(3) Hartman, §7, says "there is ... nothing in the distribution to suggest that the aspirates [/ph th ch kh/] are anything but unit consonants." Here our phonemic working assumptions differ; Hartman is analyzing occurrent material into phonemes; I analyze into determining (and determined) features; Hartman uses distributional facts as the basis for phonemic conclusions, I find the determining features and then state the distributions. The aspiration of /ph th ch kh/ sounds like the independent aspiration /h/; so I take it to be the same thing.

(4) Hartman does not, in his paper, recognize tone /5/ ("raised third") as phonemically distinct; but it was Hartman who, after publication of his paper, discovered cases of minimal contrast which establish /5/ as different from both /2/ and /3/.

Addenda: Since this report was submitted several facts have come to light (mainly through the courtesy of Yuenren Chao) which bear on the discussion.

(1) (To §8.23) The following monosyllabic microsegments, excluded in the text, are attested: /sen, seng, sengr, liun, nen, neu, tiang, tiangr, thei, pia, phia, chrei, rua/. /len/ occurs elsewhere than in the name referred to (§8.23, (5)). Many of the above additions are from onomatopoetic forms.

(2) (To §§9.31, 9.33) In the sequence /Ci .../, if /C/ is /t th n l/ the consonant is not palatalized; rather, the /i/ is a rather low high-front glide or vowel, not overlapping the consonant.

(3) There are a number of interjections having vocalic structure not subsumable under the system here set up. In addition, there is a minimum contrast between /ˌieu2 ˈcing3/ 'oil wells' and what we would here transcribe identically in the meaning 'there are wells,' where the first syllable of the latter might also bear tone /5/. /ieu2/ in the first of this pair has a relatively high back rounded vowel; in the second of the pair, a considerably lower vowel (identical with that of /ieu3/ or /ieu4/). In the light of this phenomenon there is serious doubt as to whether the two-vowel system, or its three-vowel predecessor, can be maintained. Perhaps it can, if we accept the modification suggested in §9.55.

[[On page 108 I spoke of the later trend towards what I called 'phonemics without the phoneme', and now this paper appears as its first exemplification here; others will follow. Some credit is due, at least, (and Hockett among others has given it) to stimulation from the Trubetzkoy-Jakobson theory of distinctive features, but it would also be possible to call this an autonomous American development. Its autonomy is evident in both negative and positive ways.

Negatively, there is for example the refusal to take the distinctive features as universals, exactly nine polar sets of them being extant—or is it ten?—for all the world's languages to use, and the same distinctive oppositions being found in disparate parts of the phonology, e.g. 'acute-versus-grave' alike in the stop-system and the vowel-system; instead, the distinctive features are established ad hoc for each language or even dialect. (The Trager theory of the 'over-all pattern', identical for all the dialects of a language, could be derived from these distinctive features; but as a matter of historical fact seems to have developed independently of them.) Again, any resolution of distinctive-feature sets into binary components is treated as a game, not a discovery of ultimate truth.

Negative or positive would both be good interpretations of the American rejection of the limitation to segmental elements and corresponding insistence on the importance (or even primacy) of the 'suprasegmental' frame in which the segmental elements are found to occur in all languages. Related to this is the readiness to recognize dynamic segmentals such as semivowels, against the Jakobson way of treating e.g. English /w/ as simply an /u/ vowel.

In short, there is an air of independence in it all.

The positive trend of the American development, which makes it look like an autonomous sequel to earlier American phonemics, is the one which I have characterized ((108)) as 'making the theory independent of the accidents of analytic practice'. The practicing analyst, in the American sense, is the one who has a language previously quite unknown to him to work on and has to do everything from the ground up. (He is likely to sympathize with those who complain that European phonologists avoid the spade-work and just shuffle other people's data, even when he knows that the accusation is not true.) For him, analysis begins with his own phonetic transcribing. This is segmental because his marks on paper have to be countable. Hence, when (if ever) he becomes a sophisticated theorist, he views his segmented transcription with dark suspicion. The segmentation needs justifying. In his Boas tradition (languages can differ without limit as to either extent or direction), no universal theory of segments can be called upon to settle the moot points. Only the particular language can yield proper criteria. Where these fail, segmentation is arbitrary. There have to be enough units to picture the significant differences among utterances: this is the crosswise dimension, 'at right angles' to the time-flow of speech, and it is in this dimension that one makes discoveries which increase the number of phonemic symbols. To reduce their number, one looks for distributional criteria in the other, the time-flow dimension. In Chinese, these criteria fail with at least one question still unresolved.

To make sure that the description has been freed from segmental prejudice, it is expedient to redo it as if one had started out from the other end. Hence the first paragraph on this page.

The word 'phoneme' can be and often is avoided, but the adjective 'phonemic' still labels distinctive differences; then the orphaned noun 'phonemics' is attached to it, and the adverb 'phonemically' is kept.]]

PROBLEMS OF MORPHEMIC ANALYSIS

CHARLES F. HOCKETT
Language 23.321–43—1947

I. Introductory

1. This paper develops further the theory of morphemic analysis presented by Zellig S. Harris in 1942.[1] Morphemic analysis is the operation by which the analyst isolates minimum meaningful elements in the utterances of a language, and decides which occurrences of such elements shall be regarded as occurrences of 'the same' element.

This operation does not constitute all of grammatical analysis: when it is completed, there remains the task of describing the arrangements in which the minimum meaningful elements occur, and—where the same elements are observed to occur in more than one arrangement with a difference in meaning—the features other than morphemes (the 'tagmemes') that are involved. For this latter phase, I have proposed the term 'tactics'.[2]

The fact that John hit Bill and Bill hit John have different meanings,[3] or that old men and women is sometimes approximately the same as 'old men and old women', sometimes rather comparable to 'women and old men',[4] proves that features other than morphemes have to be recognized—unless, of course, we wish to redefine 'morpheme' to cover features of this type too. There is other evidence: a graduate student in a hurry to prepare himself for a French reading exam, or a scholar deciphering a dead language written in a non-phonetic or semi-phonetic orthography, may achieve good control of the tactics and semantics of the language, but remain in almost total ignorance of anything submorphemic. To do this he needs some mnemonically satisfactory device for keeping morphemes apart. The device probably consists of speech sounds;[4a] but these may be purely private. Thus a western sinologist may know Confucius backwards and yet stumble in passing the time of day with any speaker of a modern Chinese dialect.

Although, then, morphemics and tactics are both necessarily involved in grammar, we nevertheless have considerable range of choice in drawing the line between them.[5] Faced with a language of a certain degree of complexity, we may prefer to describe it with simple morphemics and complicated tactics, or conversely, or somewhere in between. The language is not disturbed by our choice; its complexities remain whether itemized in one part or another of our description. But the resulting descriptions may vary a great deal in the clarity with which they depict the situation. Presumably we should try to obtain that distribution of data between morphemics and tactics which produces the greatest clarity. In this paper we assume, without steadfast conviction, that this end is achieved by the simplest possible tactics, whatever submorphemic complications may be necessitated.

2. The same assumption was apparently involved in Harris's formulation of 1942. Yet Harris realized that this cannot stand as the only assumption. We must have, also, a set of principles on the basis of which we identify, or refuse to identify, different stretches of speech as morphemically the same. The great value of Harris's paper lies in this: that although he does not add any individual method of morphemic identification to those currently used, he demonstrates how all the superficially diverse methods can be regarded as cases of one general procedure. This general procedure we outline herewith, with such minor modifications of terminology as will be useful to us:

Step 1. The utterances of a language are examined.[6] Recurrent partials with constant meaning (ran away in John ran away and Bill ran away) are discovered; recurrent partials not composed of smaller ones (-way) are alternants or morphs.[7] So are any partials not recurrent but left over when all recurrent ones are accounted for. The citable case most nearly approaching this is the cran- of cranberry, which does indeed recur, but always with berry following. By definition, a morph has the same phonemic shape in all its occurrences. Because we

[1] Zellig S. Harris, Morpheme Alternants in Linguistic Analysis, Lang. 18.169–80 (1942) ((109)).

[2] Review of Eugene A. Nida, Morphology: the Descriptive Analysis of Words, Lang. 23.273–85 (1947).

[3] Leonard Bloomfield, Language §10.4 (New York, 1933).

[4] Rulon S. Wells, Immediate Constituents, Lang. 23.81–117, esp. 93 ff. (§§30 ff.) (1947).

It would be possible to say that this ambiguity of old men and women was grammatically irrelevant; but features of order of the type involved in John hit Bill versus Bill hit John cannot be ignored. This being so, Bloomfield's term 'tagmeme' for a feature of meaningful arrangement is useful.

[4a] This is a reasonable assumption because of man's million years or so of natural selection, in which ability in aural memory and oral mimicry has been a factor making for survival.

[5] Cf. Zellig S. Harris, From Morpheme to Utterance, Lang. 22.161–83, esp. 162–3 (§2) (1946) ((142–3)).

[6] Obviously not all of them, but a sampling which we hope will be statistically valid. By working with successively larger samplings, and by predicting on the basis of each what else will occur, we approach, at least asymptotically, a complete description.

[7] A convenient term, because it (1) eliminates the lengthy expressions 'morpheme alternant' and 'morpheme unit', and (2) suggests a valid analogy (allo)phone : phoneme = morph : morpheme.

operate with whole utterances, morphs are not always composed of continuous uninterrupted stretches of phonemes,[8] but they are always composed of phonemes. Every utterance is composed entirely of morphs. The division of a stretch of speech between one morph and another, even if the two are simultaneous, overlapping, or staggered, we shall call a cut.

Step 2. Two or more morphs are grouped into a morpheme if they '(a) have the same meaning, (b) never occur in identical environments, and (c) have combined environments no greater than the environments of some single alternant in the language',[9] e.g -en in oxen, /z/ in cows, and various others, all meaning 'noun plural', with combined environments, or range, paralleling the range of zero with meaning 'noun singular'.[10]

Step 3. The differences in the phonemic shape of alternants of morphemes are organized and stated; this constitutes morphophonemics. Morphophonemic statements may involve morphophonemes—that is, the symbols used for phonemes, plus supplementary ones, with special definitions as to phonemic value under varying circumstances—or they may not; often lists are more convenient, and sometimes they are unavoidable. But regardless of the methods used in describing them, such alternations are morphophonemic.

3. In several ways a rigorous adherence to Harris's system as here stated is troublesome.

(1) Sometimes we are confronted with a set of alternants with apparently identical meaning which are almost, but not quite, in complementary distribution. So with the two alternants meaning 'noun plural' in hoofs and hooves, or laths with /θ/ and with /ð/. These would forbid the tactically desirable conclusion that there is but one noun-plural morpheme in English.

(2) Sometimes a set of alternants with identical meaning and completely in complementary distribution have to be kept apart because we can find no single alternant whose range parallels that of the given group. In Latin, for instance, there is no case-number combination represented after all noun stems by the same suffix; therefore we may not legitimately (by Harris's criteria) speak of a single 'nominative-singular' morpheme, or a single morpheme for any other case-number meaning.

(3) Sometimes a stretch of speech may be cut at either of two places, so as to produce equally satisfactory—and equally unsatisfactory—morphs. In Menomini, when an element ending (otherwise) in a consonant precedes, in the same word, an element beginning (otherwise) in a consonant, an /e/ appears between them. Do we cut before or after this /e/? Either cut will do; either choice is arbitrary. Harris proposes that we cut in both places, and regard /e/ as an alternant of /-/ 'morpheme juncture'.[11] In this proposal he does not adhere to his own rules, for morpheme juncture has no meaning, and is not a morpheme; yet any Algonquianist will say that his solution is correct, and the problem is to readjust the rules so that the interpretation does not violate them.

(4) Since there is no way in which French /o/ 'to the (masc.)' can be cut, we must take it as a single morph. But the tactical survey suggests rather that it be taken as two successive morphemes, à 'to' plus le 'the (masc.)'. There is at present no way in which the latter conclusion can be reached without doing violence to our criteria.

(5) As we perform step 1 of Harris's procedure, only morphs of overt phonemic content turn up. It is suggested that the definition of morph be extended to cover also the following: minus-features, such as that which added to French bonne 'good (fem.)' produces bon 'good (masc.)'; replacement features, as man : men; zero features, as in sheep (sg.) : sheep (pl.); and combinations of these, such as the difference between child and children. This is a difficult maneuver, however desirable; Harris (within the scope of his paper) tells us neither under what conditions it is called for nor how to perform it.

The items just listed are not criticisms, but points on which improvement is clearly possible within Harris's general framework. The first two difficulties are easily handled; the remaining three are more serious, but respond to a single modification in plan of attack.

Many of the problems of morphemic identification met with in dealing with any language are trivial. Before turning to the full-scale discussion of the five difficulties listed above (in Parts III and IV below), we attempt to show how the more trivial problems can be solved quickly and easily, in a fashion that sheds light on the more intricate questions to which one must eventually turn.[12]

8 The possibilities are investigated by Harris, Discontinuous Morphemes, Lang. 21.121–7 (1945)—but the added complication of this is avoided in the examples of the present paper.

9 Harris, Morpheme Alternants §7.1.

10 The zero element with meaning 'noun singular' is one of Harris's parallels (Morpheme Alternants §2.2). Such a morpheme has very dubious status, having no alternant of other than zero shape (see fn. 37 and reference cited there). Harris lists also the parallel -ful; given the modification of criteria proposed in §13 of this paper, one could add also 's 'genitive'.

11 Harris, Morpheme Alternants §4.2.

12 We propose to say both 'the morph x occurs in such-and-such an utterance', and 'the morpheme x occurs in such-and-such an utterance'. By our definition, a morpheme is a class of morphs, so that the latter type of expression, without further qualification, is logically invalid. We render it valid by stating that an expression of the form 'the morpheme x' shall be taken in some cases as a class-name, in other cases as a variable indicating the appropriate though unspecified member of the class, depending on what the context requires. No ambiguity results; this is customary usage in linguistics; but it is a point on which more care is needed than is usual.

II. Preliminary Normalization

4. Let us assume that we have before us a display of a large number of utterances of a language, in a phonemic notation. As we begin the search for recurrent partials, we may discover that a phonemic notation other than the one we have used—for there are always several mutually convertible possibilities—would simplify the task.

In Yawelmani,[13] for example, the point of syllable division is phonemic. One way to write Yawelmani is to use a hyphen for syllable juncture; then the phonemic content of syllables can be indicated with a relatively small number of vowel and consonant letters. If our display of utterances is in this notation, we find such obviously related forms as /g?ads?/ 'obsidian' : /g?ads?-ni/ 'obsidian (dative)' : /g?a-ds?a/ 'obsidian (accusative)'. The second form contains a stretch identical with the first, plus /-ni/ 'dative'; the third form contains, before /a/ 'accusative', a stretch identical with the first part of the other two save for an inserted hyphen. If we are to identify the non-identical stretches /g?ads?/ and /g?a-ds?/ as being morphemically the same (whether one morpheme or more is another matter), and make similar identifications in other cases where the presence or absence of a hyphen is the phonemically differentiating factor, then we must handle this evanescent hyphen in our morphophonemic statements.

Yawelmani can also be written phonemically without the hyphen. If we want to write it so, we must use unit symbols for certain consonants which otherwise might be interpreted either as belonging wholly to a single syllable or as being divided between two: for example, /Vg?V/ would be an ambiguous notation for both /V-g?V/ and /Vg-?V/; but if we replace /g?/ in a single syllable by /k'/, the ambiguity is removed.[13a] With all such changes as are necessary, we reach a notation which does not write the point of syllable division with a separate symbol, but which nevertheless indicates it unambiguously: when a single consonant symbol stands between two vowel symbols, a point of syllable division falls before the consonant symbol; when two consonant symbols stand between two vowel symbols, a point of syllable division falls between them.

In this notation, the forms given above appear as /k'ac'/ : /k'ac'ni/ : /k'ac'a/. The partial /k'ac'/ appears to be identical in all three forms. Phonemically, of course, it is not; but the only phonemic difference has been relegated to a status of notational predictability, and can be ignored in our further manipulations. There are so many intricate problems in Yawelmani morphemics that any advantage of this kind that we can obtain is greatly to be desired.[14]

5. A notation is phonemic if it indicates, in every position, only those phonemic contrasts which occur in that position, but indicates all of them. Once one has found the morphemically most desirable phonemic notation, one can often handle certain additional simple morphemic problems by modifying it in such a way that, in addition to indicating unambiguously all the phonemic contrasts occurring in a position, it also indicates in certain positions contrasts which are not there phonemic.

If our display is of Navaho utterances,[15] we notice, sooner or later, that vowel symbols do not occur before pause (P). The sequence /VhP/ (with V for any vowel symbol) does occur. When we examine the display for recurrent partials, we find certain stretches with constant meaning that end in /Vh/ both before P and elsewhere, others that end in /Vh/ before P but without the /h/ elsewhere. Thus /bìtàhP/ 'among them' : /bìtàh níhP/ 'among them, he says', but /dò·dàhP/ 'not' : /dò·dà níhP/ 'he says no'. The morphemic identification of /dò·dàh/ (P) and /dò·dà/ (no P) is elementary, as are other such cases. So we modify our notation throughout the display, by erasing certain pre-pause h's—namely, those at the end of stretches that occur medially with the same meaning but without the /h/. In our notes we enter the memorandum: both V and Vh before pause represent phonemic /Vh/. Thereafter, save when reading off our transcription with Navaho speech sounds, we ignore the memorandum; the new /dò·dà/ now has the same shape, to the eye, before P and elsewhere.[16]

Or suppose that we are dealing with Latin. We find pairs like ars : artis, noks : noktis, urps : urbis,

13 Stanley S. Newman, The Yokuts Language of California (New York, 1944). The phonemicity of the point of syllable division is my conclusion from the evidence he gives.

13a In this notation the letter k and the apostrophe ' are meant to constitute one symbol together. Similarly c and ' below.

14 The second notation is that used by Newman. It may be wondered why anyone would be led to investigate the potentialities of our first notation, the one that we decided to reject. But in Southern Athabascan (see citations in fn. 15) an entirely similar problem arises, and Hoijer chooses an orthography comparable to our first Yokuts orthography, not to our second. The complexity of morphophonemic statement which results is considerable, and could be rendered measurably easier if a phonemic notation were used in which syllable division is marked indirectly instead of overtly.

15 Harry Hoijer, Navaho Phonology, University of New Mexico Publications in Anthropology I (Albuquerque, 1945); a similar phenomenon in Chiricahua Apache: Harry Hoijer, Chiricahua Apache, Linguistic Structures of Native America 55–84 (New York, 1946).

16 Hoijer's working notation incorporates this normalization, though he calls the contrast between 'constant' pre-pause h and the evanescent type phonemic instead of morphophonemic. A similar normalization leads to the writing, within word borders, of phonemes both of the s-series and of the š-series, although, at least in rapid speech, only those of a single series occur within the stretch bounded by word junctures; see Zellig S. Harris, Navaho Phonology and Hoijer's Analysis, IJAL 11.239–48 (1945).

re·ks : re·gis, niks : niwis. The semantic and morphemic difference between ars and artis recurs with other pairs, but the difference in phonemic shape between the members of a pair is not so constant from one pair to another. Whatever may be our ultimate morphemic conclusions (e.g. that artis is ars plus something, or that ars and artis are both art plus something), they will be more easily reached if we can make the difference in shape to the eye parallel the morphemic difference.

This can be done. From phonemics we know that the sequences rts, kts, rbs, gs, gws do not occur before word juncture,[17] and that gw does not occur intervocalically. We may therefore rewrite the forms with precisely these non-occurrent sequences: arts : artis, nokts : noktis, urbs : urbis, re·gs : re·gis, nigws : nigwis. We note that in the modified orthography, rts and rs (before word juncture) are both representations of phonemic /rs/, and so on.[18] In the new notation, the second form of each pair differs from the first only in the presence of an i before the final consonant.

6. Sometimes it helps to perform this type of normalizing operation more than once.

In Potawatomi, a first normalization introduces, at certain points within utterances, a mark (say a space) indicating potential pause. Then we examine the stretches between successive points of potential pause to see which ones recur in various positions (relative to actual pause, or to adjacent stretches of varying structure) with the same meaning and the same or almost the same phonemic shape. Neither preceded nor followed by (actual) pause, we find kak 'porcupine', k·we 'woman', muk· 'beaver', k·uk· 'bucket'. Preceded but not followed by pause, we find rather kak, kwe, muk·, kuk·. Followed but not preceded by pause, the forms with the same meanings are kak, k·we, muk, k·uk. Both preceded and followed by pauses, the forms are kak, kwe, muk, kuk.

Now observation shows that the phonemes /p·/, /t·/, /č·/, /k·/ occur neither directly after nor directly before pause. Therefore we normalize all such forms as those itemized above, in all positions in which they occur, to kak, k·we, muk·, k·uk·, with the necessary memorandum that when either preceded or followed by pause, both k and k·, both p and p·, etc., represent phonemic /k/, /p/, etc.

This same normalization also accounts for all other alternations turning on the presence or absence of pause. When we have retranscribed our entire display of forms, we look further, and discover such pairs as nkutšuwe 'he wins a race' : nnuktušwe 'I win a race', kwtumočke 'he's fishing' : nkwutmočke 'I'm fishing', pmos·e 'he's walking' : npums·e 'I'm walking', msun?ukun 'paper' : nmusnu?kun 'my paper'. If we want to, we can begin at this level to cut our forms up into smaller recurrent partials. Clearly n means 'I, my'; the remainder of the first form, nuktušwe, presumably means 'win a race', but is not the same morph as nkutšuwe '(he) win(s) a race', because of the difference in phonemic shape. For that matter, each of these may be more than one morph. Whatever comes of this, our morphophonemic statements are going to have to be complicated at every stage by an alternation of the positions in which vowels appear.

So instead of continuing our comparison and cutting, we can try first to make a further notational normalization that will take care of the alternating vowels, or many of them. We do this by rewriting the original forms, writing a vowel in both forms of a pair whenever it appears in either form, and in a few other places for good measure: nŪkUtUšUwe : nUnUkUtUšUwe, kwUtUmočUke : nUkwUtUmočUke, pUmOs·e : nUpUmOs·e, mUsUnU?UkUn : nUmUsUnU?UkUn. Our memorandum this time states how these graphs are to be interpreted, not directly into a phonemic notation, but into the supraphonemic notation achieved by the previously applied normalizations: the first and each alternate one of a series of capitalized vowel symbols counts as zero, unless it precedes a final consonant.[19]

The second form of each pair now consists, to the eye, of nU (presumably 'I, my') followed by something that is identical with the first form of the pair. All the remaining problems of morphemic isolation and identification are rendered simpler.

7. There is no real drawback to counter the advantages of this kind of preliminary notational normalization, but there is a caution which must be observed. Our notational changes make morphs of differing phonemic shape look alike—indeed, that is why we make them. But the ultimate problem of the grouping of such morphs into morphemes is one which must be solved in a manner consistent with our handling of less patent cases—that is, on the basis of Harris's criteria (Step 2, §2) or of some other set. Performed as we have here suggested, notational regularizing is not apt to obscure more desirable morphemic identifications; but in extensive work with any specific language, one needs to check back over such preliminary operations from time to time to make sure.

17 Though it is not clear what word juncture is in Latin; it may be a non-phonemic matter introduced by a previous notational normalization.

18 Full phonemic information is still given, since such a graph as rts before word juncture stands always for phonemic /rs/, never anything alse. In the new notation we have multiple writings for certain phonemic sequences, but only one phonemic sequence for each writing.

19 We are forced to use capitals or some other device for evanescent vowels, because other vowels, phonemically the same, are not evanescent. This fact marks these alternations as non-automatic. Where no extra symbols are needed—where the symbols already used phonemically are merely extended to positions in which they do not phonemically occur—the alternations are automatic.

III. Revision of the Grouping-Requirements

8. In step 2 (§2) are stated Harris's three grouping-requirements—the three conditions which must be met by two or more morphs if they are to be regarded as belonging to the same morpheme. Some of the troubles itemized in §3 result from the particular way in which these grouping-requirements are formulated. The first of them, involving meaning,[20] is obviously the most difficult to handle. We attempt no revision of it here, and any choices dependent on it under the Harris procedure will remain in the present scheme. The second and third requirements are purely distributional, and more easily subject to analysis and modification.

Preliminary to our proposal for a modification of the second grouping-requirement we define non-contrastive distribution. Two elements of the same kind (i.e. both allophones, both morphs, or the like) are in non-contrastive distribution if either (1) they are in complementary distribution, or (2) they are in partial complementation, and in those environments in which both occur, they are in free alternation. By free alternation is meant (a) that one cannot predict, save perhaps statistically, which form will occur in a particular instance, and (b) that the occurrence of one, rather than of the other, does not produce an utterance different in meaning.

In phonemic analysis, non-contrastive distribution is often used as a criterion permitting the grouping of two or more allophones into a single phoneme. Thus the unaspirated [t] of stick and the aspirated [t'] of tick are both found at utterance-final and in certain other positions: He's in the skit may end with [t] or with [t']. But this utterance is 'the same' utterance whether the aspiration is present or not; similarly with any other pair differentiated only in this respect. The occurrence of both allophones in certain environments does not deter us from classing both in the same phoneme, /t/.

We propose, then, to revise the second grouping-requirement from 'never occur in identical environments'—which is another way of saying 'are in complementary distribution'—to read 'are in non-contrastive distribution'.[21]

The examples which follow (§§9–11) will demonstrate both the way in which this change increases the efficiency of our analysis, and also a danger inherent in it.

