplest possible morphological structure, since each is a single morpheme. *Treats* consists of the two morphemes *treat* and *-s*; *sisters* of *sister* and *-s*; and so on.

In many cases, sets of words which have similar syntactical privileges of occurrence also have parallel morphological structure, and vice versa. Consider, thus, the English verbs (I) *go, come, run, sing*, and so on, and (II) *goes, comes, runs, sings*, and so on. Verbs of type I occur freely with a subject like *I, we, you, they, the men*, but not ordinarily with *he, she, it, the man*. Verbs of type II reverse this syntactical situation: they occur with subjects of the latter variety (*he goes, the man comes*) but not with

those of the former. But *goes* is morphologically related to *go* as *comes* is to *come* and so on: each verb of type II consists of a verb of type I plus a suffix (in phonemic shape /-z/ in all our examples).

What this amounts to is a *morphological marking* of the syntactical privileges of some words. Some languages have little of this, and the marking is not complete in any language. In English, verbs like (III) *can*, *will*, *must*, *may* are single morphemes just as are those of type I. There is thus no overt morphological difference to mark the syntactical fact that verbs of type III, unlike those of type I, can occur with either variety of subject (*I can*, *he can*, etc.).

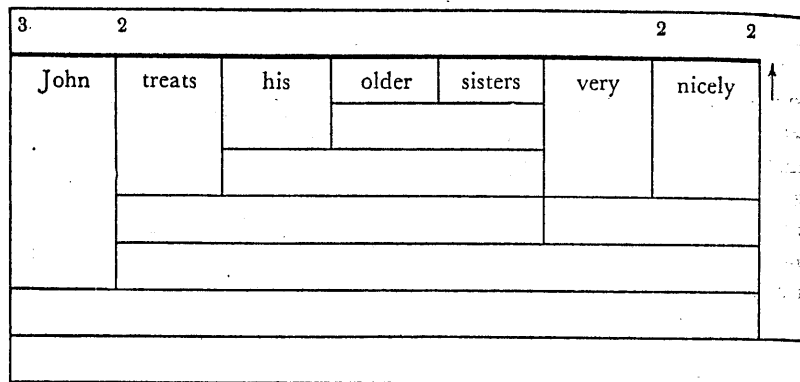


FIGURE 20.1

20.2. The Morphology-Syntax Boundary. The line of demarcation between morphology and syntax is not always as clear-cut as our discussion so far may have suggested.

English *twenty-eighth* illustrates the difficulty. The ICs are *twenty-eight* and *-th*. The construction in which these ICs stand is not clearly morphological, because the constitute built by the construction is a phrase, not a single word. Nor is it clearly syntactical, since one of the ICs (*-th*) is less than a word. Phrases of this special structure are quite common in English, and not rare in other languages.

The best procedure seems to be to adjust the definition of "morphology" so as to include all constructions in which one IC is less than a word, even if the other IC and the constitute are sometimes phrases rather than single words or parts of words. Thus we class the English

construction by which ordinalizing *-th* is added to a number as morphological, though with the special property that the number may be a phrase (*twenty-eighth*, *three hundred seventy-fourth*) rather than a word. We do the same with the English construction by which genitival *'s* is added to a word or phrase (*John's*, *the Mayor of Boston's*). Constructions which we should hardly expect to turn up in this special use sometimes do. The formation of an adjective from a noun by adding *-ial* with a shift of stress is fairly common: *mánor* : *manórial*, *díctàtor* : *díctatórial*, *repórter* : *repórtórial*.¹ The writer has heard (a) *lord-of-the-manorial* (*air*), in which *-ial* was added to the phrase *lord of the manor*, not just to *manor*.

Menomini affords a parallel:

/ahsa·ma·w/ 'he is fed' /ke·s·ahsa·ma·w/ 'he has been fed'
/neta·hsamaw/ 'I feed him' /neke·s·ahsa·ma·w/ 'I have fed him.'