20 In a manuscript not yet published, Harris demonstrates how, at least in theory, this criterion can be eliminated, thus appealing to semantic considerations at only one step of the whole process of descriptive analysis: the step at which one must decide whether two utterances, as historic events, are 'the same' or not (Bloomfield's fundamental assumption of linguistics, Language §5.3, §9.5). The first grouping-criterion (same meaning) thus becomes a practical shortcut; as such it is used here.

21 In Yokuts Structure and Newman's Grammar, IJAL 10.196–211 (1944), Harris makes explicit use (§6) of the second grouping-requirement as modified, but without theoretical discussion.

9. In Modern West Armenian[22] a number of morphs occur with meaning 'genitive singular': /o/ and /0/ with one stem of preceding noun, /van/, /u/, /an/, and /i/ with other stems. The environments in which these occur can be differentiated in terms of the nouns which immediately precede, and such nouns fall into a series of classes (of purely morphophonemic importance) by virtue of the morph or morphs of this meaning which follow them. One occurs only with /o/ and the oblique stem: /asdəvaʒ/ 'God' : /asduʒ.o/ 'of God'. Some occur only with /0/ and the oblique stem: /kuyr/ 'sister' : /kəroč/ 'of sister, sister's'. Some occur only with /i/ and the singular stem; this is the most common pattern: /atoř/ 'chair' : /atoř.i/. Most of those which occur with /van/ and the singular stem, for example /irigun/ 'evening' : /irig.van/, are also observed to occur with /i/: /irigun.i/. The same is true for most of those which occur with /u/ or with /an/: /meg/ 'one' : /meg.u/ or /meg.i/; /axčig/ 'girl' : /axčəg.an/ or /axčig.i/.

The morphs meaning 'genitive singular' are thus not in complementary distribution, and by the original form of the second grouping-requirement could not be combined into a single morpheme. However, there is no observable difference in meaning between, say, /axčəgan/ and /axčigi/; nor is the speaker's choice of one or the other of these on any particular occasion predictable, save perhaps statistically. Therefore, within the limits of semantic judgment available to us at present, the various morphs in question are indeed in non-contrastive distribution, and by the modified second grouping-requirement—providing the third grouping-requirement is also met—are classifiable as a single morpheme.

10. In Peiping Chinese there are, as the elementary texts usually put it, two words for 'two': èr and lyǎng. The latter graph, by a preliminary normalization that need not concern us here, subsumes two distinct forms, lyǎng and lyáng. Unit numerals (those for 'one' through 'nine') occur in Chinese in the following positions: before a measure; after a group numeral ('ten' etc.); before a group numeral; after an ordinal demonstrative (dì 'the ...th'); preceded and followed by unit numerals (in counting 'one, two, three,...'). Sān 'three' appears in all these positions with a single phonemic shape. Èr occurs in the second, third, fourth, and fifth positions, lyǎng in the first and third. In the third position, before a group numeral, only èr occurs before shí 'ten', but either may occur before bǎi 'hundred', chyān 'thousand', and wàn 'ten thousand' and its multiples. The choice here is free. Èr also occurs in position one, to the exclusion of lyǎng, before the measure lyǎng 'tael, ounce'; and before a few other measures either èr

22 Information and forms were kindly supplied by Gordon H. Fairbanks. It does not matter for the present discussion whether the stem differences are submorphemic within the stem or are part of the suffix. Some other complications, which do not alter the picture materially, are omitted here. In the cited genitive forms, a dot separates stem from ending.

or lyǎng may be found, with no difference in meaning.

This statement of distribution is not quite exhaustive, but a completely exhaustive one reveals the same facts. Clearly, èr and lyǎng cannot be regarded as a single morpheme under the old form of the second grouping-requirement, but can under the new form.

The case of èr and lyǎng is unique in Chinese; most other sets of morphs differing in phonemic shape but classed nevertheless as belonging to single morphemes have at least some phonemic feature in common. If this factor is to deter us in grouping èr and lyǎng together, then some additional grouping-requirement, not mentioned by Harris, and apparently quite difficult to formulate in a strict fashion, must be involved. We mention this possibility because we feel the Chinese example 'instinctively' to be somewhat different in nature from the Armenian.

11. As indicated in §3 (1), Harris's example of English noun-plural morphs will not hold by a strict application of his criteria, because of such pairs as hoofs : hooves. One of the more interesting of such pairs is brothers : brethren. Cases like hoofs : hooves present no difficulty under the modified second grouping-requirement because the morphs involved (hoof and hoove-, /s/ and /z/) are in non-contrastive distribution: hoofs and hooves do not differ in meaning. This is not true of brothers and brethren.

There are several possible ways of handling the problem. One way, which fits both versions of the second grouping-requirement but seems not too pleasing, is to group the /z/ of brothers, along with most other morphs meaning 'noun plural', into a single morpheme, but to exclude from this morpheme the morph found in brethren, both because it is not in free alternation with /z/ in this environment and because it has a different meaning: 'plural, with semantic specialization, producing a form of address for fellow lodge- or church-members of the male sex'.

Another solution is to postulate two distinct, though homophonous, morphemes brother: brother_1 'male child of same parents', and brother_2 'fellow lodge- or church-member of the male sex'. The plural of brother_1 is brothers; that of brother_2 is either brothers or brethren, in free alternation. The morphs meaning 'plural of noun' in these two cases, together with others of the same meaning, are now in non-contrastive distribution, and can be grouped into a single morpheme.

This breakdown of brother into two homophonous morphs, in order to achieve a greater differentiation of the environments in which various morphs meaning 'noun plural' occur, may seem artificial; but if one starts with the plural forms and then works to the singular, it seems less so. For brethren occurs in larger environments in which brothers also occurs, whereas brothers occurs in some larger environments in which brethren does not occur, e. g. I want you to meet my ( ), John and Bill. If one groups the cases in which both may occur, and contrasts the non-linguistic environment of these cases with that of the cases in which only brothers occurs, the semantic difference is fairly clear. Extension by analogy to the singular forms then seems justified. The source of difficulty here, as often, lies of course in the complexity of manipulating any type of semantic criterion.

12. The proposed revision of the second grouping-requirement leads rather clearly to simpler tactics; but it raises a problem for which I have no answer.

There is a generally accepted working assumption in descriptive analysis to the effect that while there may be homophonous morphemes, there are no exactly synonymous ones.[23] No matter how subtle the difference in meaning may be between, say, twenty and score,[24] the difference in phonemic shape implies non-identity morphemically. Now the revised version of the second grouping-requirement implies that we will violate this working assumption when the evidence leads us to believe that the violation is desirable. But I can make no statement as to the formal conditions under which the principle should be suspended. In every case of not quite complete complementation, we have to examine the positions in which more than one morph of the set appears, and decide whether in these positions they are in free alternation or not. In every case, this decision seems to turn on semantic considerations. If this is true, then for a long time to come such decisions are going to be partly a matter of individual taste. This need not deter us; for in any such case we need only suspend judgment, state both or all the alternative analyses, and indicate that our choice of one for further analytical purposes is only tentative.

13. The modification of Harris's third grouping-requirement that we propose is somewhat simpler; cf. §3 (2). Instead of requiring that the morphs to be grouped 'have combined environments no greater than the environment of some single alternant in the language', we require that they have a total range which is not unique. The range of a morpheme is the class of all environments in which the member morphs of that morpheme occur. Our revised requirement still stipulates that a morpheme obtained by grouping several morphs together shall have a range identical with (or paralleling) that of some other morpheme, but no longer requires that the second morpheme (the test morpheme) shall consist of a single morph.

The tactical advantage to be gained by either form of the requirement is that we thereby avoid the need to list separately the ranges of individual morphemes; we prefer to handle them in terms of classes having identical or closely parallel ranges. With either form

23 Bloomfield, Language §9.5.

24 Harris, Morpheme Alternants §2.2.

of the requirement, there may remain morphemes containing only a single morph, which have unique ranges; but this we cannot handle in the present connection.[25] For those morphs which perhaps can be grouped into complex morphemes, the tactical advantage is worthwhile.

The revised form of the requirement enables us to gain this tactical advantage in cases where it is impossible under the older form. Latin case-endings are a clear example. Since no single case-number category is represented after all noun stems by the same morph, it is impossible under the old form of the requirement to group all the morphs of any single case-number combination into a single morpheme. Under the new requirement, we may do so for one case-number combination, providing we also do so for at least one other case-number combination; the natural conclusion is to do so for every such combination. The set of eight or ten case-number morphemes can now be handled tactically as a class: they occur after noun stems; and a noun stem occurs before a case-number morpheme.

IV. Morph and Morpheme

14. We now attempt to remove the source of the remaining difficulties mentioned in §3 (3–5). This we do by a single rather fundamental alteration of the relationship between morph and morpheme.

Both before and after this alteration, an utterance consists wholly of morphs: every bit of phonemic material in an utterance is part of one morph or another. Before the alteration, every morph belongs to one and only one morpheme, so that there are as many morphemes in an utterance as there are morphs. After the alteration, the number of morphs in an utterance and the number of morphemes therein may not be identical: some of the morphs, and hence some bits of phonemic material, of some utterances, are morphemically irrelevant. How this change is brought about, and with what utility, will be demonstrated presently. In making it, we must conform to a principle which Harris does not state but which he adheres to rigorously: the principle of total accountability. Every morph, and every bit of phonemic material, must be determined by (i.e. predictable from) the morphemes and the tagmemes (if any) of which the utterance is composed.[26]

Two morphemic analyses of an utterance are tactically equivalent if they give the structure of the utterance in terms of the same morphemes and tagmemes—whatever the differences in the handling of submorphemic matters. For example, according to one analysis, Fox[27] poonimeewa 'he stops talking to him' consists of the morphs pooni 'cease', m 'act by speech on an animate object', and certain succeeding elements which do not concern us. A different analysis breaks the form into poon and im, with the same meanings. These two analyses are tactically equivalent. By the first one, the morpheme 'cease' has form poon before morphs beginning with a vowel, form pooni before those beginning with a consonant, and the morpheme 'act by speech on an animate object' has everywhere form m. By the second analysis, 'cease' has everywhere form poon, and the second morpheme has the two forms m and im, depending on what precedes. In either case, the sequence of morphemes involved can be indicated as ⟨poon⟩ + ⟨m⟩; it is only below the tactical level that there is any difference.

If, on the other hand, we divide the given form into poon 'cease', i, and m 'act by speech etc.', and consider each of these a morpheme as well as a morph, the analysis will not be tactically equivalent to the first two. For in this case the sequence of morphemes must be indicated as ⟨poon⟩ + ⟨i⟩ + ⟨m⟩—there are more morphemes in the word than by the first two analyses. It is easy to see also why this analysis is tactically inferior to the first two: the statement as to the occurrence of the morpheme i—to which no meaning can be assigned—will have to operate in terms of submorphemic (phonemic) properties of environments, whereas on the tactical level we should like to be able to state environments of occurrence and non-occurrence of classes of morphemes in terms of other classes of morphemes, without regard to submorphemic matters.

But can we find any valid basis for preferring the first of the above alternative treatments to the second, or vice versa? Clearly, there can be no tactical reason for choosing any one of two or more tactically equivalent analyses. If any reason at all is discoverable, it will be within the submorphemic realm: a matter of patterning, or perhaps simply of greater convenience. And although convenience is a legitimate basis for a choice, we must recognize such a criterion as different in kind from others, and as more open to disagreement. A more convenient analysis tells us nothing more about a language than a less convenient one that is otherwise equivalent; but what it does tell it tells more clearly.

25 In his unpublished material [see now his Methods in Structural Linguistics 202 (Univ. of Chicago Press, 1951)] Harris shows how this can be handled. His example is English /tuw/ (to, two, too), which in the absence of semantic criteria first appears as a single morph.

26 No defect of many older grammars of less-well-known languages is more marked than the confusion, or at best fuzziness, which results from a neglect of the principle of total accountability. Of course we do not condemn their writers for being 'men of their times rather than of ours'; for one thing, this doctrine could hardly be stated explicitly until the phonemic principle had been discovered.

27 I choose Fox rather than Menomini because the examples are a bit easier to cite; the same principles apply. The Fox forms are from Leonard Bloomfield, Notes on the Fox Language, IJAL 3.219–32, 4.181–226 (1924–7), and from the same writer's Algonquian, Linguistic Structures of Native America 85–129 (New York, 1946).

15. The alteration by which the number of morphemes in an utterance fails in some cases to coincide with the number of morphs consists of recognizing two special kinds of morphs: empty morphs, which have no meaning and belong to no morpheme; and portmanteau morphs, which belong simultaneously to two (or, theoretically, more) morphemes, and have simultaneously the meanings of both.

If for some submorphemic reason (patterning or convenience), the breakup of Fox poonimeewa into poon+i+m (+) is to be preferred to either of the two alternative procedures outlined in §14, this breakup can be made tactically equivalent to the latter two, rather than to the analysis which requires the occurrence of i to be taken care of on a tactical level, by calling i an empty morph. Total accountability is maintained because we say, on the submorphemic level, that when a morph ending in a consonant is followed in the same word by one beginning with a consonant, the empty morph i appears between them.

The simplest example of a portmanteau morph is French /o/ 'to the (masc.)' (§3). If this be taken as a single morpheme, tactical difficulties ensue. What other morpheme has a range of positions of occurrence parallel to the range of this one? On the other hand, since /o/ is a single phoneme, it is hardly possible to make a cut and produce two morphs. But if we interpret it as a portmanteau morph, the representative of the morpheme sequence {à le}, we not only eliminate a forlorn morpheme, but round out the distribution of {à} and of {le}, both otherwise somewhat defective. For à 'to' parallels to a great extent the distribution of sur 'on', après 'after', and other morphs, but—unless the proposed interpretation is accepted—does not occur in one important position where the others occur: before le 'the (masc.)' when the following noun begins with a consonant. Similarly, the suggested treatment of /o/ makes the parallelism between le and la 'the (fem.)' much neater. The case is so clear-cut that there is nothing remarkable in the fact that au has been traditionally so interpreted.

It is to be noted that our morphemic expansion of /o/ to {à le} involves not only the morphemes {à} and {le}, but also a specific order thereof: /o/ is not morphemically {le à}. This specific order, like the morphemes themselves, is given not by the portmanteau as such, but by its distribution and that of the morphs to which we propose to relate it. Choice of the order {à le} leads to the parallelism indicated above; choice of the reverse order leads to nothing at all.

16. The simple examples just given speak, it is believed, for the naturalness of this approach; but as yet we have given no formal statement of the conditions under which an empty morph or a portmanteau morph is to be set up.

Because of the possible importance of submorphemic patterning, it will be necessary to consider the typical phonemic shapes, or canonical forms, of morphs. It is a well-recognized fact that in any particular language, if we examine and classify those cases of morphs which do not patently involve the questions here being raised, we find that many different morphs have much the same general phonemic shape.[28] Fijian affords an elementary example.[29] A large number of morphs have the shape $\#C_1V_1C_2V_2(C_3)$, where # is word juncture, the C's indicate any consonant (or none), the V's any vowel, and C_3 is lacking when word juncture or a consonant follows: koro 'village' (C_3 = zero), sala 'path', dina 'true, truth', selev 'cut, knife', ðabet 'go up', kaðiv 'call, announce'. A second, much smaller, group have the shape $\#C_1V_1C_2V_2\,C_3V_3(C_4)$: tanane 'man, male', yalewa 'woman, female'. A third group have the shape V or V_1CV_2: a 'transitive with common object', as in ðabeta 'go up (a hill)', raiða 'to see (a child, etc.)'; i 'transitive with proper object', as in raiði 'to see (John, me, etc.)'; aka 'transitive indirective', as in ðabetaka 'carry (someone) upwards'. Lastly, there is a group of structure #CV#, occasionally $\#C_1V_1C_2V_2\#$: na 'the (common)', as in na koro 'the village'; ni 'of the (common)', as in na yaða ni koro 'the name of the village'; i 'that connected with the act of', as in na i sele 'the knife'; ko 'the (proper)', as in ko viti 'Fiji' or ko ðei 'who?'.

In some languages the variety of canonical forms is far greater than in Fijian, but in every language the total number—however assessed, for there is some choice in the process of abstraction from specific phonemes to symbols like C and V—is relatively small. In English many morphs have the shape of a single syllable with # preceding[30] (girl, act); others consist of a single consonant, or of a single syllable with initial vowel, with no preceding # (-s, -ing, -ed, -or). In both Fijian and English, and probably generally, some canonical forms can be expressed as the 'sum' of certain smaller ones: Fijian #CVCVCV(C) as #CVCV(C) plus V; the English type of author as the type of act (or watch) plus the type of -or (or -ing). Moreover, in these cases the 'sums' occur as sequences of several morphs (actor) as well as in single morphs (author). Those canonical forms which cannot be so expressed may conveniently be called minimum.

28 A point discussed in detail by Benjamin L. Whorf in various unpublished material, and orally.

29 C. Maxwell Churchward, A New Fijian Grammar (1941).

30 This avoids the risky complications which result from calling word-juncture a morpheme, as Rulon S. Wells does in his Immediate Constituents §64 (see fn. 4). The semantic contrast between Thank you with word juncture and the same without it means that word juncture is morphemic, but in such cases it might just as well be concluded—I think, a little better so—that absence of word juncture is the morpheme.

17. If in analyzing the morphemics of a language we make a preliminary classification of canonical forms, based only on those morphs whose status is perfectly clear, this classification can serve as a guide in handling the less obvious cases.

Multiplicity of analytical choice turns on two things: the location and number of cuts to be made in certain utterances; and the classification of the resultant morphs as ordinary, empty, or portmanteau. When faced with alternatives, we base our decision, wherever possible, on the relative desirability of the resulting tactics. It is on this account that the treatment of French /o/ 'to the (masc.)' as a single morpheme, or of Fox connective i as a morpheme, is rejected. When this factor cannot play a part, we turn next to morphophonemic simplicity. Morphemes of constant phonemic shape are simplest; when we cannot find these, we look next for sets of morphemes showing similar alternations in phonemic shape, since then we can describe the alternations of many different morphemes at once. When this also is not decisive, we turn to canonical forms, and prefer that analysis which produces morphs most closely conforming to the canonical forms already established—if possible, to minimum canonical forms. It may be that the second and third of these considerations should be assigned the other order of priority; apparently they are not often in conflict.

When we are confronted with three tactically equivalent alternatives for Fox poonimeewa (poonim, poonim, and poonimm with i as an empty morph), we need only proceed to the second consideration to reach an answer. If we make either the first or the second choice, one of the morphemes involved will have two alternants (poon and pooni, or else im and m). If we make the third, both poon and m become morphemes of constant phonemic shape. If this were not enough, the third criterion would show us that failure to set up i as an empty morph would force us to recognize some morphs, beginning or ending in i, of canonical forms not otherwise required (though not in the case of the elements in the particular word poonimeewa), whereas the decision to set up the i by itself produces only morphs of shapes necessary anyway.

Likewise in the French case: tactical considerations rule out a monomorphemic interpretation of /o/, but do not decide whether we must take it as a single (portmanteau) morph or may cut it further. Now by a criterion mentioned in §2 under Step 1, a morph must have overt phonemic content. In order to cut /o/ into two morphs, we must break up the phoneme /o/ into two components, say mid-back tongue position and lip-rounding. Neither of these components fits into any otherwise necessary canonical form of morphs in French (though in some other languages some morphs do have a shape definable in components rather than in phonemes). On the other hand, /o/ taken as a single morph fits into a canonical form, represented by such clear cases as /o/ 'water', /e/ 'and', /u/ 'or', /a/ 'to'. The vote is clearly for the interpretation as a portmanteau morph.

In the succeeding sections of Part IV we give further examples in which these same principles call for the recognition of empty or portmanteau morphs.

18. In Nootka[31] a word consists—submorphemically, as we shall show—of a stem plus one or more suffixes. (This statement is circular: a stem is a morpheme which begins with word-juncture, a suffix one which does not; but it will suffice for orientation.) Certain suffixes, which we may call location suffixes, occur both after ordinary stems and after a stem hina-, hin-, hita-,[32] which Sapir and Swadesh label an empty stem.

Thus ƛ̓iḥ- 'red' + -(q)oˀ(ɬ) = ƛ̓iḥo·ɬ 'red on the face'; sixʷ- 'sores, pox' + the same suffix = sixo·ɬ 'having sores on the face'; but hina- + the same suffix = hino·ɬ 'on the face, being on the face'. Similarly, ƛ̓iḥ- 'red' + -(ʔ)akso(ɬ) = ƛ̓iḥaksoɬ 'red at the lips'; hap- 'hair, fur' + the same suffix = hapaksoɬ 'having a moustache'; but hina- + the same suffix = hinaksoɬ 'at the lips or mouth, being at the lips or mouth'. Finally, maƛ̓- 'tied' + -aḥo·p 'cause momentarily to be in front' = maƛ̓aḥo·p 'ties in front'; but hina- + the same suffix = hinaḥo·p 'places in front'.

The empty stem has no meaning. Our tactics are just as well suited, and our morphophonemics are not complicated, by interpreting each form of the empty stem as an empty morph. The remaining stems constitute a class of morphs which begin with #. Suffixes other than the location suffixes constitute a class of morphs which do not begin with #. When a location suffix is preceded by a stem, neither # nor any other phonemic material intervenes. When a location suffix is not preceded by a stem, it is preceded instead by the appropriate (predictable) form #hina-, #hin-, #hita-, which in any case is meaningless and tactically irrelevant. The principle of total accountability is not violated; and the empty morphs conform to canonical forms. The alternative of taking #hina- etc. as a stem (a morpheme) is undesirable because of the meaninglessness of this element. The alternative of taking hinaḥo·p as an alternant going with -aḥo·p, and similarly for every other such combination, produces a greater complication of canonical forms.

31 Edward Sapir and Morris Swadesh, Nootka Texts (Philadelphia, 1939), esp. Part III, The Primary Structural Elements of Nootka; Morris Swadesh, Nootka Internal Syntax [sic], IJAL 9.77–102 (1936–8). The specific examples were generously supplied by Swadesh.

32 The alternation among these three shapes of the element will not concern us; it is covered by statements on a lower level of morphophonemic treatment.

19. In most of the central Algonquian languages occurs a phenomenon which we shall here illustrate with Potawatomi examples. Nouns appear in both unallocated and allocated forms: wUkUma 'chief' : nUtOkUmam 'my chief' or kUtOkUmamwa 'your (pl.) chief'. Some nouns, however, appear only in allocated forms; so nos· 'my father', kos· 'your (sg.) father'. A noun in an allocated form contains a personal prefix before the noun stem, and after the noun stem one or more of several suffixes (including, in the allocation is plural, a personal suffix) and various inflectional endings. Some of these nouns which occur in both unallocated and allocated forms contain, after the noun stem in an allocated form, a morph m, Um, im, om; so, for example, the forms for 'my chief' and 'your chief' above. Other such nouns appear sometimes with a morph of this shape, sometimes without it, but with no semantic contrast —the presence and absence of the morph are in free alternation. Still others, including all nouns which appear only in allocated forms, never occur with the m element.

The m elements are more satisfactorily regarded as morphs than as parts of the preceding morphs, particularly since the choice among the various forms of the m element depends on the environment in much the same way as does the choice among the alternants of the morpheme ⟨k⟩ 'locative' and (though with less similarity) among the various alternants of a number of other suffix morphemes. But the m elements are meaningless, even where forms appear both with and without one of them, and it is tactically convenient to eliminate them from the picture before tactical discussion begins. So we take them to be empty morphs.

20. The English interjections written conventionally as hm!, eh?, and the like, consist phonemically of an intonation-sequence, a stress, and a segmental 'carrier' for these features. In my dialect, this segmental component may have any vocalic quality (whether this occurs elsewhere or not), or any oral closure or closures, but it must be nasalized. Such a segmental structure is atypical in its wide range of nondistinctive variation, but the articulatory feature involved distinctively—nasalization—is one which does recur in more typical segmental structure, for example as that which distinguishes /m/ from /b/, /n/ from /d/.

If we compare the utterance hm (intonation 32)[33] with yes(32) and with hm(24), we see that the meaning of hm(32) is that of yes(32) minus the meaning of the intonationless abstraction yes; between hm(32) and hm(24) there is no semantic similarity. Hm itself, then, apart from the intonation which it serves to carry, has no meaning at all.

[33] Following Pike (The Intonation of American English; Ann Arbor, 1946), and Wells (Immediate Constituents §79) in the assignment of figures, and numbering the four levels from top down.

We may conclude that the hm part of such interjections is an empty morph. The intonational morphs which accompany it are also found spread through such morphs or morph-sequences as yes, I know, maybe, he didn't come, and so on. Accountability is maintained: if an utterance consists morphemically of an intonational morpheme alone, the empty morph hm will be present; otherwise hm is absent.

The tactical implications are interesting: the only free morphemes of English, in Bloomfield's sense of 'free', are intonational morphemes, and the only monomorphemic utterances of the language are those consisting of such a morpheme.

21. In certain Spanish verb forms there appears, between the stem and the endings, an element often called a conjugation vowel: the á, é, and í of amar 'to love', beber 'to drink', vivir 'to live'; the áb, í, and í of amábamos 'we loved', bebíamos 'we drank', vivíamos 'we lived', etc. Which vowel (or in one case vowel plus consonant) appears, depends on the stem and on the ending: the infinitive ending r, for example, requires á after a stem of the first conjugation, é after one of the second, and í after one of the third. The three conjugations are classes of stems, in fact, based precisely on this feature of behavior.

The conjugation vowels have no meaning. The meaning of amar is that of the two component morphemes, stem am 'love' and the infinitive ending, whether we treat the latter as ár in this case, alternating with ér and ír elsewhere, or simply as r. The latter alternative relegates the á, and all other such conjugation vowels, to the status of empty morphs.

Not all the post-stem vowels which occur in Spanish verbs have this status. The a of amas 'thou lovest', for example, is the only thing which distinguishes this form from ames 'that thou love' (subjunctive). Here the a is no empty morph, but an ordinary morph with meaning 'present indicative'. In one possible analysis of Spanish verbs, which would perhaps be the simplest from the morphophonemic standpoint, distinctions such as that between a 'present indicative' and the meaningless á of amar are not made. But this somewhat greater morphophonemic ease is outweighed by more complicated tactics.

22. Our first additional example of a portmanteau morph comes from Yokuts.[34] In the Yawelmani dialect there are about a dozen morphemes which occur after a verb stem and before a finite or gerundial suffix. One of these is WZa·la· (with alternants WZla·, FRla·, WW'e·, WLe·, FZWZla·, ila·, la·, WSla·, variously apportioned among different types of preceding element, but in non-contrastive distribution) 'cause someone to x': tisa·la·hin '(he) caused (it) to come out, (he) took (it) out', with stem tisi 'come out' and finite suffix hin 'aorist'. Another is WAda· (with variants da·, R = reduplication) 'x often or repeatedly': sodoxdo? 'will throw him repeatedly', with stem sodox 'throw' and final suffix ? 'future'.

[34] See fn. 11. The capital letters at the beginning of cited suffixes are components of the vowels in the part of the word which precedes the specific phonemes of the suffix. Thus the stem me·k'i 'swallow' contains two consonants and parts of two vowels: after the first consonant, the vowel components high-front and long, and after the second consonant, the vowel components high-front and short. When this stem occurs with the suffix FRit 'passive aorist', the component F merges with the first group of vowel components in the stem to give e·, and the component R merges with the second group of vowel components in the stem to give zero; the resulting form is me·k'it 'was swallowed'. With a different set of components contributed by the suffix WA?an 'durative present', the resulting form is mik'a·?an 'is swallowing'. For the details of this, see Zellig S. Harris, Yokuts Structure and Newman's Grammar, IJAL 10.196–211 (1944).

In some cases two of these elements occur in succession, within the position of occurrence stated above. Indeed, the alternant R of the morpheme 'x repeatedly' occurs followed by the alternant FZWZla· of the morpheme 'cause to x': muhmuhlat 'was made to dive repeatedly', with stem muhu 'dive' and final suffix t 'passive aorist'. But other alternants of the morpheme 'x repeatedly' do not occur before any alternant of the morpheme 'cause to x'. No semantic gap results, however, for there is an element WE·lsa· ∿ WE·sa· of the same positional class, meaning 'cause to x repeatedly': nine·lsa·hin 'got him to keep still several times', with stem nine· 'keep still, become quiet', and final suffix hin 'aorist'.

It is far from convenient, within the morphophonemic economy of Yokuts, to cut WE·lsa· ∿ WE·sa· into smaller morphs; each alternant subsumed by this notation is best taken as a single morph. The distribution and meaning lead one to interpret each of these morphs as a portmanteau representative of the sequence of two morphemes 'x repeatedly' + 'cause to x'.

23. In finite forms of the Spanish verb the tense-mode is usually indicated by one morph, and the person-number by another, in that order: amáb|a|mos 'we loved', ama|ré|is 'you (pl.) will love'. In a few cases, it is difficult or impossible to separate the element meaning a tense-mode from that meaning a person-number; in these cases, we may regard the undivided endings (after any conjugation vowel that may occur) as portmanteau morphs: o 'present indicative + first person singular', as in amo 'I love', é (with verbs of the first conjugation) ∿ í (with those of the second and third conjugations) 'preterit indicative + first person singular', and ó (first conj.) ∿ ió (second and third conj.) 'preterit indicative + third person singular'. This treatment, combined with the empty-morph interpretation of conjugation vowels, reduces all finite Spanish verb forms to a uniform structure: stem + tense-mode morpheme + person-number morpheme.[35]

24. In Fijian there is a construction consisting of any of certain particles followed by a noun or pronoun: na sala 'the path', na ðava '(the) what?', ni koro 'of-the village', ko viti '(the) Fiji', vei au 'to me'. One of the particles is vei 'to, with, of', as in the last example above and in vei keda 'with, to, of us', vei Joni 'with, to John'. One of the pronouns is koya 'he, she', as in ko koya '(the) he, she', nei koya 'of him, of her'. But the specific combination of vei and koya seems not to occur.[36] Where semantically it would be expected, one finds, instead, the portmanteau morph vuaa 'with, to, of him or her', as in au na vosa vuaa na ŋone 'I future speak-to-him the child' = 'I shall speak to the child'.

25. We may best approach a consideration of the fifth difficulty of §3 by examining some English cases. On the tactical level, it is certainly desirable to consider men as consisting of the morpheme man plus the morpheme s 'plural'. When cutting utterances containing men into morphs, we will not be led to men into smaller pieces; it fits a canonical form and if broken further the smaller fragments do not. So one solution, and certainly the most obvious one, is to regard men as a single portmanteau morph, representing the morpheme sequence ⟨man⟩ + ⟨s⟩.

It is true that there is a phonemic similarity between man and men—the identity of initial and final consonants—which we do not want to lose sight of. This places men in a different category from French /o/, Yokuts WE·(l)sa·, Spanish o, or Fijian vuaa, for in the latter cases the resemblance of the portmanteau to other alternants of either of the constituent morphemes is negligible. Even if men were an isolated case in English, this resemblance would be worthy of mention. But it is, of course, far from isolated; we have also mouse : mice, foot : feet, woman : women (if woman is a single morpheme), slide : slid, sing : sang, and many others.

The portmanteau interpretation of such bimorphemic forms need not obscure the phonemic resemblance of which we are speaking. In our morphophonemics we have to mention all portmanteaus. By assembling, in one section of our description, all portmanteaus which have this feature of partial phonetic identity with one of the constituent morphemes, and by organizing them into groups on the basis of the specific phonemic difference, we give ample attention to the matter.

Some may nevertheless prefer to reinterpret portmanteaus as bimorphic as well as bimorphemic, even though to do so one must extend the definition

[35] The 'irregular' verbs present more complex cases of both portmanteau and empty morphs, but are tactically quite the same, save where one or another form is missing.

[36] Churchward is not entirely clear on the matter: he says (op.cit. I.24.3) that vei koya is 'unusual'. If it does indeed occur, then the interpretation proposed is wrong; rather vei koya is like English with it and vuaa like therewith.

of 'morph' to cover elements of other than overt phonemic content. If this is considered desirable, then in the notion of portmanteau we have at least achieved a more rigorous way of extending the coverage of the term 'morph' in such a manner, as follows:

In our initial cutting of utterances, we obtain only morphs of overt phonemic content. Further examination, along the lines detailed in this paper so far, reveals the possibility that certain of our morphs are portmanteaus; but for our present purpose we may call them rather tentative portmanteaus. We then examine each tentative portmanteau and compare its phonemic shape with that of the other alternants of the constituent morphemes. If we find that the tentative portmanteau has some phonemes (or components) in common with one of the non-portmanteau alternants of one of the constituent morphemes, we may set up the entire non-portmanteau alternant as one constituent morph of the form which has tentatively been regarded as a portmanteau, and the alternation from this shape to that of the tentative portmanteau as the other constituent morph. Alternatively, we may set up the tentative portmanteau, as a whole, as an alternant of that constituent morpheme which it resembles phonemically, and set up a zero morph as an alternant of the other constituent morpheme.[37] For example, our initial cutting produces men, which does not look like more than one morph. The sequence man plus s does not occur. Men fills the tactical place which one might expect to be filled by the sequence man plus s. Men is therefore morphemically {man} + {s}. But—so runs the argument that would set up alternation morphs—men and man resemble each other in phonemic shape, both containing m-n. So men is not a portmanteau. One morph in men is man. The other is the alternation a∿e. Or—arguing now for a zero morph—men is not a portmanteau, but consists of an alternant men of {man} plus an alternant /0/ of {s}.

If a language contains only a few isolated instances of this kind, probably everyone would agree to reject the last steps of the above argument and return to the portmanteau interpretation, relying on the organization of one's morphophonemic statements to put the matter of partial phonemic resemblance into clear relief. But if the language contains a sufficient number of such cases that one is warranted in setting up a canonical form for morphs like a∿e, or like /0/, then some may prefer the extension.

Somewhat similar considerations apply to French bon 'good (masc.)' and to English sheep (pl.). In the course of examination, the portmanteau interpretation is that which first presents itself; from it we may proceed to the recognition of morphs of other than overt phonemic content if we find factors comparable to those in the case of men. It is to be emphasized that when portmanteaus are eliminated in this way, the new definition of 'morph' is no longer that with which we began; perhaps, therefore, it would be advisable to distinguish terminologically between, say, 'primary morphs' (those of overt phonemic content) and 'extended morphs' (including primary ones and morphs of the zero, replacement, or subtraction types).

English children, however, remains recalcitrant. Obviously it is morphemically {child} + {s}; so that whatever submorphemic interpretation we chose, the tactical picture is clear. The first part of children resembles child, and the last part is identical with one of the alternants of {s}, namely the -en of oxen. The alternative analyses are (1) child|ren, (2) childr|en, (3) child|r|en, (4) child + vowel change and -ren, and (5) no cut, i.e. portmanteau. The first gives a morph /čild/, the difference between which and child recurs in other contexts, e.g. slide : slid, bite : bit; but then the morph -ren is unique. The second gives a morph -en which recurs; but then the difference between /čildr/ and child is unique. The third has the merits of each of the first two, without the defects, but involves an empty morph r, which is not observed to recur and therefore requires a special statement for this occurrence. The fourth produces a morph (vowel change and -ren) which fits no canonical form, unless the vowel-change-plus-en of bitten, hidden, and others is grouped with it from the point of view of shape. Apparently this is one of the cases in which all our preferential criteria (§14) fail, and nothing remains but a resort to convenience.[38]

[37] This second alternative is that proposed by Bernard Bloch, English Verb Inflection, Lang. 23.399–418 (1947). Bloch rejects all alternation or subtraction morphs, and interprets all tentative portmanteaus as an alternant of one of the constituent morphemes plus a zero alternant of the other. One special criterion is introduced for dealing with zero alternants: no morpheme is postulated which has only a zero alternant.

[38] The unsolved case of children is discussed in detail for a reason. There is no merit in an analytical procedure which 'eliminates' all but one of a set of alternative analyses simply by fiat—by saying that when such-and-such types of alternative present themselves we shall accept the one which has certain characteristics and reject the others. Our aim is to achieve the most accurate and clearest picture possible of the workings of a language, on all levels —phonemic, morphemic, and tactical; in some cases this is attained not by giving a single treatment, but precisely by indicating the alternatives. For in some cases a range of choice is determined not by our approach, but by the nature of the language; and when this is so, the existence of a range of choice in a particular portion of the language is one of the facts about the language that ought to be portrayed in our description. In one sense, any method of description which conforms to the principle of total accountability is correct; if we nevertheless discuss the relative merits of one procedure or another within this fundamental framework, the purpose is to attain greater mutual intelligibility among the writers of grammars and, in terms thereof, more accurate pictures of the languages we describe.

V. Conclusions

26. We now summarize the procedure of morphemic analysis worked out in the course of our discussion, and end with an example from English which illustrates several of the points that have been made. Our summary of the procedure is given in steps, as in §2; but in actually working with a particular language one has to skip back and forth, operating by trial and error.

Step 1. We assemble the utterances of the language before us, recorded in some phonemic notation. If preliminary examination reveals that a different (also phonemic) notation would make the task simpler, we retranscribe them all. If further preliminary examination shows that some normalization of notation, maintaining all phonemic distinctions but adding thereto, would further simplify the task, we retranscribe again, and perhaps again. As we proceed to other steps, we check back from time to time to be sure we have not involved ourselves in contradictions.

Step 2. The utterances are now examined in the notation finally chosen. Recurrent partials with constant meaning are discovered; those not composed of smaller ones are morphs. So are any partials not recurrent but left over when all recurrent ones are accounted for; therefore every bit of phonemic material belongs to one morph or another.[39] By definition, a morph has the same phonemic shape in all its occurrences; and (at this stage) every morph has an overt phonemic shape, but a morph is not necessarily composed of a continuous uninterrupted stretch of phonemes. The line between two contiguous morphs is a cut.

Step 3. Omitting doubtful cases, morphs are classed on the basis of shape and the canonical forms are tentatively determined.

Step 4. Two or more morphs are grouped into a single morpheme if they fit the following grouping-requirements: (a) they have the same meaning; (b) they are in non-contrastive distribution; (c) the range of the resultant morpheme is not unique. Some morphs, however, may be assigned to no morpheme at all, and some may be assigned simultaneously to two (or more) morphemes. An empty morph, assigned to no morpheme,[40] must have no meaning, and must be predictable in terms of non-empty morphs. A portmanteau morph must have the meanings of two or more morphemes simultaneously, and must be in non-contrastive distribution with the combination of any alternant of one of the member morphemes and any alternant of the other (usually because no such combination occurs).

Step 5. Where there are alternative possibilities, choice is based on (a) tactical simplicity, (b) morphophonemic simplicity, and (c) conformity to canonical forms, in this order of priority.

Step 6. The differences in the phonemic shape of morphs as alternants of morphemes are organized and stated; this (in some cases already partly accomplished in Step 1) constitutes morphophonemics. In particular, portmanteaus are compared with the other alternants of the morphemes involved, and if resemblances in phonemic shape and the number of cases warrant it, morphs of other than over phonemic content are recognized, some of the portmanteaus being thus eliminated.

27. Our final example is the system of personal pronouns in English (including who, whom, whose).

At least in certain dialects, the morphs I and me (and similarly we and us, he and him, etc.) are in non-contrastive distribution; in some dialects, indeed, the complementation is probably complete. We may suspect that if it were not for the Latinizing school tradition, the complementation would be complete for most speakers: I initially except in isolation, me directly after a verb or a preposition and in isolation. Actual exceptions to this are either on the Latin pattern (It's I, or Who's there?—I, instead of Me), or are overcorrections (between you and I). For many speakers whose usage of I and me does not put them in complete complementation, there is no contrast between, for example, It's I and It's me. In other dialects and styles, on the other hand, the forms are in contrast: literary English, schoolteachers' on-duty English, and certain whimsical styles.[41] The remainder of this discussion applies only to a dialect in which the distribution is non-contrastive.

My and mine (and similarly our and ours, your and yours, etc.) are in complete complementation: my occurs when a noun follows without pause, mine otherwise.

If the above statements are to hold, we must split the occurrences of her into those which parallel those of his and those which parallel those of him; the former, her_1, is morphemically identical with hers, while her_2 is morphemically identical with she.

Paralleling John in John came, Bill saw John, John's book, the book is John's, and virtually every other utterance containing the morpheme John, we have I came, Bill saw me, my book, The book is mine, etc. John's is two morphs and two morphemes; we conclude that my and mine are two morphemes each, though each is only a single morph.

[39] We say 'phonemic' for simplicity's sake; if our notation has been normalized, then more accurately this should read 'every bit of orthographic material'.

[40] All the empty morphs in a language are in complementary distribution and have the same meaning (none). They could, if there were any advantage in it, be grouped into a single empty morpheme—but one which had the unique characteristic of being tactically irrelevant.

[41] For example, that style in which one says 'me, myself, and I' as if the reference were to three people. This is not unrelated to a style which obviously has to be excluded, both here in the discussion of English pronouns and in any other discussion of morpheme alternants: the style of the discussion itself, in which such forms as me and I contrast because they are are used as names of particular morphs.

We conclude, therefore, that the English personal pronouns have the following morphemic structure:

{I} I, me	{she} she, her$_2$
{I} + {s} my, mine	{she} + {s} her$_1$, hers
{we} we, us	{it} it
{we} + {s} our, ours	{it} + {s} its
{you} you	{they} they, them
{you} + {s} your, yours	{they} + {s} their, theirs
{he} he, him	{who} who, whom
{he} + {s} his	{who} + {s} whose

The forms it, its, and whose are the same morphically and morphemically; the others illustrate one or more of the grouping-requirements that we have discussed. Together, the twenty-six forms are analyzed into only nine different morphemes.[42]

The tactical implications are considerable. Except for the category of number, the pronouns are now exactly like any proper noun in their tactics, and can be classed as a subdivision of proper nouns. There is no longer any justification for speaking of case in English; for the distinction between subjective and objective 'cases' (under whatever name) disappears as soon as I and me, etc., are shown to belong to the same morpheme. A form with added -'s is not a case-form either, but simply a form with added -'s: the -'s is simply another morpheme, with a statable range of positions in which it occurs.

42 We might go further, interpreting we, us as {I} + pluralizing {s}, with a similar treatment for the other plural pronouns. We are deterred from this step not because plural you is identical with singular you (since after all sheep and other nouns manifest this property), but because {he} + {s}, {she} + {s}, and {it} + {s} would all add up to they, them.

[[This re-codification of morphemic analysis, five years after Harris ((109)), in part represents what most workers including Harris had come to agree on, in part is original and presents notions that did not all find favor, though all of them did stimulate the professional readers to clarify their own thinking. Hockett himself, characteristically, treated the new ideas just as if someone else had propounded them; for instance, in his Peiping Morphophonemics of 1950 ((317)) he cheerfully left unused the notion of 'empty morphs' and instead spoke of morphemically irrelevant phonemic material. This sort of behavior is, as a matter of fact, quite usual among American linguistic theorists and has caused unsympathetic critics to speak of irresponsible zig-zagging. Rather than stigmatize it thus, I should prefer to call it a protection of one's own flexibility and would maintain that a free indulgence in, or even cultivation of, this habit has been one of the most favorable traits in the growth of descriptive linguistics among us. The opposite would have required replacement, not reinforcement, of the front-line fighters; and the new men would have been less effective replacements if they had felt it expedient to withhold theses that they might someday have to alter or replace.

This paper was composed at the same time as the following one by Bloch. They were both read at the same summer meeting of the Linguistic Society. And both authors considered it best to publish them with no adjustment to each other—for similar reasons.]]

ENGLISH VERB INFLECTION

BERNARD BLOCH

Language 23.399–418—1947

1. Introductory

1.1 The inflection of verbs in present-day colloquial English[1] has been described in many works—most clearly and exhaustively, perhaps, by Sweet, Palmer, Curme, Fries, Jespersen, and Hockett.[2] In view of the number and fullness of these descriptions, no new treatment can hope to add any facts hitherto overlooked: at most, a new treatment may be able to arrange the known facts more systematically than has been done before, or in a way that will be more useful to other linguists.

In all previous works, the inflection of English verbs is described in terms of the processes by which various inflected forms are derived from underlying bases. Thus, the preterit waited is said to be derived from wait by the addition of a suffix, took from take by vowel change, built from build by consonant change, sold from sell by vowel change and suffixation together, went from go by suppletion, put from put by zero change, and so on. Statements of this kind, referring to processes of derivation, are useful for showing the relation of any inflected form to its base; but they have at least one serious shortcoming: they cannot be readily used for the description of specific forms, viewed as words in their own right.

1.2 To describe the structure of a language as a whole, the linguist must be able to describe also the structure of any single sentence or part of a sentence that occurs in the language. He does this in terms of constructions—essentially, in terms of morphemes and their order.[3] Any sentence, phrase, or complex word can be described as consisting of such-and-such morphemes in such-and-such an order; each morpheme has a meaning, and so also has the order in which they occur (the 'constructional meaning').

For the purposes of this paper we adopt Bloomfield's definition of a morpheme, which has been accepted by nearly all descriptive linguists. A morpheme, according to Bloomfield, is 'a linguistic form which bears no partial phonetic-semantic resemblance to any other form'; a linguistic form is 'any combination of phonemes ... which has a meaning'.[4]

To illustrate: the preterit form waited—considered simply as a word, without the morphemes of stress and pitch that would accompany it in a real utterance—can be described as follows. It consists of two morphemes, /weyt/ and /ed/,[5] occurring in that order. The meaning of the first morpheme is a par-

1 The dialect here studied is a somewhat generalized northeastern variety of standard American English. For the methodological groundwork of this paper see Zellig S. Harris, Morpheme alternants in linguistic analysis, Lang. 18.169–80 (1942) ((109)); Rulon S. Wells, Immediate constituents, Lang. 23.81–117 (1947) ((186)); and C. F. Voegelin, A problem in morpheme alternants and their distribution, Lang. 23.245–54 (1947). Compare now also Charles F. Hockett, Problems of morphemic analysis, Lang. 23.321–43 (1947) ((229)). Though my views in general agree with Hockett's, it will be observed that his treatment of certain problems of English inflection (especially in §24 of his paper) differs markedly from the one here proposed.

I have profited from discussions with R. S. Wells, W. F. Twaddell, and Martin Joos.

2 Henry Sweet, A new English grammar logical and historical 1.391–428 = §§1283–493 (Oxford, 1892); Harold E. Palmer, A grammar of spoken English on a strictly phonetic basis 88–122 = §§199–270 (Cambridge, 1930); George O. Curme, A grammar of the English language, II. Parts of speech and accidence 241ff., esp. 269–96 = §60, 304–19 = §63 (Boston etc., 1935); Charles Carpenter Fries, American English grammar 59–71 (New York and London, 1940); Otto Jespersen, A modern English grammar on historical principles, VI. Morphology 28–83 = Chaps. 4–5 (Copenhagen, 1942); Charles F. Hockett, English verb inflection, SIL, Vol. 1, No. 2 (1942). For an entertaining travesty of American English verb inflection see H. L. Mencken, The American Language[4] 427–47 (New York, 1936).

3 Wells, Lang. 23.93–8.

4 Leonard Bloomfield, Language 161 and 138 (New York, 1933).

5 The phonemic transcriptions in this paper necessarily reflect my own speech, except that a few distinctions not commonly made in other dialects of English have been disregarded. On the transcription of vowels and diphthongs see George L. Trager and Bernard Bloch, The syllabic phonemes of English, Lang. 17.223–46 (1941), and cf. Lang. 19.189 fn. 15 (1943). The following stressed syllabics occur in the dialect here transcribed: /i/ in pit, /e/ in pet, /æ/ in pat, /a/ in pot, /ʌ/ in cut, /o/ in coffin, /u/ in put, /ə/ in just (adverb); /iy/ in beat, /ey/ in bait, /ay/ in bite, /oy/ in boil, /uy/ in ruin (monosyllabic); /aw/ in bout, /ow/ in boat, /uw/ in boot; /ih/ in theater, /eh/ in yeah, /æh/ in mad, /ah/ in calm, /oh/ in law, /əh/ in er (hesitation form); /ir/ in here, /er/ in there, /ar/ in part, /or/ in port, /ur/ in sure, /ər/ in curt; /ihr/ in beer, /ehr/ in bare, /ahr/ in bar, /ohr/ in bore, /uhr/ in boor, /əhr/ in burr. Pitch will not be marked in this paper; stress will be marked only in transcriptions of whole utterances. A space between words has no phonetic or phonemic significance.

Needless to say, the treatment of verb inflection offered here does not depend on the system of transcription. The cited forms could be written just as well according to any other system, so long as it recognized the existing phonemic distinctions in American English.

ticular action that we need not specifically describe here; that of the second is 'past time' or the like. The constructional meaning of the order in which the two morphemes occur is approximately 'perform a certain action at a certain time'.

How shall we describe, now, the preterit form took? Its relation to the uninflected form take is irrelevant, because we are concerned here simply with the structure of this one word, not with its derivation. Either took is one morpheme, or it is two morphemes; the possibility of its being more than two may be neglected as improbable. If it is one morpheme, it is either the same morpheme as take, or a different morpheme: as a morpheme, it cannot be partly the same and partly different. If it is the same morpheme as take, its meaning must be the same also; but of course we know that took and take are not synonymous. If took is a single morpheme different from take, then there can be no morphological connection between them—just as there is none between took and talk; but since the semantic and syntactic relation of took to take is exactly the same as that of waited to wait, we do not like to give up the possibility of connecting them in a morphological system. Finally, if took consists of two morphemes, what are they? Perhaps they are (1) /teyk/ and (2) vowel change; but if a morpheme is ultimately a combination of phonemes, then it is clear that vowel change, a process, is not a morpheme. Perhaps, instead, the two morphemes in took are (1) /t...k/ and (2) /u/; but then what about take? Does take then consist of the two morphemes /t...k/ and /ey/? If so, it differs in a fundamental respect from a verb like wait, which consists of only a single morpheme; and we must classify English verbal bases into two morphological groups according to the number of their morphemes. Or perhaps take is only one morpheme, and the two morphemes in took are (1) /teyk/ and (2) /−ey+u/;[6] but again the second of these entities fits no definition of a morpheme that linguists are commonly agreed upon.

The difficulty is even greater with an 'unchanged' preterit like put (He put it there yesterday). How can we phrase a description of this word that will be different from a description of put, the corresponding base form?