(The alternations of vowel length are morphophonemic and here irrelevant.) The form 'he has been fed' is built by a syntactical construction from 'he is fed' with a separate word /ke·s/ 'completion.' The form 'I feed him' is built by a morphological construction from 'he is fed,' with a prefix /ne-/ 'I, me.' This same morphological construction, with the same prefix, applied to the phrase 'he has been fed,' yields the phrase 'I have fed him.'

A consequence of our adjustment in the definition of "morphology" is that we cannot always break a sentence down into successive layers of ICs, finding only syntactical constructions until we have reached the level of words, and only morphological constructions from then on. If we analyze *She arrived on the twenty-eighth day*, we find only syntactical constructions until we reach *twenty-eighth*. At this point a morphological construction appears (*twenty-eight* and *-th*). But then, at a still lower hierarchical level, we again find a syntactical construction: that of *twenty* and *eight*. In this sense, the line of demarcation between morphology and syntax remains ill-defined even though we are able to class constructions themselves unambiguously as morphological or as syntactical.

20.3. Complexity. English words rarely achieve great morphological complexity. *Ungentlemanliness* and *impressionistically* are about as far as English goes, barring a few artificial monstrosities like *honorificabili-*

¹ The colon preceded and followed by space sets off contrasting forms for comparison.

tudity. Figure 20.2 shows the structure of *ungentlemanliness*: there are six morphemes and four layers of ICs. Fox /e-howi-kiči/, diagrammed in Figure 20.3, shows about the same complexity; but the latter word is of only average complexity for Fox.

On the other hand, English syntax can get even more complicated than the morphology of Fox or of other languages which, like Fox, tend

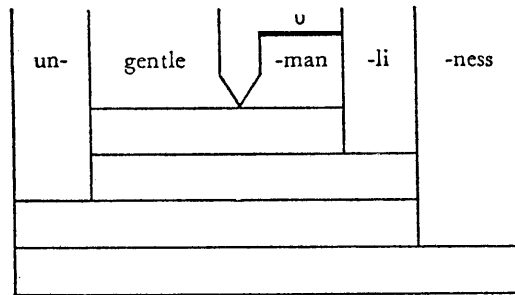


FIGURE 20.2

The absence of stress on *-man* is taken as a separate morpheme; the mark "u" represents this distinctive lack of stress.

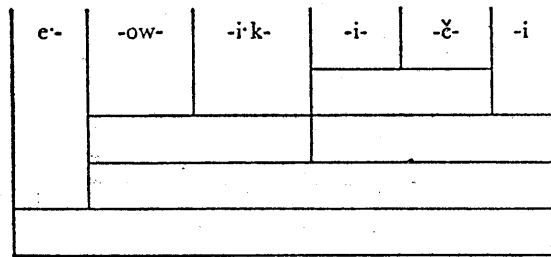


FIGURE 20.3

to put much material into individual words. The first sentence in this paragraph is average for expository English. If one diagrams it one finds about twenty IC layers.

Objective measurement is difficult, but impressionistically it would seem that the total grammatical complexity of any language, counting both morphology and syntax, is about the same as that of any other. This is not surprising, since all languages have about equally complex jobs to do, and what is not done morphologically has to be done syntac-

tically. Fox, with a more complex morphology than English, thus ought to have a somewhat simpler syntax; and this is the case.

Thus one scale for the comparison of the grammatical systems of different languages is that of average degree of morphological complexity—carrying with it an inverse implication as to degree of syntactical complexity. The easiest rough measure of morphological complexity is the average number of morphemes per word in a representative sample. Mandarin Chinese scores very low, with barely more than one morpheme per word on the average. English shows nearer to two morphemes per word; Spanish about two and one half; Latin about three; and Fox nearly four.