1.3 The treatment of inflection to be given here is intended to make possible a clear and unambiguous description of all verb forms. By analyzing every inflected form as a combination of morphemes in a particular order, and by avoiding all reference to the process by which the form is derived, we shall be able to systematize the facts of English verb inflection in a way that will be not only more useful to the descriptive linguist than the treatments hitherto published, but also more uniform and in the long run simpler.

6 Read: minus /ey/ plus /u/. Cf. Zellig S. Harris, Lang. 21.121 and fn. 3 (1945).

2. Categories and Assumptions

2.1 A verbal base, in English, is used without any suffix in several different functions: as an infinitive (I can't wait; I don't want to wait), as an imperative (Wait a minute), as a finite present with a subject in the 1st or 2d person singular or in the plural [[or as A. A. Hill puts it, 'with a genderless subject']] (I wait here every day; If you wait for him; They wait in line for it), and in some other ways. In other functions, the verbal base appears with a following inflectional suffix:[7] as a finite present with a subject in the 3d singular [['a gender subject']] (He waits here every day), as a finite preterit (I waited for him), as a participle (I've waited long enough), and as a gerund (I'm still waiting; perhaps also Waiting is tiresome).

We shall speak of 3d singular, preterit, participle, and gerund as the four inflectional categories of English verbs; and we shall speak of every verb form that is used in one of these four functions as an inflected form.[8]

2.2 To prepare the ground for further discussion, we shall briefly examine a number of typical verb forms. Comparing the 3d-singular forms passes, waits, and lives /pæhs-ez, weyt-s, liv-z/,[9] we note that the inflectional suffix appears in three different phonemic shapes /ez, s, z/, whose choice depends on the last phoneme in the base. If we now add the 3d-singular form need (He need not go) and interpret this like the others as consisting of a base plus an inflectional suffix, we find that the suffix has yet another phonemic shape, namely zero, and that the occurrence of this shape cannot be predicted from the last phoneme in the base but only from the base itself; i.e. the choice of zero instead of /z/ depends on the fact that the base here is need and not some other base, such as lead.

Again, the preterit forms waited, passed, and lived /weyt-ed, pæhs-t, liv-d/ reveal three phonemic shapes of the preterit suffix. In general, the choice among them depends on the last phoneme in the base; but a form like dwelt /dwel-t/ instead of the expected /dwel-d/ shows that the alternation among the three shapes is not wholly determined by this criterion.

7 We shall pay no attention in this paper to derivational suffixes (like the -er in waiter, the -ful in wakeful, or the -th in growth), except for a brief mention in §4.5. On the difference between inflection and derivation see Bloch and Trager, Outline of linguistic analysis 54–5 = §4.3 (Baltimore, 1942).

8 The verb be has inflected forms not included among these four categories. Since this verb presents a special and vexing problem, we shall postpone all reference to it until §5. Inflected verb forms used only with a subject in the 2d person singular (archaic forms like waitest, preterit waitedst, and Quaker forms like [thee] waits) are not regarded in this paper as part of present-day standard American English.

9 Hyphens in the phonemic transcription throughout this paper indicate morphological boundaries, not phonemic junctures.

The preterit form put (He put it there yesterday) contains an additional shape of the same suffix, namely zero again.

Not only the inflectional suffixes but bases also may have more than one phonemic shape. If we compare the preterit forms cried and fled /kray-d, fle-d/ with the corresponding uninflected forms, we find that the base cry /kray/ remains unchanged before the suffix /d/, whereas the base flee /fliy/ appears before this suffix in the special shape /fle/. In just the same way, the base take /teyk/ appears before the same suffix in the special shape /tuk/—the only difference being that after this particular base the preterit suffix has the phonemic shape zero, as it has also after the base put.

2.3 We are now ready to state the special assumptions that underlie our treatment of English verb inflection. (1) Every verb form functioning as a 3d-singular finite present, as a finite preterit, as a participle, or as a gerund consists of a base and an inflectional suffix. (2) Different phonemic shapes of a given base appearing before different suffixes, and different phonemic shapes of a given suffix appearing after different bases, are morpheme alternants of the same morpheme.[10] (3) One of the alternants of a given morpheme may be zero; but no morpheme has zero as its only alternant. (4) Different morphemes may have one or more alternants (including zero) in common. (5) Phonemically different forms that occur in the same environment, and are not in completely free variation with each other, are morphemically different.

3. Inflectional Suffixes

3.1 We list here the four morphemes that appear as inflectional suffixes after verbal bases, together with their morpheme alternants. Alternants whose choice depends on the last preceding phoneme (automatic alternants) are connected by a curve (∿); alternants whose choice depends not on a phonemic feature but on the base itself (non-automatic alternants) are separated by a semicolon. The four suffix morphemes are designated by underlined numerals.

Suffix 1 (3d singular): /ez/ after sibilant ∿ /s/ after voiceless non-sibilant ∿ /z/ elsewhere; /0/ (zero).

Suffix 2 (preterit): /ed/ after apical stop ∿ /d/ after voiced sound other than apical stop; /t/; /0/.

Suffix 3 (participle): /ed/ after apical stop ∿ /d/ after voiced sound other than apical stop; /t/; /n/ after syllabic[11] ∿ /ən/ elsewhere; /0/.

Suffix 4 (gerund): /iŋ/.

10 On morpheme alternants and morpheme units see Harris, Lang. 18.170–3 ((110–111)); and cf. Hockett, Lang. 23.341–2 ((241)).

11 A syllabic is a vowel alone or a vowel plus /y, w, h, r, hr/; cf. fn. 5.

3.2 Certain alternants of different suffix morphemes are associated, in the sense that if a given base is followed by one of them, it will be followed also by the other; thus the /t/ alternant of suffix 2 occurs only after bases that are followed also by the /t/ alternant of suffix 3, and conversely. On the other hand, certain alternants of different suffix morphemes are mutually exclusive, in the sense that they never appear after the same base; thus no base that is followed by the zero alternant of suffix 1 is followed also by the zero alternant of suffix 2 or 3.

Among verbs inflected for all four categories we have listed, we find that the alternants of the four suffix morphemes appear in seven different combinations. (Automatic alternants are not differentiated in this count.) These combinations provide a basis for grouping such verbs into seven inflectional classes (A to G). To these we must add two more (H and I) to accommodate verbs that are not inflected for all four of the categories. Some verbs have no participle or gerund; other verbs lack not only these forms but a preterit also.

The following list shows the nine inflectional classes of English verbs, based on the combinations of suffix alternants that accompany their bases. To simplify the listing, the automatic alternants /ez ∿ s ∿ z/ are represented by /z/, the alternants /ed ∿ d/ by /d/, and the alternants /n ∿ ən/ by /n/. One example is given for each class.

Class	Suffix 1	Suffix 2	Suffix 3	Suffix 4	Example
A	z	d	d	iŋ	live
B	z	t	t	iŋ	pass
C	z	0	n	iŋ	fall
D	z	0	0	iŋ	put
E	z	d	n	iŋ	show
F	z	0	d	iŋ	dive
G	0	d	d	iŋ	need
H	0	d			can
I	0				must

3.3 In traditional terminology, verbs in classes A and B are weak verbs; those in classes C and D are strong verbs; those in classes E, F, and G are mixed or anomalous verbs; and those in classes H and I are auxiliaries.

Auxiliaries are not only defective in their morphology, but syntactically peculiar as well. The uninflected form (can, shall, must, may, etc.) is used as a finite present with a subject in the 1st or 2d person singular or in the plural (cf. §2.1), but not as an infinitive or as an imperative.

4. Verbal Bases

4.1 Turning now from the suffix morphemes to the base morphemes, we observe that some verbs (wait, live, beat, show, etc.) have a base with only one phonemic shape—in other words, with only one morpheme alternant.[12] Other verbs (flee, take, fall, bite, etc.) have a base with two different morpheme alternants: one that appears when the base is used alone and before certain of the inflectional suffixes, another that appears before certain other suffixes. And some verbs (sing, fly, drive, do, etc.) have a base with three different morpheme alternants.

This divergence among the base morphemes allows us to divide them into seven base groups according to the number of their morpheme alternants and the particular suffix morphemes before which the alternants occur. When a given base has two alternants or more, they are designated as 'first alternant', 'second alternant', and so on: the first alternant being the one that appears when the base is used alone, and the others being numbered arbitrarily.

Base group 1. Single alternant. (Example: wait.)
Base group 2. Second alternant before suffix 2. (Example: take.)
Base group 3. Second alternant before suffixes 2 and 3. (Example: break.)
Base group 4. Second alternant before suffixes 1, 2, and 3. (Example: say.)
Base group 5. Second alternant before suffix 2; third alternant before suffix 3. (Example: sing.)
Base group 6. Second alternant before suffix 2; third alternant before suffixes 1 and 3. (Only example: do.)
Base group 7. Second alternant before suffixes 2 and 3; third alternant before suffix 1; fourth alternant before suffix 4. (For the only example see §6.1 s.v. have.)

4.2 In citing a verbal base (in the lexicon or elsewhere) we must give all its morpheme alternants, listed in the order determined by the tabulation in §4.1. If we do this, it is obvious that the complete inflection of the verb can then be defined simply by noting the inflectional class and the base group to which it belongs. The following illustrations will make this clear:

wait (A1) /weyt/	take (C2) /teyk; tuk/
build (B3) /bild; bil/	put (D1) /put/
beat (C1) /biyt/	sing (D5) /siŋ; sæŋ; sʌŋ/

The indication A1 after wait means that the suffix morphemes appear after this base in the alternants /s, ed, ed, iŋ/, and that the base appears before all these suffixes in the single form /weyt/; the complete inflection of this verb is accordingly /weyt-s, weyt-ed, weyt-ed, weyt-iŋ/. The indication B3 after build means that the suffix morphemes appear in the alternants /z, t, t, iŋ/, and that the base has two alternants of its own: one, /bild/, appearing before the suffixes /z/ and /iŋ/, the other, /bil/, appearing before /t/ and /t/. The indication C1 after beat means that the suffix morphemes appear as /s, 0, ən, iŋ/, and that the base is uniformly /biyt/ before these suffixes. In the same way, the indications C2, D1, and D5 after take, put, and sing respectively, mean that the complete inflection of these verbs is /teyk-s, tuk-0, teyk-ən, teyk-iŋ; put-s, put-0, put-0, put-iŋ; siŋ-z, sæŋ-0, sʌŋ-0, siŋ-iŋ/.

The twofold classification of a verb according to its inflection class and its base group (indicated by a double symbol such as A1, B3, C2) we shall call its conjugation type.

4.3 The overwhelming majority of English verbs belong to one of two conjugation types: to B1 if the base ends in a voiceless consonant except /t/, otherwise to A1. Such verbs are regular; all others are irregular.

A considerable number of verbs belong to both a regular and an irregular conjugation type. Thus, burn belongs to both A1 and B1 (with inflected forms /bəhrn-z, -d, -d, -iŋ/ and /bəhrn-z, -t, -t, -iŋ/ respectively; fit belongs to both A1 and D1 (with inflected forms /fit-s, -ed, -ed, -iŋ/ and /fit-s, -0, -0, -iŋ/ respectively); heave belongs to both A1 and D3 (with inflected forms /hiyv-z, -d, -d, -iŋ/ and /hiyv-z, howv-0, howv-0, hiyv-iŋ/ respectively). There are also verbs that belong to two or more different irregular conjugation types. Thus, spit belongs to D1 and D3 (with inflected forms /spit-s, -0, -0, -iŋ/ and /spit-s, spæt-0, spæt-0, spit-iŋ/ respectively); tread belongs to C3 and D3 (with inflected forms /tred-z, trad-0, trad-ən, tred-iŋ/ and /tred-z, trad-0, trad-0, tred-iŋ/ respectively); and shrink belongs to four different types—C3, C5, D3, and D5 (with inflected forms /šriŋk-s, šrʌŋk-0, šrʌŋk-ən, šriŋk-iŋ/, /šriŋk-s, šræŋk-0, šrʌŋk-ən, šriŋk-iŋ/, /šriŋk-s, šrʌŋk-0, šrʌŋk-0, šriŋk-iŋ/, and /šriŋk-s, šræŋk-0, šrʌŋk-0, šriŋk-iŋ/ respectively).

[12] Strictly considered, every verbal base has at least two alternants differing in stress; see Wells, Lang. 23.108–14 ((202–5)) = §§66–78. Thus, wait has the following alternants: (1) /weyt/, without inherent stress, when the base is accompanied in an utterance by the emphatic stress morpheme /ˊ/, e.g. Wait a minute /wéyt ə mînit/; (2) /wêyt/, with reduced-loud stress, when the base is not accompanied by the emphatic stress morpheme, e.g. Let's wait here /lèt s wêyt hír/; (3) /wèyt/ with medial stress, again when the base is not accompanied by the emphatic stress morpheme, e.g. Let's not wait up /lèt s nât wèyt ʌ́p/. (The first of these alternants would appear also if the base were pronounced without any phonemic stress; but this would scarcely happen with a word like wait.) In our discussion, we shall disregard all stress differences among morpheme alternants.

Some verbs exhibit a difference in meaning according to their conjugation type; thus, shine is transitive in A1 (I shined my shoes) but intransitive in D3 (The sun shone). Other verbs have the same or approximately the same denotation, but slightly different stylistic and social connotations: the participle shown (E1) is for many speakers more elegant than the participle showed (A1). Still other verbs are apparently identical in both meaning and connotation regardless of their conjugation type, so that the different inflected forms (e.g. burned and burnt) occur interchangeably in all situations—i.e. in completely free variation. According to our assumptions (§2.3), if a verb that belongs to a given conjugation type differs in meaning or connotation, however slightly, from a verb with a phonemically identical base that belongs to another type, the verbs are different morphemes: the shine whose preterit is shined is a different verb from the shine whose preterit is shone; and by the same argument the show whose participle is shown is a different verb from the show whose participle is showed.[13]

Since in practice it is often difficult to decide whether a given phonemic shape that belongs to two or more conjugation types is one verbal base or more than one, we shall not attempt the distinction. Hereafter, we shall use an asterisk to identify all bases that either (1) belong to both a regular and an irregular type, or (2) belong to an irregular type and are homonymous (in at least their first morpheme alternant) with a base belonging to a regular type.

4.4 A number of bases have morpheme alternants whose appearance is controlled not by an inflectional suffix but by some other following element. Chief among such elements is the unstressed morpheme n't — /nt/ after a syllabic, /ənt/ elsewhere — that occurs after eleven of the uninflected bases and after some of the corresponding inflected forms.[14]

The following verb forms appear before this morpheme unchanged:[15]

can, pret. /kud/ : /kud-ənt/.	may, /mey/ : /mey-nt/.
dare, /dehr/ : /dehr-nt/.	might, /mayt/ : /mayt-ənt/.
do, 3d sg. /dʌz/, pret. /did/ : /dʌz-ənt, did-ənt/.	ought, /oht/ : /oht-ənt/.
have, uninflected /hæv/, 3d sg. /hæz/, pret. /hæd/ : /hæv-ənt, hæz-ənt, hæd-ənt/.	need, /niyd/ : /niyd-ənt/.
	shall, pret. /šud/ : /šud-ənt/.
	will, pret. /wud/ : /wud-ənt/.

Certain other bases (and homonymous 3d-singular forms including the zero alternant of suffix 1) appear before the morpheme n't in an alternant that occurs nowhere else:

can, /kæn/ : /kæh-nt/.	must, /mʌst/ : /mʌs-ənt/.
do, uninflected /duw/ : /dow-nt/.	shall, /šæl/ : /šæh-nt/.
	will, /wil/ : /wow-nt/.

4.5 There are also bases that appear in a special alternant shape before certain derivational suffixes. Thus, the base of the verb see has the alternants /siy/ and /soh/ before the inflectional morphemes, but the alternant /say/ before /t/ in the derivative noun sight /say-t/; the base of the verb sing has the alternants /siŋ/, /sæŋ/, and /sʌŋ/ before the inflectional morphemes, but the alternant /sohŋ/ before a zero derivational suffix in the noun song.[16]

Finally, some bases have special sandhi alternants when followed in the same phrase by a word with an initial consonant, especially one that is homorganic or identical with the last consonant of the base in its fuller form. This is especially true of bases ending with /t/ or /d/ after another consonant. Thus, the verb form last, normally /læhst/, may appear as /læhs/ in the phrase How long will it last today? or the like; the verb forms find and found, normally /faynd/ and /fawnd/, may lack the final /d/ in the phrases Did you find time to do it? and I found two dollars. The use of such sandhi alternants is optional: the same person will speak sometimes the shorter form, sometimes the longer in the same context. In the remainder of this paper we shall ignore them entirely.

5. The Verb be

5.1 The base of the verb be has a greater number of morpheme alternants than any other. As the first alternant—the one that appears when the base is used without any inflectional suffix—we may set up the shape /biy/; but this has a more limited use than the first alternant of other verbs except the auxiliaries (§2.1, §3.3). It occurs freely as an infinitive

[13] The argument: Phonemically different forms that occur in the same environment, and are not in completely free variation with each other, are morphemically different (§2.3). In the two participles shown /šow-n/ and showed /šow-d/, the phonemically different elements /n/ and /d/ both follow the base /šow/. They are not in free variation, since the two forms have different connotations of elegance and hence are not interchangeable. Therefore, either /n/ and /d/ are different morphemes, or the environment in which they occur is after all not the same for both. Since we wish to identify /n/ and /d/ as alternants of the same morpheme (suffix 3), we assume that the /šow/ that precedes /n/ is morphemically different from the /šow/ that precedes /d/. In other words, we choose to set up two different but homonymous morphemes /šow/, and to refer to them—rather than to the suffix alternants—the stylistic or connotative difference between the inflected forms shown and showed.

[14] The form n't is best regarded as a separate morpheme, not as an alternant of the full form not. The two forms contrast, at least stylistically and in their connotations, in such phrases as I cannot go : I can't go.

[15] Any form in this list not otherwise identified is both the uninflected form of the verb and the homonymous 3d-singular form with zero suffix.

[16] It goes without saying that historical considerations play no part in a structural description. The actual historical relation between sing and song is irrelevant here; all that is relevant is their morphological relation in the structure of present-day English.

(I can't be there; I don't want to be there) and as an imperative (Be quiet); but as a finite present it occurs only after if, lest, etc., and only in rather formal style (If they be still there; If such there be). In the latter use, /biy/ occurs also with a 3d-singular subject (If it be not presumptuous; Lest it be objected that...), thus differing again from other verbal bases. Moreover, /biy/ is unique in its optative and concessive uses (God be praised; The public be damned; Be it never so humble; Be that as it may).[17]

5.2 The inflected forms of the verb be in the four usual inflectional categories are as follows: 3d singular is (base alternant /i/, suffix alternant /z/); preterit was and were (base alternants /wʌz/ and /wəhr/ in complementary distribution,[18] suffix alternant zero); participle been (base alternant /bi/,[19] suffix alternant /n/); gerund being (base alternant /biy/, suffix alternant /iŋ/).

But there are still other inflected forms of be that belong to none of these four categories. This constitutes the most striking idiosyncrasy of the verb: that its inflection distinguishes categories not recognized by the morphology of any other verb. We find, first, the forms am and are /æm, ahr/. These might be regarded as uninflected forms, in complementary distribution with the alternant /biy/; for the three words together have the same distribution as the uninflected form of such a verb as wait. The objection to this view is that the alternant /biy/ occurs, as already mentioned, in clauses with if (§5.1), contrasting in this position with /æm/ and /ahr/ (If I be not mistaken : If I am not mistaken; If you be he : If you are he). If the contrast is valid, we must set up a new inflectional category for am and are, perhaps to be called general present (i. e. non-3d-singular present). But in that case am and are include an inflectional suffix; and this cannot be zero, since we have assumed (§2.3) that no morpheme has zero as its only alternant. Any solution of the problem is inescapably ad hoc; we propose to regard the alternants of the base here as /æ/ and /ah/ in complementary distribution, and to posit a suffix morpheme with two alternants /m/ and /r/, the choice between them being regulated by the shape of the base alternant: /m/ after /æ/, /r/ after /ah/.

Complementary distribution and identity of meaning have allowed us to treat the two preterit forms was and were as both containing the zero alternant of suffix 2. But there is another form were which is not preterit and which contrast with was: the form that appears in conditional clauses after if with a subject in the 1st or 3d person singular (If I were rude, I'd apologize; If he were here, he'd see it; contrast If I was rude, I apologize; If he was here, he saw it). This were must be an inflected form different from the preterit were; its category we may call (with Hockett) the unreal. Again we analyze the form ad hoc: base alternant /wəh/, suffix morpheme /r/.[20]

5.3 Our analysis results in the following array of forms for the verb be: uninflected /biy/; 3d singular /i-z/; preterit /wʌz-0/ and /wəhr-0/ in complementary distribution; participle /bi-n/; gerund /biy-iŋ/; general present /æ-m/ and /ah-r/ in complementary distribution; unreal /wəh-r/. The base morpheme has eight alternants: /biy, i, wʌz, wəhr, bi, æ, ah, wəh/; and the six inflectional morphemes appear in the following alternants: 1 /z/, 2 /0/, 3 /n/, 4 /iŋ/, 5 /m/ and /r/ in complementary distribution, 6 /r/. This multiplicity of forms calls for an addition both to our list of inflectional classes (§3.2) and to our list of base groups (§4.1). Accordingly, we set up (again ad hoc) class J and base group 8 to accommodate the verb be. The former has been defined already in the list of suffix alternants just given; the latter is defined as follows:

Base group 8. Second alternant before suffix 1; third and fourth alternants before suffix 2; fifth alternant before suffix 3; sixth and seventh alternants before suffix 5; eight alternant before suffix 6.

5.4 Three of the inflected forms of be (but not the uninflected form) appear unchanged before the morpheme n't (§4.4). These are 3d singular /iz/, preterit /wʌz/ and /wəhr/, general present /ahr/ but not /æm/ : /iz-ənt, wʌz-ənt, wəhr-nt, ahr-nt/. In the substandard form ain't /ey-nt/, the /ey/ is a ninth alternant of the base, here followed by the zero alternants of suffixes 1 and 5.

6. List of Irregular Bases

6.1 The following list shows the bases of irregular verbs as they might appear in the lexicon. The list is thought to be complete, including all irregular verbs current in standard colloquial English, together with a few that are no longer used in conversation but are still occasionally spoken in a formal or literary style. Completely obsolete forms and all clearly substandard forms are omitted.

17 The optative use of be is paralleled by other verbs in a few formulas: God have mercy; God forbid; Perish the thought. The concessive use of other verbs than be is limited to such archaic locutions as Try they never so hard.

18 The alternant /wʌz/ occurs only with a subject in the 1st or 3d person singular; the alternant /wəhr/ occurs with all other subjects.

19 Or /biy/ in British English, identical with the alternant that appears in the uninflected form.

20 Since be has more inflected forms than any other verb, we might have begun our discussion with it instead of saving it for the end. In that case we should have said that all other verbs (except the auxiliaries) have the same inflectional categories as be, but that only be formally distinguishes the general present from the uninflected form or the unreal from the preterit. This is essentially what Hockett did in his rigorously systematic treatment of English verbs (op. cit. in fn. 2). The treatment here adopted seems preferable because it results in a simpler statement.

Verbs are alphabetized according to the conventional spelling of their uninflected forms. Each entry includes a symbol denoting the conjugation type of the verb, and a phonemic transcription of the base alternants that appear before the inflectional suffixes. If there is also a special alternant of the base appearing before the morpheme n't (§4.4), this is added after the symbol N; but alternants that appear only before derivational suffixes (like the /say/ in sight) and all sandhi alternants (§4.5) are omitted. Verbs that belong also to a regular conjugation type (AI) are marked with an asterisk. Verbs that belong to more than one irregular conjugation type are entered separately for each type: beat, bite, cleave, get, have, hide, spit, spring, stink, stride, tread, wake, are entered twice each; bid and shrink are entered four times each.

Compound verbs containing an irregular verb as the second member (e. g. broadcast) and derivative verbs formed with a prefix from an irregular verb (e. g. be-come) are not entered. Where a given base does not occur without a prefix (e. g. be-gin), the full form of the verb is given, but the prefix is disregarded in alphabetizing and in citing the base alternants.

Comments on individual verbs are included where they seem necessary.[21]

be J8 /biy, i, wʌz, wəhr, bi, æ, ah, wəh/. The only member of class J and of base group 8. In some dialects the alternant /bi/ is lacking; in others, it is /be/.
bear C3 /behr, bohr/.
beat C1 /biyt/.
beat D1 /biyt/.
bend B3 /bend, ben/.
*bet D1 /bet/.
better I1 /betər/. The only dissyllabic base. Used as a verb in colloquial speech without any form of had: I better; You better go (cf. I must; You must go). When had or 'd is present (I'd better; You'd better go), the word is not a verb but an adverb like rather.
bid C1 /bid/. Also outbid, underbid, overbid; not forbid.
bid C2 /bid, bæhd/. Also forbid.
bid D1 /bid/. See bid C1.
bid D2 /bid, bæhd/. See bid C2.
*bide D3 /bayd, bowd/.
bind D3 /baynd, bawnd/.
bite C3 /bayt, bit/.
bite D3 /bayt, bit/.
bleed D3 /bliyd, bled/.
blow C2 /blow, bluw/.
break C3 /breyk, browk/.
breed D3 /briyd, bred/.
bring B3 /briŋ, broh/.
build B3 /bild, bil/.
*burn B1 /bəhrn/.
burst D1 /bəhrst/.
buy B3 /bay, boh/.
can H2 /kæn, ku/; N /kæh/. Not homonymous with /kæhn/ 'tin'.
cast D1 /kæhst/.
catch B3 /kæc, koh/.
choose C3 /cuwz, cowz/.
*cleave 'split' B3 /kliyv, klef/.
*cleave 'split, adhere' C3 /kliyv, klowv/.
cling D3 /kliŋ, klʌŋ/.
come D2 /kʌm, keym/.
cost D1 /kohst/.
*crow F2 /krow, kruw/.
creep B3 /kriyp, krep/.
cut D1 /kʌt/.
*dare G1 /dehr/.
deal B3 /diyl, del/.
dig D3 /dig, dʌg/.
*dive F2 /dayv, dowv/.
do E6 /duw, di, dʌ/; N /dow/. The only member of base group 6.
draw C2 /droh, druw/.
*dream B3 /driym, drem/. In the pronunciation [drempt], the [p] is predictable and hence not phonemically distinctive.
drink D5 /driŋk, dræŋk, drʌŋk/.
drive C5 /drayv, drowv, driv/.
dwell B1 /dwel/.
eat C2 /iyt, eyt/. In standard British English, base alternants /iyt, et/.
fall C2 /fohl, fel/.
feed D3 /fiyd, fed/.
feel B3 /fiyl, fel/.
fight D3 /fayt, foht/.
find D3 /faynd, fawnd/.
*fit D1 /fit/.
flee A3 /fliy, fle/.
fling D3 /fliŋ, flʌŋ/. There is a jocular preterit /flæŋ/, one of a number of non-standard 'strong' preterits occasionally heard in facetious utterance.
fly C5 /flay, fluw, flow/.
freeze C3 /friyz, frowz/.
get C3 /get, gat/. Also beget, forget.
get D3 /get, gat/. Not beget, forget.
begin D5 /gin, gæn, gʌn/.
*gird B3 /gəhrd, gər/.
give C2 /giv, geyv/.
go C5 /gow, went, goh/. The classical instance of 'suppletive' inflection. In went, the /t/ could also be regarded as an alternant of suffix 2; but this interpretation would require the creation of a new inflectional class for this verb alone, with suffix alternants /z, t, n, iŋ/. By regarding went as consisting of a base alternant /went/ and the zero alternant of suffix 2, we escape this necessity. —The phonemic dissimilarity between /gow/ and /went/ is no bar to grouping them together as alternants of the same morpheme; for we have not assumed (in §2.3) that morpheme alternants must resemble each other phonemically.
grind D3 /graynd, grawnd/.
grow C2 /grow, gruw/.
*hang D3 /hæŋ, hʌŋ/.
have 'possess, etc.' A4 /hæv, hæ/.
have 'be obliged' G7 /hæf, hæ, hæs, hæv/. The only member of base group 7. In colloquial speech, this verb has the following forms: uninflected /hæf/ (I have to leave), 3d singular /hæs-0/ (He has to leave), preterit /hæ-d/ (I had to leave), participle /hæv-iŋ/ (I'm having to leave tomorrow). The form /hæs/ is analyzed as including the zero alternant of suffix 1; another possible analysis would be into the base alternant /hæ/ and the suffix alternant /s/; but this would be the only case of /s/ after a voiced sound, and would make it impossible to regard the alternation of /ez ∿ s ∿ z/ as automatic.
hear A3 /hihr, həhr/.
*heave D3 /hiyv, howv/.
*hew E1 /hyuw/.
hide C3 /hayd, hid/.
hide D3 /hayd, hid/.
hit D1 /hit/.
hold D3 /howld, held/.
hurt D1 /hərt/.
keep B3 /kiyp, kep/.
*kneel B3 /niyl, nel/.
*knit D1 /nit/.
know C2 /now, nuw/. In some dialects the second alternant is /nyuw/ or /niw/.
lead D3 /liyd, led/.
*lean B3 /liyn, len/.
*leap B3 /liyp, lep/.
*learn B1 /ləhrn/.
leave B3 /liyv, lef/.
lend B3 /lend, len/.
let D1 /let/.
*light D3 /layt, lit/.

21 One verb not listed in this section requires a word of comment. Beware is used as an infinitive and as an imperative (I told him to beware; Beware of the dog); inflected forms (bewares, bewared, bewaring) are either not used at all in present-day English or at best extremely rare. We may tentatively assign beware to conjugation type AI as a regular verb whose inflected forms happen to be not in use.

lose B3 /luwz, lohs/.
make A3 /meyk, mey/.
may I1 /mey/. See might.
mean B3 /miyn, men/.
meet D3 /miyt, met/.
might I1 /mayt/. Historically but no longer in colloquial usage the preterit of may.
must I1 /mʌst/; N /mʌs/.
*need G1 /niyd/.
ought I1 /oht/.
*pen B1 /pen/.
*plead D3 /pliyd, pled/.
*prove E1 /pruwv/.
put D1 /put/.
*quit D1 /kwit/.
read D3 /riyd, red/.
*bereave B3 /riyv, ref/.
rend B3 /rend, ren/.
*rid D1 /rid/.
ride C5 /rayd, rowd, rid/.
*ring D5 /riŋ, ræŋ, rʌŋ/. Regular in the meaning 'encircle'.
rise C5 /rayz, rowz, riz/.
run D2 /rʌn, ræhn/.
forsake C2 /seyk, suk/.
say A4 /sey, se/.
see C2 /siy, soh/.
*beseech B3 /siyc, soh/.
seek B3 /siyk, soh/.
sell A3 /sel, sowl/.
send B3 /send, sen/.
set D1 /set/.
*sew E1 /sow/. Same as *sow.
shake C2 /šeyk, šuk/.
shall H2 /šæl, šu/; N /šæh/. Though may and might (qq.v.) have become different verbs in colloquial usage, should can still be viewed as the preterit of shall. Those who deny this possibility would set up two independent verbs: shall I1 /šæl/, N /šæh/; should I1 /šud/.
*shear C3 /šihr, šohr/.
shed D1 /šed/.
*shine D3 /šayn, šown/. In standard British English the base alternants are /šayn, šon/.
*shoe A3 /šuw, ša/.
shoot C3 /šuwt, šat/.
*show E1 /šow/.
shrink C3 /šriŋk, šrʌŋk/.
shrink C5 /šriŋk, šræŋk, šrʌŋk/.
shrink D3 /šriŋk, šrʌŋk/.
shrink D5 /šriŋk, šræŋk, šrʌŋk/.
*shrive C5 /šrayv, šrowv, šriv/.
shut D1 /šʌt/.
sing D5 /siŋ, sæŋ, sʌŋ/.
sink D5 /siŋk, sæŋk, sʌŋk/.
sit D3 /sit, sæt/.
slay C2 /sley, sluw/.
sleep B3 /sliyp, slep/.
slide D3 /slayd, slid/.
sling D3 /sliŋ, slʌŋ/.
slink D3 /sliŋk, slʌŋk/.
*slit D1 /slit/.
*smell B1 /smel/.
smite C5 /smayt, smowt, smit/.
*sow E1 /sow/. Same as *sew.
speak C3 /spiyk, spowk/.
*speed D3 /spiyd, sped/.
*spell B1 /spel/.
spend B3 /spend, spen/.
*spill B1 /spil/.
spin D3 /spin, spʌn/.
*spit D1 /spit/. Regular 'impale'.
spit D3 /spit, spæt/.
split D1 /split/.
*spoil B1 /spoyl/.
spread D1 /spred/.
spring D3 /spriŋ, spʌŋ/.
spring D5 /spriŋ, spræŋ, spʌŋ/.
stand A3 /stændn, stu/.
*stave D3 /steyv, stowv/.
steal C3 /stiyl, stowl/.
stick D3 /stik, stʌk/.
sting D3 /stiŋ, stʌŋ/.
stink D3 /stiŋk, stʌŋk/.
stink D5 /stiŋk, stæŋk, stʌŋk/.
*strew E1 /struw/.
stride C5 /strayd, strowd, strid/.
stride D3 /strayd, strowd/.
strike D3 /strayk, strʌk/.
*string D3 /striŋ, strʌŋ/.
*strive C5 /strayv, strowv, striv/.
swear C3 /swehr, swohr/.
*sweat D1 /swet/.
sweep B3 /swiyp, swep/.
*swell C3 /swel, swowl/.
swim D5 /swim, swæm, swʌm/.
take C2 /teyk, tuk/.
teach B3 /tiyc, toh/.
tear C3 /tehr, tohr/.
tell A3 /tel, towl/.
think B3 /θiŋk, θoh/.
*thrive C5 /θrayv, θrowv, θriv/.
throw C2 /θrow, θruw/.
thrust D1 /θrʌst/.
tread C3 /tred, trad/.
tread D3 /tred, trad/.
used I1 /yuws/. Though spelled with -d, this acts in colloquial speech like a finite present, not a preterit—the past-time meaning in this verb can be regarded as inherent in the base itself.
*wake C3 /weyk, wowk/.
*wake D3 /weyk, wowk/.
wear C3 /wehr, wohr/.
*weave C3 /wiyv, wowv/.
*wed D1 /wed/.
weep B3 /wiyp, wep/.
*wet D1 /wet/.
*will H2 /wil, wu/; N /wow/. On would as the preterit of will, cf. shall. Regular 'effect by the will' or the like.
win D3 /win, wʌn/.
wind D3 /waynd, wawnd/.
*wreak B3 /riyk, roh/. Wrought seems to be semantically closer to wreak than to work; and the relation wreak:wrought has a parallel in seek:sought, whereas work:wrought would be unique. Except in very formal discourse and in a few clichés, this verb is now obsolete. The adjective wrought (in wrought iron and the like) is of course not a form of the verb.
wring D3 /riŋ, rʌŋ/.
write C5 /rayt, rowt, rit/.

6.2 The foregoing list contains 200 entries, representing 182 orthographically different words. The number of verbs in each inflectional class, base group, and conjugation type is shown in the table below. The letters across the top denote inflectional classes; the numerals from 1 to 8 at the left denote base groups.

	A	B	C	D	E	F	G	H	I	J	Total
1		8	2	26	6	–	2	–	6	–	50
2	–	–	14	3	–	2	–	3	–	–	22
3	7	31	20	44	–	–	–	–	–	–	102
4	2	–	–	–	–	–	–	–	–	–	2
5	–	–	12	9	–	–	–	–	–	–	21
6	–	–	–	–	1	–	–	–	–	–	1
7	–	–	–	–	–	–	1	–	–	–	1
8	–	–	–	–	–	–	–	–	–	1	1
Totals	9	39	48	82	7	2	3	3	6	1	200

[The inadvertent omission of the verb lie (lay, lain) has resulted in certain statistical and other errors, especially in the section on morphophonemics. The following entry should be inserted in the above list:
lie C3 /lay, ley/ [[addendum supplied in 1957]]]

6.3 We now show in full the membership of each of the twenty conjugation types to which irregular verbs belong. The ten inflectional classes (A to J) are again defined by the suffix alternants that appear after the verbal base; within each class, the bases that belong to it are grouped according to their conjugation types. Since the morpheme alternants of the bases have already been given (§6.1), bases are cited here only in the conventional spelling of the uninflected form.

An asterisk, as before, marks bases that belong also to a regular conjugation type; a plus sign marks bases that belong to more than one irregular type.

Class A. Suffix alternants: 1 /z/, 2 /d/, 3 /d/, 4 /iŋ/.
Type A3: flee, hear, make, sell, *shoe, stand, tell.
Type A4: have+, say.

Class B. Suffix alternants: 1 /z/, 2 /t/, 3 /t/, 4 /iŋ/.
Type B1: *burn, dwell, *learn, *pen, *smell, *spell, *spill, *spoil. Bases end with /n/ or /l/. Except for dwell, all verbs belong also to A1.

Type B3: bend, bring, build, buy, catch, *cleave+, creep, deal, *dream, feel, *gird, keep, *kneel, *lean, *leap, leave, lend, lose, mean, *bereave, rend, *beseech, seek, send, sleep, spend, sweep, teach, think, weep, *wreak.

Class C. Suffix alternants: 1 /z/, 2 /0/, 3 /n/, 4 /iŋ/.
Type C1: beat+, bid+.

Type C2: bid+, blow, draw, eat, fall, give, grow, know, forsake, see, shake, slay, take, throw.

Type C3: bear, bite+, break, choose, *cleave+, freeze, get+, hide+, *shear, shoot, shrink+, speak, steal, swear, *swell, tear, tread+, *wake+, wear, *weave.

Type C5: drive, fly, go, ride, rise, shrink+, stride+, *strive, smite, *shrive, *thrive, write.

Class D. Suffix alternants: 1 /z/, 2 /0/, 3 /0/, 4 /iŋ/.
Type D1: beat+, *bet, bid+, burst, cast, cost, cut, *fit, hit, hurt, *knit, let, put, *quit, *rid, set, shed, shut, *slit, *spit+, split, spread, *sweat, thrust, *wed, *wet. All bases end in /t/ or /d/.

Type D2: bid+, come, run.

Type D3: *bide, bind, bite+, bleed, breed, cling, dig, feed, fight, find, fling, get+, grind, *hang, *heave, hide+, hold, lead, *light, meet, *plead, read, *shine, shrink+, sit, slide, sling, slink, *speed, spin, *spit+, spring+, *stave, stick, sting, stink+, stride+, strike, *string, tread+, wake+, win, wind, wring. By far the largest conjugation type of irregular verbs.

Type D5: drink, begin, *ring, shrink+, sing, sink, spring+, stink+, swim. Except for swim, bases end with /ŋ/ or /ŋk/; without exception, they end with a nasal or a nasal + /k/.

Class E. Suffix alternants: 1 /z/, 2 /d/, 3 /n/, 4 /iŋ/.
Type E1: *hew, *prove, *sew, *show, *sow, *strew. All bases belong also to A1. Sew and sow differ only in meaning (and spelling), but perhaps not more widely than different senses of certain verbs listed only once.

Type E6: do.

Class F. Suffix alternants: 1 /z/, 2 /0/, 3 /d/, 4 /iŋ/.
Type F2: *crow, *dive. Both bases belong also to A1.

Class G. Suffix alternants: 1 /0/, 2 /d/, 3 /d/, 4 /iŋ/.
Type G1: *dare, *need. Both bases belong also to A1.
Type G7: have+ 'be obliged'.

Class H. Suffix alternants: 1 /0/, 2 /d/.
Type H2: can, shall, *will.

Class I. Suffix alternant: 1 /0/.
Type I1: better, may, might, must, ought, used.

Class J. Suffix alternants: 1 /z/, 2 /0/, 3 /n/, 4 /iŋ/, 5 /m, r/, 6 /r/. Type J8: be.

7. Morphophonemics

7.1 Morphophonemics is the study of the alternation between corresponding phonemes in alternant shapes of the same morpheme. When the morpheme alternants of a language, or of some form-class in a language, have been listed in full, a statement of the morphophonemics will serve as a convenient index to the listing. Since the base alternants of English irregular verbs have been listed in §6.1, the present section is an index to that one.

We shall group together here all bases that exhibit the same phonemic difference among their alternants, regardless of the inflectional classes to which they belong. There is no necessary connection between any morphophonemic set and any inflectional class. Some of the sets to be listed include verbs of several different classes; for instance, the alternation /iy/ ∿ /e/ is found in bases of class A, class B, and class D. Other sets include verbs of only one class, but not all the verbs that belong there; for instance, the alternation /ay/ ∿ /aw/ is found only in bases of class D, but by no means in all the bases of that class. No set is coextensive in its membership with any class.

In listing the alternations between phonemes, we shall regard every syllabic as a unit, whether it consists of a vowel alone or of a vowel and a following semivowel /y, w, h, r, hr/. Thus, we shall express the difference between eat /iyt/ and ate /eyt/ as /iy/ ∿ /ey/, not simply as /i/ ∿ /e/, and the difference between bind /baynd/ and bound /bawnd/ as /ay/ ∿ /aw/, not simply as /y/ ∿ /w/.

Verbs of base group group 1, whose bases have only a single alternant, naturally do not appear in the listing.

7.2 There are ten types of morphophonemic alternation between base alternants of English irregular verbs. The first four types are found in bases with two alternants, the next three in bases with three alternants, and the rest in bases with more than three alternants. The types may be characterized as follows:[22]

Type I. Alternation between two different syllabics ($V_1 \sim V_2$).

Type II. Alternation between a syllabic with one or two following non-syllabics and a different syllabic without a following non-syllabic (V_1C or $V_1CC \sim V_2$).

Type III. Alternation between a syllabic with a following non-syllabic and a different syllabic with a different following non-syllabic ($V_1C_1 \sim V_2C_2$).

Type IV. Alternation between the presence and the absence of a non-syllabic ($C \sim 0$).

22 The following symbols are used to characterize the several types of alternation: V = syllabic (not merely vowel), C = non-syllabic consonant, 0 = zero (i.e. absence of a phoneme), V_1 and V_2 = different syllabics. The symbol ∿ everywhere means 'in alternation with'.

Type V. Alternation among three different syllabics ($V_1 \sim V_2 \sim V_3$).

Type VI. Alternation between a syllabic with a following non-syllabic and two different syllabics without a following non-syllabic ($V_1C \sim V_2 \sim V_3$).

Type VII. Suppletion: alternation among three phonemic shapes that have no phoneme in common.

Type VIII. Alternation among four different syllabics ($V_1 \sim V_2 \sim V_3 \sim V_4$).

Type IX. Alternation among three different non-syllabics and the absence of a non-syllabic ($C_1 \sim 0 \sim C_2 \sim C_3$).

Type X. Suppletion: alternation among eight phonemic shapes that have no phoneme in common.

Type I

/ay/ ∼ /i/ bite, hide, light, slide.
/iy/ ∼ /e/ bleed, breed, creep, deal, dream, feed, feel, flee, keep, kneel, lead, lean, leap, mean, meet, plead, read, sleep, speed, sweep, weep.
/ey/ ∼ /e/ say.
/ow/ ∼ /e/ hold.
/oh/ ∼ /e/ fall.
/i/ ∼ /æ/ sit, spit.
/e/ ∼ /a/ get, tread.
/uw/ ∼ /a/ shoe, shoot.
/i/ ∼ /ʌ/ cling, dig, fling, shrink, sling, slink, spin, spring, stick, sting, stink, string, win, wring.
/æ/ ∼ /ʌ/ hang.
/ay/ ∼ /ʌ/ strike.
/ey/ ∼ /u/ forsake, shake, take.
/i/ ∼ /ey/ give.
/ʌ/ ∼ /ey/ come.
/iy/ ∼ /ey/ eat.
/ay/ ∼ /aw/ bind, find, grind, wind.
/e/ ∼ /ow/ sell, swell, tell.
/iy/ ∼ /ow/ cleave, freeze, heave, speak, steal, weave.
/ey/ ∼ /ow/ break, stave, wake.
/ay/ ∼ /ow/ bide, dive, shine, stride.
/uw/ ∼ /ow/ choose.
/ey/ ∼ /uw/ slay.
/ow/ ∼ /uw/ blow, crow, grow, know, throw.
/oh/ ∼ /uw/ draw.
/i/ ∼ /æh/ bid.
/ʌ/ ∼ /æh/ run.
/iy/ ∼ /oh/ see.
/ay/ ∼ /oh/ buy, fight.
/ihr/ ∼ /ehr/ hear.
/ihr/ ∼ /ohr/ shear.
/ehr/ ∼ /ohr/ bear, swear, tear, wear.

Type II

/iyk/ ∼ /oh/ seek, wreak.
/iyc/ ∼ /oh/ beseech, teach.
/æc/ ∼ /oh/ catch.
/iŋ/ ∼ /oh/ bring.
/iŋk/ ∼ /oh/ think.
/æhnd/ ∼ /u/ stand.
/ehrd/ ∼ /er/ gird.

Type III

/iyv/ ∼ /ef/ cleave, leave, bereave.
/uwz/ ∼ /ohs/ lose.

Type IV

/d/ ∼ 0 bend, build, lend, rend, send, spend.
/t/ ∼ 0 must.
/k/ ∼ 0 make.
/v/ ∼ 0 have 'possess'.

Type V

/i/ ∼ /æ/ ∼ /ʌ/ drink, begin, ring, shrink, sing, sink, spring, stink, swim.
/ay/ ∼ /ow/ ∼ /i/ drive, ride, rise, shrive, smite, stride, strive, thrive, write.
/ay/ ∼ /uw/ ∼ /ow/ fly.

Type VI

/æn/ ∼ /u/ ∼ /æh/ can.
/æl/ ∼ /u/ ∼ /æh/ shall.
/il/ ∼ /u/ ∼ /ow/ will.

Type VII

/gow/ ∼ /went/ ∼ /goh/ go.

Type VIII

/uw/ ∼ /i/ ∼ /ʌ/ ∼ /ow/ do.

Type IX

/f/ ∼ 0 ∼ /s/ ∼ /v/ have 'be obliged'.

Type X

/biy/ ∼ /i/ ∼ /wəz/ ∼ /wəhr/ ∼ /bi/ ∼ /æ/ ∼ /ah/ ∼ /wəh/ be.

8. Atonic Verbs

8.1 We have not quite finished. In the sentence I have seen it /ày hǽv síyn it/, the verb have /hæv/ is accompanied by the emphatic stress morpheme /ˊ/ (cf. fn. 12). In the sentence I've seen it /ày v síyn it/, where the stress morpheme accompanies the participle seen, the verb have appears as /v/. This difference between the two phonemic shapes of the finite verb might be attributed to the different locations of the stress—a view that would add to the number of morpheme alternants of have and of some other bases, but would not essentially complicate the system. Unfortunately, this obvious and convenient explanation will not work. The alternant /hæv/ also occurs without the stress morpheme /ˊ/, as in I have seen it /ày hæv síyn it/,[23] which contrasts at least in style and connotation with I've seen it /ày v síyn it/. Phonemically different forms that occur in the same environment, and are not in completely free variation with each other, are morphemically different (§2.3); therefore /hæv/ and /v/, so far as they differ in their social connotation, are not merely alternants of the same morpheme, but different morphemes.

[23] We have already agreed to disregard the difference between such alternants as /hǽv, hèv, hæv/—a difference in stress alone (fn. 12). But even if we discriminate such alternants, the situation will not be affected; for the completely unstressed alternant /hæv/ is common and perfectly natural in the cited context.

A similar contrast exists between some of the inflected forms of the verb /hæv/ and those of the verb /v/; 3d singular /hæz/ in He has left : /z/ in He's left; preterit /hæd/ in He had left : /d/ in He'd left. But no such contrast is possible between different participles or gerunds: the verb /v/ simply lacks these categories. Now the form /hæz/ obviously consists of the base alternant /hæ/ and the suffix alternant /z/; /hæd/ consists of /hæ/ and the suffix alternant /d/. In parallel fashion, the verb forms /z/ in He's left and /d/ in He'd left are analyzed as consisting of a base alternant and the suffixes /z/ and /d/ respectively. The base alternant, accordingly, is zero. In the light of our assumptions (§2.3) and of our treatment of suffix morphemes (§3.1), there is no theoretical objection to such a view: if /0/ can be admitted as a morpheme alternant of /ez ∼ s ∼ z/ (suffix 1), it can equally well be admitted as a morpheme alternant of such a base as /v/.

The verb form /s/ in Jack's left is of course an automatic alternant of the /z/ in He's left: the base alternant in both forms is zero, and the choice of the suffix alternant, as already noted (§3.1), depends on the preceding phoneme.

8.2 But have is not yet fully disposed of. We can say not only I have seen it /ày hæv síyn it/ and I've seen it /ày v síyn it/, both with the emphatic stress morpheme on seen, but also /ày həv síyn it/ and /ày əv síyn it/—with forms of have that are phonetically intermediate between the 'full form' /hæv/ and the 'completely reduced form' /v/. To what morphemes do these intermediate pronunciations belong? Reluctant as we may be to allow the multiplication of elements, we cannot escape the conclusion that the verb forms /həv/ and /əv/ belong to none of the morphemes mentioned so far. Since they occur in the same environments as /hæv/ and /v/, and differ from them—though perhaps only minutely—in their social flavor, they are morphemically different from both.

In the same way, the inflected forms /həz/ and /əz/, /həd/ and /əd/, as in He has left and He had left, are to be analyzed as consisting of the base alternants /hə/ and /ə/ respectively and the suffix alternants already mentioned.

Accordingly, we set up four verbal bases of related meaning but different connotation: one with morpheme alternants /hæv/ and /hæ/; one with morpheme alternants /həv/ and /hə/; one with morpheme alternants /əv/ and /ə/; and one with morpheme alternants /v/ and /0/. The first of these belongs to conjugation type A4; the other three, forming their 3d-singular and preterit forms with /z/ and /d/ respectively but lacking a participle and a gerund, belong to none of the types that have been established.

8.3 It is a syntactic peculiarity of the verbs /həv/, /əv/, and /v/ that they are never, like other verbs, accompanied by the emphatic stress morpheme /ˊ/. For this reason we may call them atonic verbs.

The number of atonic verbs is at least eleven. In addition to the verbs /həv/, /əv/, and /v/, we find the following:

Be, with morpheme alternants /bi/ uninflected, /0/ before suffixes 1 and 5; /wəz/ and /wər/ in complementary distribution before suffix 2; /wə/ before suffix 6. Inflected forms: 3d singular /0-z/ and /0-s/, preterit /wəz-0/ and /wər-0/, general present /0-m/ and /0-r/, unreal /wə-r/; participle and gerund lacking.