Nineteenth-century scholars tried to class languages, not along a scale, but into one or another of a limited set of pigeonholes. Among their classificatory terms were *analytic*, *synthetic*, and *polysynthetic*. Thus, they thought that Chinese words were always just one morpheme, and classed the language as "analytic." Greek, Latin, and Spanish, with more morphemes per word, were "synthetic." When languages like Fox were discovered, showing an even higher count, "synthetic" seemed inadequate and the term "polysynthetic" was added. A continuous scale is better than this sort of pigeonholing. But we can conveniently use the terms "analytic" and "synthetic" in a relative way, saying, for example, that Spanish is more synthetic than Chinese but more analytic than Fox.

There is no discernible correlation between the placement of a language on the analytic-synthetic scale and anything else about either the language or other aspects of the life of its speakers. Some Nineteenth-century scholars proposed theories to the contrary, some of which have become part of the folklore about language current among educated laymen today. For that reason it is important to mention these theories and emphasize their falsity.

One false theory was that in course of time all languages tend to become increasingly analytic. There are attested instances of this direction of change: Old English to modern English, Latin to the French of the Fourteenth or Fifteenth century. But there are also instances of the opposite direction of change: French is somewhat more synthetic now than it was a few centuries ago.

A further misconception, a sort of corollary of that just mentioned, was the notion that some languages of today, especially English, are

more "progressive" than others, like Spanish and German, because they have developed further in the analytic direction. For this there is no evidence at all. The Turks of today manage all the business of everyday life, and the complexities of modern technology, with a highly synthetic language; the Chinese of today do just as well with a markedly analytic language.

NOTES

New terms: *morphology* and *syntax*; (*ultimate*) *syntactical constituents* (= words and suprasegmental forms); *analytic* versus *synthetic*. "Polysynthetic" is superfluous. Some contemporary linguists use the term "morphology" to subsume all that we divide up into morphology and syntax; some logicians and semanticists use the term "syntax" or "syntactics" in this same broader way; our usage of the two terms follows that of Bloomfield 1933.

Problem. Determine and list the ultimate syntactical constituents of each of the following English sentences:

- (a) The scientist walked like a man on an errand that was too important to be interrupted.
- (b) His two brothers-in-law joined the Air Force.
- (c) I put on my prayers.
- (d) Good thirty-five to forty; utility thirty to thirty-five.
- (e) I roaded the car five miles this side of Grinnell.
- (f) They tape recorded the whole series of lectures.
- (g) He's a swash-buckling buccaneer.
- (h) These differences were noted as a fact of immediate apprehension.
- (i) It looks like on this bet we're going to what they call in black-jack push.
- (j) Aspirin is monoacetosalicylic acid.

21.

SYNTACTICAL CONSTRUCTION-TYPES: ENDOCENTRIC

21.1. Construction-Types. The sentence *The old dog lay in the corner* contains two composite forms, *old | dog* and *lay | in the corner*, built by different constructions but nevertheless showing certain similarities. In terms of meaning, an old dog is one kind of dog, and lying in the corner is one kind of lying. In each case, then, one of the ICs modifies the meaning of the other. This is not true of all composite forms. *Men and women*, with ICs *men* and *women*, refers neither to one kind of men nor to one kind of women; *visit Bill* refers neither to one kind of visiting nor to one kind of Bill.

If a constitute built by one construction (say construction A) and a constitute built by another (say B) show a certain similarity, then any other pair of constitutes, one built by A and one by B, show the same similarity. We can thus speak directly of similarities between constructions. A *construction-type* is a group of constructions which are similar in some specified way. *Old | dog* and *lay | in the corner* are built by different constructions, but the constructions are of the same type in that both involve the modification of one IC by the other.

Construction-types are useful in comparing languages. Constructions themselves have to be defined separately for each language: a French composite form and an English one cannot be built by the "same" construction. But a French construction and an English one can be of the same type.

Construction-types are also helpful in dealing with a single language, because there are instances in which it is not easy to tell whether two