Be, with morpheme alternants as above except for /ə/ instead of /0/ before suffix 5. Inflected forms: as above except for general present /ə-m/ and /ə-r/.

Can, with morpheme alternants /kən/ and /kəŋ/ in complementary distribution[24] uninflected and before suffix 1, /kə/ before suffix 2. Inflected forms: 3d singular /kən-0/ and /kəŋ-0/, preterit /kə-d/; other categories lacking, as in the verb /kæn/.

Do, with unique morpheme alternant /də/. Inflected forms: 3d singular /də-z/; other categories lacking.

Shall, with morpheme alternants /šəl/ uninflected and before suffix 1, /šə/ before suffix 2. Inflected forms: 3d singular /šəl-0/, preterit /šə-d/; other categories lacking, as in the verb /šæl/.[25]

Will, with morpheme alternants /wəl/ uninflected and before suffix 1, /wə/ before suffix 2. Inflected forms: 3d singular /wəl-0/, preterit /wə-d/; other categories lacking, as in the verb /wil/.

Will, with morpheme alternants /əl/ uninflected and before suffix 1, /ə/ before suffix 2. Inflected forms: 3d singular /əl-0/, preterit /ə-d/; other categories lacking, as in the verb /wil/.

Will, with morpheme alternants /l/ uninflected and before suffix 1, /0/ before suffix 2. Inflected forms: 3d singular /l-0/, preterit /0-d/; other categories lacking, as in the verb /wil/.

8.4 The most efficient way of describing these atonic forms is to set them up as a major subclass of English verbs, coordinate with a subclass of tonic verbs that will include all the rest.[26] The inflectional classes and base groups that were established in the first part of this paper apply to tonic verbs only; for atonic verbs we must add a separate classification:

[24] /kəŋ/ before /k/ or /g/ (I can come; I can go), /kən/ elsewhere.

[25] Some speakers use also a preterit form /št/, as in I should think so /ày št θíŋk sów/. For such dialects we must set up another verb shall, with morpheme alternants /šəl/ uninflected and before suffix 1, /š/ before suffix 2.

[26] For pedagogical purposes, this description of atonic verbs—and of tonic verbs too, for that matter—would certainly be over-meticulous and ineffective. To the student learning to speak English, busy at his primary task of memorizing model sentences, it would be unhelpful at best, if not actively confusing. But the intent of this paper is not pedagogical.

Class	Suffix 1	Suffix 2	Suffix 5	Suffix 6
A′	z	d		
B′	0	d		
C′	z			
D′	z	0	m; r	r

Base group 1′. Single alternant.
Base group 2′. Second alternant before suffix 2.
Base group 3′. Second alternant before suffixes 1 and 2.
Base group 4′. Second alternant before suffixes 1 and 5; third and fourth alternants before suffix 2; fifth alternant before suffix 6.
Base group 5′. Second alternant before suffix 1; third and fourth alternants before suffix 2; fifth alternant before suffix 5; sixth alternant before suffix 6.

Accordingly, we list our eleven atonic verbs as follows:

be D′4′ /bi, 0, wəz, wər, wə/.
be D′5′ /bi, 0, wəz, wər, ə, wə/.
can B′2′ /kən ∿ kəŋ, kə/.
do C′1′ /də/.
have A′3′ /həv, hə/.
have A′3′ /əv, ə/.
have A′3′ /v, 0/.
shall B′2′ /šəl, šə/.
will B′2′ /wəl, wə/.
will B′2′ /əl, ə/.
will B′2′ /l, 0/.

[[The foregoing paper by Bloch and the following one by Nida, taken together, could easily lead us into far more extensive discussion than I can find room for. I content myself with a few hints; interested readers will find others in the texts.

This is no one-dimensional opposition between an A and its not-A. There is, rather, something like a dramatic encounter. I don't want this to be taken in the sense of a polemic or personal conflict between the two writers; I intend it to refer rather to ideas and to modes of thought. These are, between the two sides, as multiply different as could be imagined, or more so. Further, as if to make the whole thing all but inaccessible to discussion, there is even a total difference in the kind of participation or involvement in the encounter.

In principle, Bloch's procedure is to adopt a set of axioms and then to develop the consequences of the set à outrance. (In another capacity, that of the interested bystander, he may express something like dismay at this or that outré result; but that has just exactly nothing to do with his main business.) This paper, as published, does not group all the axioms first and then present the development in strictly logical sequence. But the difference between that and what he actually does is merely literary. His interspersed arrangement is intended to hold the reader's interest, which it adequately does. (This, incidentally, is one of the leading characteristics of Bloch's papers. In its own right, it is a great merit. But it can result in his being misapprehended by any reader of a different cast of mind.)

The ultimate justification for such a proceeding would have to be an understanding that other writers are expected to adopt other sets of axioms, and to develop their private sets with equal ruthlessness. As such publications accumulate, practical descriptions (cf. fn. 26 above) may be expected to improve gradually, as their makers (any one of whom may be one of those theorists, acting now in another capacity) gradually learn from the various theoretical papers (their own and others) how to avoid inconsistency while working to satisfy other requirements, such as pedagogical effectiveness or easy reading. There is an assumption concealed here, to the effect that a division of functions of this sort is at least one of the better ways to procure progress in theoretical and practical work both. I don't mean that Bloch and the others who behave like this do so in consequence of having adopted this assumption as one of their axioms; rather, I suppose they do it because it feels right to each of them, and my 'assumption' is just a postulate that I think I might well use if I were to construct A Set of Postulates for the Scientific Study of Linguistic Thought. Anyhow, I am inclined to feel that this assumption is sound; but I leave to others the task of trying to show that Nida did wrong to depart from it. I don't feel that he did wrong in doing so; he was simply being himself, which is the only way to make one's own contribution.

Bloch as reader of Nida's paper, in turn, couldn't help but sort out the one from the other. From one of the sorts of things he found there, he got what he had in mind, I suppose, when he recently said to me that after reading Nida's paper he would do some things differently if he were to rewrite the English verb paper. What these would be, I must not try to guess. Whether, if he were to do another pedagogical book on Japanese, he would, after reading Nida's paper, do some things differently, is a separate question.]]

THE IDENTIFICATION OF MORPHEMES

EUGENE A. NIDA
Language 24.414–41—1948

1. Recent developments and publications[1] concerning the analysis and classification of morphemes have been of great importance: they have provided a variety of new approaches to old problems, refined the statements of relationships, demonstrated the applicability of the method to a number of languages, emphasized the importance of distributional criteria, and pointed out a number of unsolved or perhaps only partially solved problems. All this has been a genuine gain. Nevertheless, a close examination of some of these developments is rather disturbing, and certain rather basic objections can be raised. For one thing, the term morpheme has come to be used by Hockett, and to a somewhat lesser extent by others, as a designation for forms which are related almost wholly in terms of distribution. I do not wish to deny the significance of tactics in linguistic structure; but when the primary criterion for the choosing of alternative possibilities is stated by Hockett as being 'tactical simplicity',[2] I am inclined to doubt the validity of quite so much stress being laid upon distributional characteristics, especially when such a choice leads to the conclusions which Hockett has described. In both Bloch's and Hockett's recent papers on morphemic problems there is a conspicuous tendency to make covert distinctions more important than overt ones, though the treatment is quite different as regards detail.

I am in complete agreement with the desire to attain simplicity of statement and to reveal the whole structure of a language in terms of such statements; but it is possible that too much emphasis on structural simplicity, both of statement and language structure, can give a false impression of simplicity and can actually misrepresent some of the pertinent facts. The objections which this paper raises are not directed against any lack of logic in the presentations cited. Bloch and Hockett have been thoroughly consistent in following to the end the implications of their basic premises. Any disagreement with their treatments is not a criticism of their handling of the data as such, but is directed rather at the fundamental principles upon which all of us as descriptive linguists have been working. Some linguists have become increasingly aware of certain of the peripheral but methodologically fundamental problems of our descriptive science, and we are genuinely indebted to Hockett, Bloch, Harris, Wells, Voegelin, and others for pointing out the ramifications of our principles. It would seem, however, that if the principles upon which we operate lead us to the conclusions adopted in the recent articles by Bloch and Hockett, there should be some refining of statement, or perhaps complete change in statement. Even Bloch does not seem particularly pleased about some of his conclusions. He says, (Lang. 23.417 ((253))), 'Reluctant as we may be to allow the multiplication of elements, we cannot escape the conclusion ...'

We shall attempt to treat in §2 some of the specific problems raised by Bloch's and Hockett's papers. These will constitute a background for the development of certain principles of analysis and classification, considered in §3. A fourth section will describe the types of morphemes.

2. We shall not attempt to treat systematically all the types of problems introduced by Bloch and Hockett, nor try to give any overall criticism. Rather, we shall restrict ourselves to considering a few examples of where their systems would seem to break down, or where there appear to be other factors which must be considered.

2.1 Bloch points out (Lang. 23.406 ((247))) that in accordance with his system the stem morphemes of shown /šown/ and showed /šowd/ must be considered as different morphemes. Since he considers that the suffixal morphemes /-n/ and /-d/ are most conveniently treated as alternants of the same morpheme and hence do not differ in meaning, and since the total forms shown and showed are not in free variation, the difference in meaning must be attributed to a difference in the stems. Bloch states the problem as follows:

Since we wish to identify /n/ and /d/ as alternants of the same morpheme, we assumed that the /šow/ that precedes /n/ is morphemically different from the /šow/ that precedes /d/. In other words, we we choose to set up two different but homonymous morphemes /šow/, and to refer to them—rather than to the suffix alternants—the stylistic or connotative difference between the inflected forms shown and showed.

[1] Zellig S. Harris, Morpheme alternants in linguistic analysis, Lang. 18.169–80 (1942) ((109)); id., Discontinuous morphemes, Lang. 21.121–7 (1945); id., From morpheme to utterance, Lang. 22.161–83 (1946) ((142)); Rulon S. Wells, Immediate constituents, Lang. 23. 81–117 (1947) ((186)); C. F. Voegelin, A problem in morpheme alternants and their distribution, Lang. 23.245–54 (1947); Charles F. Hockett, Problems in morphemic analysis, Lang. 23.321–43 (1947) ((229)); Bernard Bloch, English verb inflection, Lang. 23. 399–418 (1947) ((243)).

[2] Op. cit. 342 ((241)).

This type of procedure results in attributing a distinction in meaning to formally identical items, rather than to items which are formally different. This also means assigning a meaning-difference to certain covert elements rather than to overt distinctions. It would seem that the difference in connotation between these forms when they are not in complete free variation is precisely the type of difference which exists in the range of distribution of the suffixes /-n/ and /-d/. A procedure to treat this type of problem will concern us in §3.4.

2.2 Bloch employs a similar treatment in the analysis of sang, the past tense of sing (Lang. 23.407 ((247))). The past tense form /sæŋ/ is treated as an alternant of /siŋ/. The meaning-difference is considered as expressed by a zero suffix. By this procedure an overt distinction—the replacement of /i/ by /æ/[3]—is treated as meaningless, while the covert distinction becomes the meaning-carrier. I do not deny the significance of zero in such a form, nor the importance of the pattern which leads one to recognize a zero; but it appears to me as strikingly contradictory to treat overt distinctions as meaningless and covert distinctions as meaningful. Undoubtedly the unsophisticated speaker of the language has no such reaction, and though I do not argue for following the judgment of such a person, I do insist that we should not disguise features which, as as far as the native speaker is concerned, bear all the evidence of being meaningful and distinctive. If we do so, we have given entirely too high a priority to the arrangements of items (i. e. the tactics).

2.3 A further problem is introduced by Bloch in his treatment of were (Lang. 23.408 ((248))). The two forms was and were are treated as morpheme alternants, both occurring with zero alternant of the past tense suffix. The were which occurs in conditional clauses of an 'unreal' type is considered to consist of an alternant /wə-/[4] (an alternant of /biy/) plus the suffix /-r/. Bloch insists that this conditional were is different in meaning from the past-tense were and contrasts with it in distribution. As evidence he cites the following: If I were rude, I'd apologize; If he were here, he'd see it; and If I was rude, I apologize; If he was here, he saw it. It is true that within the if clause there is contrast of occurrence with identical constituents, but one should not overlook the fact that within the apodoses there is a contrastive use of 'd apologize and 'd see vs. apologize and saw. Hence, it can be stated that was and were are still in complementary distribution on the basis of the forms occurring elsewhere in the sentence. The so-called two forms of were, the past tense and the unreal, do bear a certain phonetic-semantic resemblance to each other (following Bloomfield's definition)[5] to which the treatment by Bloch does not appear to give full recognition. The methodological problem is that the past-tense forms was and were, which are formally different, become morphemically identical, while two identical forms, past-tense were and unreal were, become morphemically different.

What seems to be basic in Bloch's development of this situation is that were occurs with different meanings in different situations. Note, however, that in the clause If they were there, we have no idea whether this is an unreal or a simple past tense. It is the choice of the apodosis that provides the clue. If the apodosis is they would do it, we conclude that the were is unreal, but if the apodosis is they did it, then we know that were is a simple past-tense form. We should not be surprised to find forms with different meanings in different syntactic combinations, e. g. run in they run away and they run him away. The meaning in the intransitive and transitive constructions is quite different, and this contrast is paralleled by a difference in distribution of occurrence. Hence, we conclude that both run's are the same morpheme.

2.4 Hockett states that 'some bits of phonemic material, of some utterances, are morphemically irrelevant'.[6] The problem involved is the familiar one: the determination whether a structural unit which has no definable meaning in terms of a biosocial context is a morpheme, or whether such a unit is to be assigned to some other element, instead of being given independent status. Hockett considers, for example, that the stem vowels in the Spanish verb are 'empty morphs.' As empty morphs, they do not belong to any morpheme. That is to say, the vowels a, e, and i of amar 'to love', beber 'to drink', and vivir 'to live' are meaningless and hence are not morphemes.

It is perfectly true that these stem formatives cannot be given a meaning which corresponds to anything in the practical world. Their meaning, if they have any, must be stated purely in terms of grammatical situations. But the inability of the linguist to assign a non-linguistic meaning to a form is not, it seems, a sufficient basis for regarding such a structural unit as not constituting a morpheme. Otherwise, one may be faced with an amazing assortment of empty morphs which Hockett would speak of as morphemically irrelevant. On the other hand, it is rather difficult to see just how the infinitive suffix -r, which Hockett describes as a morpheme, is to be regarded as having more than grammatical meaning, i. e. meaning in terms of distributional situations. It is possible that such a line of distinction could be drawn; but the establishment of empty morphs on the basis of so-called meaning-deficiency would certainly produce quite a change in the traditional treatment of primary derivatives. For example, the -er in such words as spider, hammer, otter, badger, ladder, auger (there are more than 100 nouns in English with this suffixal formative) would be ruled out as being morphemically irrelevant—unless, perhaps, some way could be found to construe this -er as an affix similar to the Spanish infinitive ending. In any case, the meaning is essentially one which is derived from the distribution of the form in the language.

Rather than postulate structural units whose position and distribution must be described (if the description of the language is complete), but which are supposedly morphemically irrelevant, it would seem to me much safer, as well as structurally more valid, to follow Bloomfield in contending that 'every complex form is made up entirely of morphemes'.[7]

2.5 In treating so-called portmanteau items (Lang. 23.333 ff. ((236 ff.))), Hockett has introduced an interesting solution to matters which have caused complication for some time. Hockett's treatment of a form such as men differs considerably in outward form from Bloch's, but is essentially the same. Hockett would not say that men is a morphemic alternant of man plus a zero alternant of the plural suffix, but rather that men is one morph, which belongs simultaneously to two morphemes, namely ⟨man⟩ and ⟨-s⟩. Except for the type of statement, such a description does not differ appreciably from Bloch's. It is only that Hockett's appears to employ zero less frequently.

In handling inflexional endings like the -o in Spanish amo 'I love' (Lang. 23.338 ((239))), Hockett says that this is a morph which belongs to two morphemes: a tense-mode morpheme and a person-number morpheme. In this instance, however, he does not provide any overt form for the separate morphemes. Only the morph has overt form; the morphemes are merely structural parallelisms as indicated by such related paradigmatic forms as amábamos 'we loved' and amaréis 'you (pl.) will love', in which the tense-mode morpheme precedes the person-number morpheme. In the instance of French au 'to the' occurring in place of the analogically demonstrable but non-occurring à le, there is more basis for speaking of a single morph belonging to two morphemes, since these morphemes may be identified as occurring elsewhere in the language. But in the case of Spanish -o, Hockett's morphemes are only statements of sequence evidenced at other points in the paradigmatic series; there are no overt forms to which this particular morph may be referred.

2.6 The complications which arise from the portmanteau concept are extreme. For example, Hockett combines she /šiy/ and her (object form) /hər/ as two morphs in complementary distribution, and assigns them both to the morpheme ⟨she⟩. Just why the subject rather than the object form should have been chosen as the basic form of the morpheme, we are not told. It is possible that the responsibility lies in an unconscious recognition of the series: he, him; they, them; who, whom. In Hockett's system the parallel development of these three sets of forms is not considered significant, even though the final /-m/'s do bear a strong phonetic-semantic resemblance to each other. Hockett considers that complementary distribution is all that fundamentally counts. But the implications of this method seem to be greater than he may have anticipated. For example, consider what could be done on that basis with number distinctions in English. Not only would it be possible to combine all the plural affixes of nouns in one morpheme (a step which we should all agree to), but one could say that these are in complementary distribution with the partly homophonous third-singular suffix of verbs. A single morpheme could then be set up with the meaning 'number distinctiveness' and with the additional distributional characteristic that if an alternant occurs after the noun it does not occur after the verb, and vice versa, e. g. /ðə boyz rən/ the boys run vs. /ðə boy rənz/ the boy runs. By slight extensions it might be possible to construct a descriptive system by which practically all the features of concord, government, and cross-reference could be treated on a submorphemic level. If this were done, we should only have succeeded in changing the meaning of the word 'morpheme' to apply to certain distributionally related forms.

If distribution should be taken as the all-important factor, we could invade the traditional lexical area with new statements of relationships. We could say that -ology is a morphemic alternant of science, since they are in complementary distribution in such series as biology : life-science, geology : earth-science. (The bound form -ology and the free form science would never occur in identical environments.) Since the meanings of the two forms are closely related (we admittedly have no scale for measuring degrees of difference), we should have no objection to relating this series, except perhaps Hockett's caveat that 'the total range should not be unique' (Lang. 23.331 ((234))).

2.7 The recent attempts to refine morphological methodology have, nevertheless, been exceedingly valuable and important; for some phases of Bloomfield's treatment have caused difficulty, primarily because of the apparent confusion as to the status of some types of changes. The distinction between processes and morphemes is not always clear, and it is sometimes hard to know when a change is to be considered as independently meaningful and hence as constituting a morpheme. The refinements of methodology which are suggested in this paper do not attempt to contradict Bloomfield's basic analysis, but rather to refine the statements and point out the specific areas of difficulty, with suggestions of how such problems can and should be treated.

[3] [The use of the word 'replacement' evades the objection raised in Lang. 23.400 ((244)).—BB]

[4] The phonemic writing of English is here adapted to conform to my own speech, in order that all citations of English in this article may be consistent.

[5] Leonard Bloomfield, Language 161 (New York, 1933).

[6] Op. cit. 331 ((235)).

[7] Op. cit. 162.

3. The analysis of morphemes proposed here[8] takes as a basic definition Bloomfield's statements:[9] 'a linguistic form which bears no partial phonetic-semantic resemblance to any other form, is a... morpheme', and 'any combination of phonemes... which has a meaning, is a linguistic form.' It should be noted that this definition is essentially a negative one, in that it prescribes limits to the morpheme by setting it off from other morphemes, rather than by defining the nature and relationship of its own parts. Its lack of phonetic-semantic resemblance to other morphemes is the essential characteristic, not necessarily its possession of an exact phonetic-semantic identity in all of its occurrences. It is by virtue of such a negative definition of the morpheme that we can include as members of one morpheme structural units which are not identical in form. For example, we admit as morphemic alternants such a series as /-əz ∿ -z ∿ -s/, plural formatives with nouns (e.g. in roses, cows, sticks). The distribution of these alternants may be phonologically defined. We also speak of /-ən/ in oxen as an alternant of the plural suffix, but in this instance we recognize that the distribution cannot be defined in terms of phonological criteria, but rather that the distribution of /-ən/ can only be described by identifying the morpheme ox. The essential difference between these two types of alternants has been recognized by everyone; in his recent article, Bloch makes a special point of distinguishing the two types of alternants (Lang. 23.402 ((244 f.)). There is not, however, complete agreement on the degree of difference which can be admitted in morphemic alternants, nor on the types of distribution which are significant in determining patterns of morphemic alternation. For that reason, we need more precise clarification of our procedure.

Before treating such details of method, it should be pointed out that Bloomfield's definition, which we are adopting here as basic, provides not only for sub-morphemic distinctions of phonetic form but also for sub-morphemic distinctions of semantic value. This will be more evident as we consider Principles 4 and 6 below. A definition which contrasts one phonetic-semantic entity with all other entities in the language still permits the sub-morphemic distinctions of phonetic form and semantic areas within the basic distinctiveness which sets off such a form from other possibly related forms. An understanding of this essentially negative character of the underlying definition is necessary for a proper evaluation and analysis of the following principles.

These principles of analysis are not here numbered by any system of priority, nor necessarily by any system of logical delimitation of the field. Rather, they are arranged as a sort of related series of working postulates, beginning with the more general and readily accepted and advancing toward the more specific and complex.

3.01 Principle 1. Forms which possess a common semantic distinctiveness[10] and an identical form in all their occurrences constitute a single morpheme.

With this principle no one can disagree, for it means that such a suffix as /-ər/ in dancer, walker, runner, worker, swimmer, etc., constitutes a morpheme, since it has a single form (i.e. consists of the same phonemes in the same order), and since in each instance it possesses a common semantic distinctiveness, which may be defined as 'agentive' with respect to the underlying verb form. Note, however, that this first principle of analytical procedure is not a definition of a morpheme; rather, within the scope of the basic definition, such a principle of identification holds true.

3.02 Principle 2. Forms which possess a common semantic distinctiveness but which differ in phonemic form (i.e. constituency or shape) constitute a single morpheme provided that the distribution of formal differences can be phonologically defined.

Principle 2 provides for the identification of phonologically defined alternants as members of the same morpheme. For example, the distribution of the English plural formatives /-əz/, /-z/, and /-s/ may be described as phonologically defined: /-əz/ occurs after sibilant phonemes, /-z/ after voiced non-sibilant phonemes, and /-s/ after voiceless non-sibilant phonemes. Such a distribution is phonologically definable, and, as such, is meaningless (see Principle 6). This morphemic series may be symbolized as /-əz ∿ -z ∿ -s/.[11]

Such phonologically defined alternation is not restricted to affixes. For example, the atonic indefinite article in English has in some dialects the forms /æn/ and /ə/. The first occurs before vowels: /æn æpəl, æn orinʃ/.[12] The second occurs before consonants: /ə mæn, ə fiš/. Accordingly these morphemic alternants may be related as /æn ∿ ə/.[13]

[8] I am indebted to my colleagues Kenneth L. Pike and William L. Wonderly for many helpful suggestions made during the development of the system described here.

[9] Op. cit. 161 and 138.

[10] For a definition of 'common semantic distinctiveness' see Principle 3, fn. 15.

[11] In this paper the symbol ∿ is used only to designate phonologically defined alternation. The symbol ∞ is used to relate alternants whose distribution is defined morphologically (see Principle 3).

[12] The space left in writing between proclitics and nouns is not morphemic.

[13] Morphemic alternants can conveniently be called allomorphs. Accordingly, allomorphs are related to morphemes as allophones are related to phonemes. In the process of analyzing a language there might be occasion to use the term morph to designate a a structural unit which has not as yet been assigned to any morpheme; but in the description of a language (as distinct from the procedure of analyzing it) every structural element except features of arrangement is either a morpheme or part of a morpheme. Hence every element is also an allomorph or part of an allomorph.

In the reduplicative formations of the Greek perfect there occur forms which exhibit a considerable variety of phonemic content. Compare the following: /leluka, bebouleuka, gegraphᵃ, dedoika, tetʰeika, keklɛˑka, memɛˑna/ from stems meaning 'loose, consider, write, give, place, call, rage' respectively. Despite the phonological differences in the forms /le ∿ be ∿ ge ∿ de ∿ te ∿ ke ∿ me/, it is easy enough to predict the form by saying that the initial consonant is identical with the initial consonant of the stem, except that the consonant of the reduplicated element is unaspirated even if the initial consonant of the stem is aspirated (Grassman's law). This type of morpheme we may set up as /C_1e/ by defining C_1 as any consonant which is first in the underlying form, with the restriction that in potential sequences of aspirated consonants the first becomes unaspirated.[14]

In the above illustrations the difference of form is primarily one of phonemic constituency. It is also possible that the same phonemes may occur in related forms, but in a different order. If, for example, a suffix -an is found to occur after consonant-final stems and -na after vowel-final stems, these two related forms can be combined into the same morpheme—provided, of course, that they possess a common semantic distinctiveness. Differences of form may therefore be treated as affecting either the constituency of phonemes or their arrangement. This type of arrangement we may speak of as the phonological shape of the form.

3.03 Principle 3. Forms which possess a common semantic distinctiveness,[15] but which differ in their phonemic form in such a way that the distribution of the forms cannot be phonologically defined, constitute a single morpheme if the forms are in complementary distribution, subject to the following restrictions:

(1) Forms that occur in complementary distribution may occur in tactically identical environments or in tactically different environments.[16] Complementary distribution in tactically different environments constitutes a basis for combining different forms into one morpheme only on the following condition: that some other morpheme—belonging to the same distribution-class, and having either a single phonemic shape or phonologically defined alternant shapes—occurs in all the tactically different environments where the forms in question are found.

(2) In determining the morphemic status of forms, tactically identical environments take precedence over tactically different environments.

(3) In determining the morphemic status of forms, the immediate tactical environment takes precedence over the larger (non-immediate) environment.

(4) Different forms that contrast in identical distributional environments may nevertheless be assigned to the same morpheme, if the difference in meaning between them is identical with the difference in meaning derived from the distribution of the related forms.[17]

By the application of Principle 3 we are able to combine such forms as /-əz ∿ -z ∿ -s/ in axes, hills, lips with the /-ən/ of oxen, even though the distribution of the alternants cannot be stated in terms of phonological environments. Only the identification of the morpheme (or morphemes) with which the suffixal elements form a construction can suffice to tell us which collocations occur. Since the morphemes must be identified, we may say that the distribution is morphologically defined. A zero alternant also occurs in this construction, e.g. in sheep, trout, elk, salmon. Principle 3 permits us to combine into a single morpheme at least three such alternants, and since the alternation is morphologically defined we may symbolize the series as /(-əz ∿ -z ∿ -s) ∞ -ən ∞ 0/.

[14] In a descriptive grammar of Classical Greek such a statement concerning the potential sequences of aspirates would be included in the morphophonemic section, and would apply to all situations described in later sections where such sequences might arise.

[15] To say that forms possess 'a common semantic distinctiveness' does not mean that all the occurrences of these forms must have an identical meaning, but rather that they possess some semantic feature in common which remains constant in all their meanings and which sets them apart from all other forms in the language. The phrase 'a common semantic distinctiveness' is essentially a negative formulation: it implies merely that the forms which possess that distinctiveness are in semantic contrast with other forms. It is this definition by exclusion that makes it possible to deal with certain sub-morphemic semantic differences, which will be treated under Principles 5–7.

[16] It is necessary to distinguish between 'immediate' and 'non-immediate' tactical environments, and between 'identical' and 'different' tactical environments. The distinctions overlap. In the expressions this man and that boy, the forms this and that occur in identical immediate tactical environments. In the sentence This man hit that boy, the immediate tactical environments of this and that are identical, but the non-immediate tactical environments are different. In the expressions for this man and by that boy, both the immediate and the non-immediate tactical environments of this and that are identical; but in the sentence That boy worked for this man, the forms this and that again occur in different non-immediate tactical environments. In the sentences They left me and They left late, the non-immediate tactical environments of me and late are identical (both occur in subject-predicate constructions), but their immediate tactical environments are different: me occurs in a verb-goal construction, late occurs in a construction of verb plus temporal attributive. For any one occurrence of a morpheme there is usually only one immediate tactical environment; but the number of its non-immediate tactical environments is limited only by the number of sets within sets of immediate constituents in which that morpheme is included. Hence immediateness of tactical environment is determined by the analysis into immediate constituents.

[17] Restriction 5 is covered by a separate principle (§3.07).

These three alternants of the plural morpheme contrast completely in phonological form; they are what Bloomfield calls 'suppletive alternants' (Language 215). Often, however, the morphologically defined alternants bear a strong phonetic resemblance to the underlying form. Compare, for example, the following sets: wife /wayf/ : wives /wayvz/; path /pæθ/ : paths /pæðz/; house /haws/ : houses /hawzəz/. The bound stems which occur in the plural formations resemble the singular except for a change from a voiceless to a voiced phoneme. This series of alternant forms cannot be phonologically defined, since there are other, phonologically similar forms which do not exhibit the same alternants, e. g. in cliffs, myths, classes. It is needless to attempt to symbolize the exact degree of phonological difference between related forms; the definable basis of distribution is more important. For example, the atonic forms of the indefinite article an /æn/ and a /ə/ have no phoneme in common, but the alternants occur in phonologically definable environments; the stem forms /wayf/ and /wayv-/ are only slightly different in phonological form, but the distribution of the alternants must be defined in terms of morphemes.

The application of Principle 3 is not particularly difficult as long as there is but one immediate tactical environment, i. e. one construction. When, however, the complementation occurs in more than one tactical environment, i.e. in more than one construction or set of constructions, the problems are much more complex. For example, the forms I and me generally occur in complementation:[18] I occurs in preverbal subject position, me in postverbal object position and after prepositions.

The situation in all languages known to me is so complicated by a number of other factors that it seems preferable, for the initial statement of the principles here described, to construct an artificial set of forms, and on the basis of these to point out the methodological implications in actual languages.

Let us suppose that we encounter the following set of pronouns:

	Subjective	Objective
1st person	na	fi
2d person	so	ka
3d person animate	ri	po

If these pronouns occur in complete complementation (e. g. the subjective set preceding verbs and the objective following verbs), we should perhaps be inclined to consider the related forms as morphemic alternants, or allomorphs. There would, however, be a basic objection to such an analysis: we could not know whether the tactical arrangement of order was itself structurally meaningful, since there are two contrastive features relating these forms to the verb, one of order and one of form. It would thus be impossible for us to describe the related forms in these two series as allomorphs unless we were able to demonstrate that a morpheme of the same class possessing only one allomorph or phonologically defined allomorphs could exist in the same environments and that hence the meaningful relationship of the parts could be indicated solely by the tactical arrangement. Thus, for the sake of developing our procedure, let us assume that we find a 3d-person-inanimate pronoun with the form go occurring in both subject and object relationships and in the same tactical environments as the subjective and objective pronominal forms. This circumstance would provide evidence of the grammatical significance of the tactical arrangement.

The situation in English is somewhat parallel. The forms I, we, she stand in complementary distribution (except for fluctuations considered under Principle 7) with me, us, and her respectively. There is also one pronoun you, which occurs in all the tactical environments of the formally contrastive set. If these were the only forms, we could combine the forms I ∞ me, we ∞ us, and she ∞ her as allomorphs in accordance with the second restriction on distribution stated under Principle 3. The situation would be completely analogous to the hypothetical pattern just described. But the situation in English is complicated by the fact that three pronouns have other types of contrasts, namely he : him, they : them, who : whom.

The forms him, them, and whom bear a partial phonetic-semantic resemblance to each other, and we are thus obliged to regard -m as a morpheme (even though its distribution is very limited). On the basis of this pattern, it is possible (and perhaps obligatory) to postulate a zero objective suffix for the other pronominal forms. But if this is the case, we must follow the application of Principle 4, which requires the recognition of overt differences in such situations as being morphemically significant, hence morphemes.[19]

In elucidating the second and third restrictions I shall use another hypothetical illustration. Let us assume that nouns occur in the subject case-form with the suffix -ma and in the object case-form with the suffix -li,[20] and that the distribution is precisely the same as in the previous hypothetical example, namely that subject forms always precede the verb and object forms always follow it. As in the previous problem, we cannot know whether the meaningful

18 This complementation holds good except for such phrases as It's me : It's I, for you and me : for you and I, which are treated as special problems under Principle 7.

19 The relations between I : me, we : us, she : her is not the same as the relation between any set of six completely unrelated morphemes in the language. Rather, me, us, and her constitute a type of replacive morpheme (see §4.32).

20 These hypothetical illustrations are intentionally extreme, since they are designed to point up the theoretical limits of the restriction.

element is the form of the nouns or their position in the construction. Nevertheless, we regard the suffixes -ma and -li as morphemes, since they contrast in the identical and immediate tactical environment in which they occur, namely as suffixes to nouns. The significance of restrictions 2 and 3 is simply that a contrast in an identical or immediate tactical environment cannot be controverted by complementation in non-identical or non-immediate tactical environments.

In practice this principle of distributional environments means that in such a language as German we may establish the morphemic value of the case endings quite apart from such differences of meaning as in dem Tempel 'in the temple' and in den Tempel 'into the temple'. If there were no prepositions which occurred with both dative and accusative case-forms, and if, in all situations in which the dative and accusative case-forms were attributive to verbs, there existed a similar complete complementation, we should still consider the dative -m as a morpheme in contrast to the accusative -n, because in the identical and immediate tactical environments there is a contrast, namely den : dem 'the', ihn : ihm 'him', meinen : meinem 'my'.

To illustrate the implications of these principles, I turn to a Huichol problem.[21] The second singular pronominal reference is indicated by the following forms: a- preceding nouns in a possessive construction (aki 'your house'), pe- preceding verbs as subject of an intransitive (peptikuye 'you are sick'), and -ma- as object of a transitive verb (pəmatsireiya 'he sees you'). On the basis of such data alone, it is impossible to determine whether the forms a-, pe-, and -ma- should be considered a single morpheme. It may be that this difference of form is essential to the defining of the relationships of constituent parts of the constructions. The three relationships may be stated as follows: (1) 'A (the prefix a-) possesses B', (2) 'A (the prefix pe-) is the actor of the action B', and (3) 'A (the prefix -ma-) is the goal of the action B'. If all the pronominal elements exhibited similar sets of formally different forms, we should be forced to conclude that for the defining of the various relationships not only a difference in tactics (tactical environment) is necessary, but also a difference in the form of the pronominal elements. This, however, is not the case. The first person pronominal element ne- occurs in all positions, and the differences of meaning are attributable purely to the tactics, e. g. neki 'my house', neptikuye 'I am sick', and pənetsireiya 'he sees me'. Accordingly, we conclude from the patterns of the two series that there are two morphemes: ne- first person singular pronoun, and a- ∞ pe- ∞ -ma- second person singular pronoun.

It is possible, however, to take an exactly opposite position. Bloomfield says (224), 'The existence of even a single over-differentiated paradigm implies homonymy in the regular paradigms'. In analyzing the structure of a language such a principle can be helpful. For example, we usually say that in the imperative expression Run away! the form of the verb is 'infinitive' because we find a distinctive infinitive form in the over-differentiated verb to be, e.g. Be good![22] A strict application of this principle of over-differentiation, however, means that the structural analysis is 'tied' to the most diverse series, and in practice this would involve the recognition and description of a tremendous amount of homonymy. Though the logic of Bloomfield's statement cannot be denied, I have not followed it because it seems to violate the basic functional operation of the language. Productive patterns, analogical constructions and the structurally principal forms and combinations of forms do not appear to be based on the most diverse paradigmatic set or sets. It is possible that the best solution to this dichotomy of form vs. distribution lies somewhere between the poles which I set up; but the defining of such a mid-point is extremely difficult, if not impossible. I have, however, introduced one important restriction to the autonomy of distribution: namely, that before complementary distribution in different tactical environments is valid for uniting different forms into the same morpheme, it must be demonstrated that at least one homophonous set of forms belonging to the same distribution-class occurs in all such positions.

This restriction on different tactical environments also becomes meaningful in limiting the types of semantically related forms which one may combine into the same morpheme. For example, in English the nominal prefix ex- in ex-president, ex-chairman, ex-official indicates past time. Similarly the morpheme /-əd ∿ -d ∿ -t/ indicates past time with verbs: accepted, turned, walked. These morphemes are in complementary distribution in that they occur in different constructions.[23] If it were not for the first restriction on Principle 3, they could be combined as alternants of the same morpheme. But in this instance, there is no set of identical forms which may occur as prefixes to nouns and suffixes to verbs and which in both positions define a tense relationship. This is the situation in Makushi,[24] a

21 I am indebted to John McIntosh of the Summer Institute of Linguistics for this illustrative material.

22 Note, however, that Bloch (Lang. 23.401 ((244)) does not use this type of analysis, but combines all formally identical items.

23 There is admittedly a certain difference in meaning between ex- and the past-tense morpheme of verbs. But since our science does not provide us with an adequate measuring-rod for handling such differences, and since our basic definition of the morpheme is primarily negative and admits of sub-morphemic distinctions, my main reason for refusing to combine ex- in one morpheme with /-əd ∿ -d ∿ -t/ is distributional.

24 For the Makushi data I am indebted to Neill Hawkins, a research fellow of the University of Oklahoma and the Summer Institute of Linguistics.

language spoken on the borders of Brazil and British Guiana. In Makushi the regular past-tense suffix is -pï: wïiya 'he kills', wïpïiya 'he killed'. This same -pï occurs as a derivative suffix with noun expressions: to the phrase a?nay ipu 'ear of corn' (a?nay 'corn' and ipu 'its foot') one can add the suffix -pï to produce a?nay-ipupï 'corn cob'—that is to say, an ear of corn in the past tense. It is true that Makushi lacks other tense formatives with the same distribution, but this occurrence of -pï could establish a pattern which would permit the combining of forms exhibiting the same distribution even though they differ in shape in such a way that the occurrence of the alternants could not be phonologically defined.

In describing and identifying morphemic alternants (allomorphs) it is frequently convenient to distinguish two types: basic and non-basic. This distinction may include phonologically defined allomorphs or morphologically defined allomorphs or both. In the series /-əz ∿ -z ∿ -s/, whose distribution is phonologically definable, it is possible to select one alternant as basic and the others as phonologically 'derived'. There are three types of criteria, in the following order of importance: (1) parallel structure, (2) general patterns of morphophonemic change, and (3) limitation of distribution.

The selection of a basic alternant or allomorph of the set /-əz ∿ -z ∿ -s/ may depend on the parallel structure in such atonic forms as /-iz ∿ -z ∿ -s/ for is (Bess's gone : Bill's gone : Dick's gone) and /-əz ∿ -z ∿ -s/ for has (Bess's done it : Bill's done it : Dick's done it). In these instances the fuller forms /-iz/ and /-əz/ are obviously closer to the full word.

The second criterion, involving phonological change, is directly supplementary to the first. By referring to general patterns of morphophonemic change we can explain the loss of a vowel more easily than the development of a vowel, although in some languages developed vowels might be frequent in certain phonological sequences. In the English forms just cited, the choice of /-z/ as a basic form would not permit us to predict the vowel of the longer form, while the choice of the fuller form permits us to make a completely regular statement for all the changes that occur. In general, there a number of phonological principles which may guide our choice, e. g. assimilation, palatalization, reduction of clusters, Verner's phenomenon. The choice of a basic form from which we may derive the alternants by frequently observed phonological developments (or sets of correlations—speaking on a purely synchronic level) constitutes a much sounder principle of procedure than to assume for a language all sorts of rare or unique patterns of phonological change or alternation.

The third criterion is based on patterns of limitation in the distribution of forms. Thus, if one allomorph of a morpheme occurs after all but three morphemes of a certain class, and another allomorph occurs only after those three, we should accept as basic the allomorph with the wider distribution. It is the observable structural patterns of a particular language, not its general phonological tendencies or the mere statistics of occurrence, that constitute the chief criterion for any given analysis. Accordingly, in English there is no difficulty in establishing the alternant /-əz/ as basic.[25] Following Harris, Hockett, and Bloch, we can symbolize the whole morpheme (including all its allomorphs) as ⟨-əz⟩.

To determine the basic alternant among allomorphs whose distribution is morphologically defined, the productiveness of a given type of formation is of prime importance as a guide in applying our third criterion. Two types of synchronic data are significant in comparing the productivity of different formations: their relative frequency of occurrence and their relative regularity. The two are of course interrelated. Allomorphs that occur in a larger number of combinations and in a larger number of actual utterances are likely to illustrate the productive patterns; but regularity of formation is important also. If two alternants have approximately the same frequency of occurrence, but if one of them combines usually with forms that exhibit considerable modification of their underlying forms, while the other combines with unmodified forms, the latter is likely to illustrate the more productive pattern, and—other things being equal—should be chosen as the basic alternant.

3.04 Principle 4. An overt formal difference among related forms (forms containing recurrent partials or occurring in complementary distribution) constitutes a morpheme, if in any of these forms this difference, together with a zero tactical difference, is the only significant feature for establishing a minimal unit of phonetic-semantic distinctiveness.

Principle 4 means that in the case of /sæŋ/, where the only differences to contrast it with /siŋ/ are a covert zero alternant of ⟨-əd⟩ and the overt replacement of /i/ by /æ/, we shall consider the overt replacement of the vowels as constituting a morpheme (see §2.2). Zero features are covert features, but minus features (or subtractives, as I prefer to call them) are not covert. A minus feature is overt; compare for instance the French feminine form platte /plat/ 'flat' with the masculine plat /pla/, which may be described as derived from the feminine by subtraction.[26]

[25] This type of pattern furnishes a reason for choosing /-əd/ as the basic alternant in the series /-əd ∿ -d ∿ -t/, the past-tense morpheme of English verbs. Compare also the parallel atonic alternants of would and had.

[26] Bloomfield, op. cit. 217.

Principle 4 does not rule out the significance of zero, though it is hard to determine how extensive a pattern must be to demand the introduction of zero. Such problems are too much a matter of proportion, balance, symmetry, pattern congruence, and the like—features closely allied to one's esthetic feeling. It is for this reason that different linguists describe the same structure somewhat differently.[27] But though Principle 4 does not rule out zero, it does mean that overt formal differences have at least equal status with zero. It will mean, for example, that we cannot say that song /soŋ/ is an alternant of sing /siŋ/ with the addition of a zero derivative suffix, as Bloch does (Lang. 23.407 ((247))). In the first place, the nominalizing suffixes in English do not constitute a complementary distributional set. There are such series as receiver, receipt, receptor, reception and creator, creation, creature. Since I adopt Hockett's general principle that zero must not constitute the only allomoprh of a morpheme (Lang. 23.340 ((240))), I am obliged to assign the zero of song to some morpheme with at least one overt allomorph (if, indeed, there is a zero derivative affix), but there is no morpheme or set of allomorphs to which such a zero can be assigned. In this instance, a derivative zero allomorph proves untenable. If one wishes to describe the form-meaning relationship between sing and song as pertinent to the language (they certainly exhibit partial phonetic-semantic resemblance), the structurally significant feature is the replacement of the vowel, not the addition of zero.

On the basis of Principle 4, in such related words as breed : bred, feed : fed, meet : met, plead : pled, the replacement of /iy/ by /e/ is meaningful. This establishes the replacement /e ← iy/ (read '/e/ replaces /iy/') as a morpheme. The same replacement occurs in the related words leave : left /liyv, left/. Having established the replacement /e ← iy/ as a morpheme at one point, we cannot deny its status as a morpheme just because the usual suffix /t/ is overt. The past-tense formation of left thus includes two morphemes in addition to the stem. This occurrences of two morphemes with the structural value of one morpheme in other situations should not unduly disturb us, since the situation occurs in a number of other languages. In Greek, for example, some perfects occur both with reduplication of the initial consonant plus /e/ and with a suffix /-k-/, e. g. /leluk-/ 'have loosed'; but others have only the reduplication, e. g. /leloip-/ 'have left'.[28]

[27] Any science of classification becomes involved in the same difficulty. Thus, taxonomists in the fields of botany and zoology are faced with the same problem of determining the degree and the nature of structural likeness or difference which are to be regarded as significant in establishing classes and subclasses.

[28] The change of vowel here need not be taken as a morpheme, since it is not independently meaningful (does not occur as the only difference between related forms, or as the only difference aside from a zero difference). The element /le...k-/ could be taken as a single discontinuous morpheme (cf. Harris, Lang. 21.121–7), except for the fact that one of the two parts occurs without the other. See §4.2.

In the English past tense form /left/ there is another change which distinguishes it from its underlying form /liyv/, namely the change of /v/ to /f/. Why is not this change of consonant just as meaningful as the change of vowel? Precisely because it does not occur as the only modification (with the exception of zero) in a related series. But in describing the related forms strive and strife such a difference of consonant does constitute the 'meaning-carrier' of the semantic difference; and so, in the series strive : strife, thieve : thief, grieve : grief, the replacement of /v/ by /f/ constitutes a morpheme.[29] It would be possible, it is true, to analyze the replacement in left somewhat differently. We could, for example, describe the difference between /liyv/ and /lef-/ as /ef ← iyv/. Such a change would be paralleled by the one in /luwz/ lose and /los-/ los(t). On the other hand, an examination of the forms in which such replacements occur shows us that wherever a consonant other than postvocalic /y/ or /w/ is replaced or lost, an overt form of the morpheme ⟨-əd⟩ also occurs. This means that only the replacement of a syllabic[30] is independently meaningful (i. e. serves as the only overt difference between forms); differences of consonantal phonemes may be treated as meaningless, since wherever they occur, an overt alternant of ⟨-əd⟩ occurs also.

One other problem confronts us in treating the replacement of syllabics in English. The replacement /o ← uw/ occurs only in the form /lost/. Since this specific replacement is not independently meaningful, as is the replacement /e ← iy/ (e. g. in bred, led, fed, met), it would be possible to treat it as 'conditioned' by the occurrence of the morpheme ⟨-əd⟩. But if English past-tense formations are considered as structurally related, the pressure of the pattern would seem to be sufficient to permit the classification of changes by types: (1) the replacement of syllabics as morphemic, and (2) the replacement of consonants as sub-morphemic.

On the basis of Principle 4 we treat noun plurals in English such as men, feet, mice, teeth as occurring with 'replacives' (i.e. replacements which are

[29] The so-called direction of derivation (strive derived from strife, or strife derived from strive) depends on factors which are not discussed here. See §4.32.

[30] In this use of the term syllabic I follow Bloch (Lang. 23.414 ((251))), who says, 'we shall regard every syllabic as a unit, whether it consists of a vowel alone or of a vowel and a semivowel.' In my own English, only the postvocalic semivowels /y/ and /w/ combine with vowels to constitute single syllabics.

morphemic). Whether we are to say that both a replacive morpheme and a zero alternant of the plural suffix occur in these words depends on how much significance we attach to that part of the structural design which provides the basis for zero.

Following the same procedure we treat was /wəz/ and were /wər/ as morphemically different from be /biy/, rather than as morpheme alternants with the past-tense zero suffix, as Bloch does (Lang. 23.416 ((252))). The morphemic difference is not, however, simply that /wəz/ is one morpheme and /wər/ is another morpheme. These forms /wəz/ and /wər/ contain a recurrent partial, namely /wə-/; the contrast between /-z/ and /-r/ may be said to be meaningful, in that they occur in certain patterns of concord. The morpheme /-z/ occurs in 'realis' expressions containing first- and third-person singular subject constituents; the morpheme /-r/ occurs in all other situations. The meanings of /-z/ and /-r/ are thus essentially grammatical (see §3.6). This problem and the one involved in the treatment of were in such expressions as If John were here, he would help can be more easily understood if we examine contrasts occurring in the present-tense forms of English.

We establish the morphemic value of the suffix ⟨-əz⟩ in tries /trayz/ by such contrasts as try : tries /tray : trayz/ in boys try /boyz tray/ vs. John tries /jan trayz/ and the like; and we establish the contrast as morphemic on the basis of the immediate tactical environment, namely /trayz/ vs. /tray/ (Principle 5, restrictions 2 and 3). Otherwise, we might regard such a suffix as meaningless, on the ground that it is conditioned by the class (singular vs. plural) of the subject expression (see §3.3).

Having once established that /trayz/ occurs after third-person singular subject expressions and /tray/ occurs after all other subject expressions, we are not to be surprised to find that this contrast is not always valid; thus, try /tray/ occurs after a third-person singular subject expression in I move John try it. On the basis of our principles of procedure[31] we regard this try as identical with the try that normally occurs after all persons except the third. We have established a morpheme on the basis of a contrast in one environment, and such a contrast is valid; but we must now go farther in our description of the distribution, so as to take in all environments. The fact that try occurs with only certain persons in one type of environment but with all persons in another environment, does not negate its original contrast with tries.

We now return to a more complex problem. On the basis of contrast in simple 'realis' contexts we distinguish between /-z/ and /-r/ in the forms was and were. In the 'irrealis' context (which demands the occurrence of some potential form such as would in the principal clause), the form were occurs with all persons, and was is absent altogether. But the distributional difference in the larger environment does not invalidate the original contrast. Nor does the difference in the meaning of were in the two situations impair the analysis, since we are prepared to recognize subordinate semantic differences within the basic semantic contrastiveness, and the differences of meaning here can be related directly to the difference in environment (see §3.6 and §3.7).

3.05 Principle 5. The meaning of any form is definable in terms of the feature or features common to the situations in which the form occurs.

This means that signification is statable only in terms of environment. Certain forms are definable (or identifiable) as symbols for objects in the environment of the practical world, e. g. animal, house, brick, moon, policeman. Other forms, such as joy, peace, love, blue, good, only identify features which exist as parts of other things. Still other forms have substantially what we may call 'grammatical meaning', e. g. ⟨-əz⟩ as the third singular suffix in verbs, /-m/ as the object suffix in him, them, and whom, and the /-ər/ in hammer, ladder, spider, otter, badger, water. 'Grammatical' or 'linguistic meaning' depends on the occurrence of a form in significant situations defined only or at least primarily in terms of the linguistic environment. Many forms—especially stems—are significant in biosocial environments; but every form has linguistic meaning as well, since every form occurs in some linguistic environment.

The linguistic meaning of a morpheme can be defined in terms of several criteria: (1) the types of construction in which it occurs, (2) its frequency of occurrence, and (3) its productiveness in new combinations. The first of these are linguistically the most meaningful.

We can give part of the meaning of boy by identifying the referent (or the referent-types) for which this morpheme occurs as a symbol; but another part of the meaning of boy is the distribution of the morpheme in particular linguistic situations. The linguistic meaning of boy includes such facts as the following: boy occurs as the subject of a sentence, the object of a verb, and the second member in a prepositional phrase; it combines with derivative formatives such as -ish (boyish); and it occurs in an exclamatory phrase Oh boy!. If we disregard entirely the biosocial distinction in the meanings of boy and girl, we can still say that the linguistic meanings of these two words differ in that boy occurs in a type of exclamatory phrase from which girl is excluded. Bloomfield (op. cit. 164) recognizes that distribution is meaningful by saying: 'Selection of forms contributes a factor of meaning'. (Bloomfield's 'selection' is identical with what I here call distribution.)

31 On the basis of Bloomfield's principle of over-differentiation (see §3.03), we might consider the second try a different morpheme, since in the verb to be the regular third-person form is is but after a principal verb such as move the form instead is be.

The very frequency of occurrence of a form also contributes a feature of linguistic meaning; for it is important on the one hand in establishing patterns of analogy, and on the other in preserving the form against levelling influence. Irregular forms of frequently occurring words (e. g. the forms of the verb to be in English) resist levelling analogies which extend to other verbs of less frequent occurrence.

The productiveness or non-productiveness of a formation is similarly a meaningful morphemic feature. As descriptive linguists we are sometimes inclined to overlook such a dynamic aspect of the language. We assume that a form is a fixed feature and that its distribution is fully defined or definable. We tend to imply that productiveness is only a diachronic fact, but it may be a synchronic reality as well.[32] In syntax it is essential to recognize that the limits of the distribution of many constructions cannot be fixed, as far as the possible collocations of morphemes are concerned; for otherwise a speaker would only be able to repeat verbatim the phrases which he had previously heard. If we recognize this openness of class membership in the constituency of many syntactic constructions, we must be prepared to recognize a similar phenomenon in morphology. When we say that the distribution of a form is open (at least in terms of certain types of environment), we say simply that the form is productive—that is, it occurs with a class of forms whose membership we cannot accurately delimit. This is a synchronic fact, and also part of the distributional meaningfulness of the form in question.

3.06 Principle 6. Morpheme alternants whose distribution is not phonologically definable exhibit sub-morphemic differences of meaning.

I have already anticipated this principle in the discussion of sub-morphemic phonetic and semantic distinctions (§3). Note also Bloomfield's assumption (op. cit. 145) that there are 'no actual synonyms'—no items that are different form but absolutely identical in meaning. If it is true that 'selection of forms contributes a factor of meaning', then the different selection (i. e. distribution) of allomorphs implies that they have different meanings.

One obvious sphere in which this principle holds true is the description of productive patterns, formations which cannot be said to have a rigidly limited distribution (§3.5). Compare the allomorphs of the plural suffix, /-əz ∿ -z ∿ -s/ vs. /-ən/: the former set is productive, the latter is not. This productiveness is a distributional characteristic, and as such is meaningful.

32 Any body of material sufficiently extensive to serve adequately as a basis for descriptive analysis must include fluctuations of forms, which are symptomatic of different degrees of acceptability. Synchronic contrasts of this kind are closely allied to contrasts of productivity vs. non-productivity.

It is often extremely hard to define the differences of meaning between allomorphs, since it depends solely on a difference of distribution. But our inability to correlate this difference with biosocial contrasts does not invalidate the principle (especially Principle 5, which underlies Principle 6), nor the basic assumption that we adopt as a premise for all descriptive analysis: that there are no actual synonyms.

3.07 Principle 7. Related forms which occur in the same environment, but which could otherwise be regarded as allomorphs, can still be so regarded if (a) there is no apparent difference of meaning between them, or (b) the difference of meaning is derivable from the distribution of the related forms. Condition (a) is known as free variation.

This principle is designed to treat instances of 'overlap', i.e. forms which are in complementary distribution except at certain points where there is a contrast resulting from fluctuation of forms. Because of difficulties with such fluctuation of forms, Bloch has treated the show in shown and the show in showed as two distinct morphemes, since he had already combined /-n/ and /-d/ as morphemic alternants (Lang. 23.406 ((247))). This type of analysis has already been discussed in §2.1.

On the basis of Principle 7, we can say that the participial suffixes /-n/ and /-d/ are nevertheless allomorphs, each having its sub-morphemic distinction of meaning derived from its own distribution. Where they both occur with the underlying form /šow/, they are no longer in complementary distribution, and there is a difference of meaning; but the difference of meaning is precisely that which depends on the distribution of the two allomorphs. The allomorph /-d/ is productive: its potential distribution is not arbitrarily fixed like that of /-n/. Both allomorphs may occur in the same person's speech, but /-d/ occurs in more colloquial socio-linguistic environments. On the other hand, /-n/ has a fixed distribution, but within this fixed range of occurrence the /-n/ may be said to have greater 'socio-linguistic acceptability'. These facts are synchronic as well as diachronic; they apply to the present functioning of the language.[33]

The descriptive linguist is always faced with the problem of fluctuations of forms. The series roof : roofs : rooves and hoof : hoofs : hooves is illustrative. The suffixes conform to the regular pattern of alternation; if there is any difference in meaning, it must be associated with the stems roof- vs. roov- and hoof- vs. hoov-. Some speakers of English may

33 I admit readily that the viewpoint here expressed is different from the traditional descriptive approach; but we are forced to treat fluctuations of forms in some way within the framework of our descriptive statements. It does not seem to me wise to go to the extremes proposed by Bloch within the rigid confines of his method.

contend that these forms are in completely free variation, and they may be right. Yet on the assumption that there are no actual synonyms, we should still say that a difference of meaning exists between roofs and rooves, hoofs and hooves — though it may be very slight, and though it may affect only the connotations of the various forms.[34] The semantic difference between roofs and rooves, hoofs and hooves derives from the different distributions of unchanging noun-final /f/ and of noun-final /f/ replaced in the plural by /v/. The former is the productive type. The latter is non-productive, with a strictly limited distribution; within that range, it has greater socio-linguistic acceptability in some social groups but is considered pedantic in others.

A more complicated problem is presented by the series /hæv, həv, əv, v/ have. Bloch's treatment considers these related forms as four different morphemes. Common sense reacts to this type of analysis as being incorrect; but in terms of the conventional handling of morphemic problems, Bloch's analysis is probably inevitable.[35] The differences of meaning that exist between the various forms of the series /hæv, həv, əv, v/ are essentially connotative and reflect usage in various types of discourse, which may be described as being increasingly colloquial. That is to say, /hæv/ occurs in more 'elegant' socio-linguistic context than the form /v/, and the other forms stand somewhere between. But we must not assume that these forms are the only distinguishing characteristics of such contexts. There are supplementary differences of lexical choice, as well as many types of phonological reduction and change which mark the various 'levels' of discourse. Compare /səm mowr, səmowr/ some more; /did yuw, didyə, diJə/ did you. There are still other features, not normally included in a phonemic statement but meaningful in the characterization of socio-linguistic context: preciseness of articulation, speed of utterance, 'voice color' (serious, happy, non-committal, etc.). The forms /hæv, həv, əv, v/ occur in environments that are connotatively distinct; and from the difference in their distribution they acquire a certain difference in meaning. When they are then contrasted in an identical environment, they exhibit the connotative differences which are characteristic of the general surroundings where they are usually found. Since, then, their difference of meanings does not violate the negatively defined phonetic-semantic distinctiveness, we may say that this series constitutes a set of allomorphs with submorphemic phonetic-semantic distinctions, symbolized as /hæv ∞ həv ∞ əv ∞ v/.[36]

Principle 7 has not been developed solely for the purpose of helping us to describe some of the awkward things about a language. It is, rather, a common-sense statement of situations with which we deal constantly, and which we must describe in such a way as not to violate the very evident relationships to which we intuitively react. We cannot afford to make fluctuations of form and overlapping of patterns the pivotal points of any system. We must fit them into the system, rather than tailoring the system to fit them.[37]

Principle 7 introduces a feature of analysis into the morphology and syntax which does not apply to phonemics. In phonemics, a difference of form in any part of an utterance, accompanied by a difference of meaning for the whole, justifies the isolation of the formal difference as meaningfully pertinent, and hence as a distinct formal unit. This is not true of the analysis of morphemes, which are themselves meaningful units. In phonemics we admit conditioned and free variation only when the meaning of the whole remains constant. In morphology we must admit submorphemic difference of both form and meaning within the phonetic-semantic distinctiveness of a single morpheme (its non-resemblance to other forms).

3.08 Principle 8. Homophonous forms possessing more than one distinct area of meaning and belonging to correspondingly different distributional classes[38]

34 When two fluctuating forms are shifting their status—when the originally more frequent form is growing less frequent and the other is climbing up—there is undoubtedly a point at which the two are equally in use. At such a point one might insist that free variation actually exists between the forms. But the semantic difference between them derives from the whole pattern of the language; and even though the forms may have the same frequency statistically, they can and will nevertheless be different in meaning, if this is measured by their distribution in terms of the language as a whole.

35 One might also consider that the terms in this series exhibit sufficient semantic-phonetic resemblance to each other to justify a different treatment. One might take /hæv/ as the basic form, and regard the other forms as consisting of the morpheme /hæv/ plus various morphemes of replacement (/ə ← æ/) or loss (/-h/ or /-hæ/). Then the form /əv/ would consist of three morphemes: /hæv/ plus /ə ← æ/ plus /-h/. Such an analysis, however, not only would be awkward but would seem to be untenable in the absence of parallel formations.

36 The occurrence of particular allomorphs here is not determined solely by phonological criteria, nor does the characteristic distribution depend on purely phonetic features. The distribution involves both phonological and morphological criteria, and the alternation between allomorphs is accordinly symbolized by ∞.

37 The fluctuation between It's I and It's me is covered by this principle, but the form me is related to the overt series him, them, whom. See §3.03.

38 The expression 'distributional classes' is used to avoid the possible ambiguity of the term 'form-classes'. The latter is now applied to two fundamentally different sets of forms; not only to distributional classes in the present sense, but also sometimes to classes of forms with a particular phonological structure (e. g. CVCV as against CV, or high-low tone as against low-high). Classes of the second kind may or may not be distributionally pertinent. There is often a correlation between the two kinds of classification, but it is nevertheless important to distinguish them.

consist of as many morphemes as there are parallel semantic and distributional classes.

Principle 8 means that we assign the verb read /riyd/ and the noun reed /riyd/ to two different morphemes. They cover two distinct areas of meaning[39] and belong to two different distribution classes, respectively verbs and nouns. This principle is accepted by practically all linguists and requires no further elaboration.

3.09 Principle 9. Homophonous forms are semantically related when they identify regularly associated aspects of the same object, process, or state.

This principle is intended to define the term 'semantically related'. The morphemic status of semantically related forms will be treated under Principles 10 and 11. Any attempt to classify semantic relationships is extremely difficult; but there are certain relationships between words which we all recognize to a greater or lesser degree. For example, we note a similarity in run as a verb and the homophonous run as a noun. Both forms identify regularly associated aspects of what may be a single process. In the phrases they run and their run, even the relationship of actor to action may be substantially the same, though we normally describe the latter phrase as expressing a possessor-possessed relationship. In the phrase the run in her stocking the word run may be said to identify the result of a process. In the phrases they fish and the fish the homophonous forms are related as process and characteristically associated object. This is not true in the phrase to pare the pear, for a pear is not an object regularly associated with the process of paring; nor is the process of paring regularly associated with pairing. Accordingly, although the homophones run and run, fish and fish, are semantically related, the homophones pair, pare, and pear are not.

The meaningful relationship between homophonous forms may be one of form and function. The term horn which designates a particular shape and/or substance (the horn of an animal or by extension of appearance the horns of the moon) is related to horn meaning a musical instrument, because the two meanings may be combined in a single object, for instance a ram's horn used as an instrument. Some speakers of English may not recognize such a relationship; for them, such homophonous forms would be completely unrelated. In this matter, we are obliged to introduce some subjectivism; for in describing the functioning of a language we must consider the relationships which words possess for the native speakers.

There are many possible semantic relationships between homophonous forms. The following very limited series will give some idea of the range. (Note that the order of meanings in this list has no reference to the order of derivation.)

(1) Form and function: horn of an animal and horn as an instrument
(2) Process and result: to run and a run in her stocking
(3) Process and associated object: to fish and the fish
(4) State and cause: foul (adjective) and to foul the job
(5) Process and agent: to man the ship and a man
(6) Process and instrument: to spear and a spear
(7) Object and associated characteristic: it is a pill and he is a pill
(8) Form and process: a cross and to cross

There are a number of historically related homophones which are probably unrelated in present-day speech. The noun board (A) meaning a particular material and the noun board (B) in the expression room and board are historically related, but today they are probably not associated by any speakers with the same object, process, or state. It is also possible that speakers do not regularly associate either A or B with another historically related form board (C) in the expression board of directors. In the verb expression to board up the house the form board is undoubtedly related to A, and it is even possible that in to board a ship the speaker may make some association with A, but we cannot be certain. It is precisely this dependence upon subjective factors which makes the situation so complex. Nevertheless, despite the inadequate tools which we possess for treating such problems, the answer to the difficulty of identifying morphemes must be found at least partially in the native speaker's response to the meaningful units in his speech.

3.10 Principle 10. Homophonous forms which are semantically related and which occur in correspondingly different distributional environments constitute a single morpheme with multiple distribution-class memberships.

The expressions (1) a run in her stocking, (2) they run away, and (3) they run the office refer to related areas of meaning. The first denotes a condition resulting from a process; the second denotes a process in which the grammatical subject performs the action directly; and the third denotes a process in which the grammatical subject causes an action. The meanings here are not identical, but the subdivision into related areas corresponds to different distributional environments. The first environment is definable as one in which nouns occur; the second, as one in which verbs without objects (i. e. intransitives) occur; and the third, as one in which verbs with objects occur. We may combine the homophonous forms in a single morpheme because the related meanings are paralleled by different distributions.

The distributional differences may be less conspicuous. The sentences it is a pill and he is a pill are structurally identical; but there is a slight difference in the environment: in the first sentence the

39 Criteria for distinguishing 'related' and 'distinct' are given in the next section.

subject is it, in the second it is he. The differences of meaning in the two occurrences of pill parallel this difference of distribution.

Relating transitive and intransitive verbs or combining various meanings of a single noun by complementary distribution does not involve any departure from the usual method; but combining noun and verb forms in the same morpheme seems more radical. Nevertheless, we must either do this or be faced by two clumsy alternatives. First, we might attempt to derive one of the forms from the other. We might set up man, fish, and angle as basic nouns and derive verbs from these by a zero affix. But zero should be a zero alternant of something, and there is no parallel series in English (see §3.04). Similarly, we might set up run, swim, and dive as basic verbs and derive nouns by a zero affix, but the same objection would hold. Furthermore, a division into basic nouns and basic verbs would be questionable on any purely descriptive basis.

The second alternative would be to consider the homophonous forms entirely different morphemes. The difficulty with such a procedure is first that it involves considerable repetition and second that it violates our basic definition of a morpheme as a form which does not exhibit partial phonetic-semantic resemblance to any other form. This definition would obviously not hold if we regarded the homophonous forms of run as more than one morpheme.

Accordingly, we are left with the conclusion that the simplest and best treatment of such homophones occurring in different distributional environments is to consider them single morphemes, but with different class memberships. In dealing with noun and verb classes in English we could list (1) forms which occur both as nouns and as verbs: man, fish, walk, run, swim, jump, mother, father, sun, moon, star, spring; (2) forms which occur as nouns only: boy, girl, grass, elephant, bee; (3) forms which occur as verbs only: be, seem, see, come.

3.11 Principle 11. Homophonous forms which are semantically related but which do not occur in correspondingly different distributional environments constitute as many morphemes as there are meaning-distribution classes.

Suppose that someone insists upon relating board (A) 'construction material' with board (B) 'food' (see §3.09). These two forms occur in the same grammatical environment, e. g. the board we got was terrible; without other information it is impossible to know whether meaning A or meaning B is intended here. The two forms are thus not in complementary distribution, though in other situations they often are. For example, board (A) occurs only with a determiner, e. g. a board, the board, some board, etc.; but board (B) occurs both with and without a determiner, e. g. board is five dollars a week, this board is expensive at any price. Whether, then, we consider these homophonous forms related in meaning or not, they constitute different morphemes because they are not in complementary distribution.

In treating the semantically related forms horn (A) 'animal's horn' and horn (B) 'instrument', we may again treat them as two morphemes. Form A occurs both as a noun and as a verb, e. g. the animal's horn and he horned in; form B occurs only as a noun, e. g. the man's horn. The sub-morphemic semantic difference under A is paralleled by a distributional difference and hence the noun and the verb are subsumed under the same morpheme. Horn (A), however, contrasts in its distribution with horn (B); and so we describe these forms with related meaning as constituting two different morphemes.[40]

3.12 Principle 12. The allomorphs of two (or possibly more) morphemes may be partially or completely coexistent, provided one of the allomorphs is not zero.

Instances in which juxtaposed morphemes fuse and thus reduce their phonemic size are not uncommon. Thus, the expression /did yə/ may reduce to /diJə/, the sequence /dy/ being replaced by /J/. This /J/ is part of both morphemes. For convenience we usually speak of the alternants /did ∞ diJ-/ and the alternants /yə ∞ ə/, but this does not describe the situation accurately. We might use an underline to mark phonemes which are 'possessed' simultaneously by two allomorphs: /did ∞ diJ-/ and /yə ∞ -J̲ə/.

Allomorphs that coincide either partially or wholly are numerous in some languages. In Yipounou (a Bantu language of the Gabon), the first-person object is denoted by a nasal, /n-/ or /m-/ according to the first phoneme of the stem; but before a stem with initial /r/, the nasal is 'lost' and the /r/ is replaced by /t/. A form consisting of the morphemes /ama/ + /n/ + /ronda/ appears as /amatonda/ 'she loved me'. In describing this form we usually say that the prefix /n- ∿ m-/ occurs in a zero alternant and the stem /-ronda/ occurs in the alternant /-tonda/; but this does not do justice to the facts. Rather, the stem morpheme /-ronda/ occurs in the alternant /-tonda/, and at the same time the prefix morpheme /n- ∿ m-/ occurs in the alternant /t-/: the underlinings indicate that the two morphemes coincide.

In French, the morphemes à 'to' and le 'the (m.)' occur in the fused form au /o/.[41] This is simply an extension of the process already noted: here the two allomorphs coincide completely. We say that the allomorph of à is /o-/ and the allomorph of le is /-o/, and further that the distribution of these allomorphs is so restricted that each occurs only with the other, the tactical order of the morphemes being à + le.

40 I admit that my classification is to some extent open to criticism, since I here classify in two or more morphemes certain forms which are identical in shape and related in meaning, but which do not occur in complementary distribution. It is true that such forms exhibit partial phonetic-semantic resemblances to each other. If I adopt the formal criterion of distribution as basic, it is because we have no technique as yet which will enable us to define the various degrees of semantic difference.

41 Cf. Hockett's discussion, Lang. 23.333 ((236)).

The final stipulation of Principle 12, that one of the allomorphs is not zero, is added to obviate the contradictory situation which arises in Hockett's treatment of she and her as allomorphs of the single morpheme ⟨ she ⟩ (Lang. 23.343 ((241))). Hockett's solution is simply another device to render a covert zero element meaningful, while setting aside the overt difference between related forms. By insisting in Principle 12 that allomorphs be overt, I only confirm what has already been discussed under Principle 4.

3.13 Principle 13. A single morpheme may be tactically equivalent to two or more morphemic categories.

In the Greek verb /lu·ɔ·/ 'I loose' the form /-ɔ·/ covers a wide area of meaning: it identifies the tense, mode, number, person, and voice of the combined form. In the form /lu·sontai/ 'they will loose for themselves' these categories are all overtly indicated: /s/ future tense, /o/ indicative mode, /n/ plural number, /t/ third person, and /ai/ middle or passive voice. Such forms, in which all five categories are separately indicated, are rare, and the combination of all these categories into a single form such as /-ɔ·/ is not uncommon. This situation in Greek is different from the one in French au /o/. In the Greek /-ɔ·/ we are not able to identify specific morphemes which would have coexisting allomorphs of the shape /-ɔ·/, for in too many instances such allomorphs would be relatable to zero. It is the particular characteristic of 'inflectional' languages that single morphemes are tactically equivalent to a number of categories, which may be fully and overtly expressed in some forms but only partially in others.[42]

4. Having discussed the principles which govern the identification of morphemes, we turn to an analysis of the types of morphemes. A classification of such types may be based upon (1) the types of phonemes which comprise the morphemes, (2) the positional relationship of the parts of the morphemes, and (3) the positional relationship of the morphemes to other morphemes.[43]

4.1 Morphemes may be classified as consisting of (1) segmental phonemes, (2) suprasegmental phonemes, and (3) both kinds together.[44]

4.11 Morphemes consisting of segmental phonemes are very common. The /-θ/ of growth, the /-t/ of lost, the replacive /æ ← i/ of sang, the /riy-/ of receive are a few examples.

4.12 Morphemes consisting entirely of suprasegmental phonemes are less common. One relatively common variety, however, is intonational pattern. The meaningful units of such patterns constitute morphemes. Thus, in English, the different glides following the last primary stress in a phrase are in meaningful contrast,[45] and are therefore morphemes.

We sometimes find that the so-called morphological tones accompanying particular segmental sequences constitute morphemes in their own right. In Ngbaka (a 'Sudanic' language of the northwest Congo) simple verbs have no inherent tone; rather, they occur with four different accompanying tone patterns: (1) low, denoting continuous action, (2) mid, denoting completive action, (3) high, denoting imperative, and (4) low-to-high, denoting future. Regardless of the length of the stem, the tone patterns are constant; cf. the two verbs sa 'to call' and yolo 'to stand': (1) sà, yòlò; (2) sā, yōlō; (3) sá, yóló; (4) sǎ, yòló. In this three-register tone language, the tones of verbs have independent status as morphemes, since there is no basic tonal form of any verb to which the four modifications can be related.

4.13 Finally, there are morphemes consisting of both segmental and suprasegmental phonemes. The English morpheme /bóy/ consists of three segmental phonemes plus the suprasegmental phoneme of stress.[46] In Ngbaka, nouns have inherent tone: lí 'face', lī 'name', lì 'water'.

Note that Ngbaka morphemes have two 'canonical' forms:[47] some include an inherent tone, like the nouns just cited; others are toneless, like the verbs cited above. There are other examples of such breaks within the basic structure of a language. In Semitic languages there are three types of morphemes:[48] (1) radicals of the type CCC, e.g. Hebrew /šbr/ 'break'; (2) patterns of vowels, or of vowels plus length, or of vowels plus a consonantal affix, e. g. --u- 'active', -a-i·- 'object with a particular quality', -a-·a- 'transitive intensive', n-a-i- 'middle'; (3) sequences of vowels and consonants, e. g. mí· 'who?', -tím 'you (masc. pl.)', baltí· 'without'. Again, the accentual system of classical Greek shows a structural break: in verbs the position of the accent is determined by the phonological structure of the form; but in other parts of speech it is not thus conditioned.

42 Hockett treats a similar problem in his discussion of tense-mode and person-number elements in Spanish verbs, Lang. 23.338 ((239)).

43 This paper makes no attempt to treat the significance of various classes of structure, such as derivational vs. inflectional formations, criteria for class determination, the relative order in descriptive statements, and the relation of constituent elements to the sets of immediate constituents. For these subjects see my book Morphology (Ann Arbor, 1946).

44 Juncture is intentionally excluded from this classification. Junctures are to be treated on the same level as order, which is the other formal feature of arrangement.

45 See Kenneth L. Pike, The intonation of American English (Ann Arbor, 1945).

46 In some positions within the intonational pattern the phoneme of full stress may be reduced.

47 The term 'canonical form' is Hockett's; see Lang. 23.333–4 ((236)).

48 Zellig S. Harris, Linguistic Structure of Hebrew, JAOS 61.143–67 (1941).

The segmental and the suprasegmental phonemes which compose a morpheme need not occur simultaneously (see §4.2). In Turu (a Bantu language of Tanganyika), the subjunctive ends in -è, but in addition the verb as a whole has a characteristic subjunctive tone pattern: the stem syllable is high, all succeeding syllables are low. Here a tonal pattern and a particular suffix combine to make a single functional unit.

Non-simultaneous components of morphemes also exist in English. The suffix /-itiy/ in such words as ability, facility, banality, regularity, similarity, and generosity always occurs with presuffixal stress. This position of the stress may be regarded as a part of the nominalizing morpheme. It is thus no longer a 'property' of the underlying form, but rather of the suffix.

4.2 Morphemes may be classified by the positional relationship of their constituent parts into continuous and discontinuous. Morphemes of which the constituent phonemes (segmental and suprasegmental) are adjacent or simultaneous constitute the usual type. Indeed, the very fact of separation generally leads one to believe that the parts constitute separate morphemes unless (1) the units can be demonstrated to be separated by replacive or infixal forms, or (2) the units never occur without each other. In a form such as /sæŋ/ the phonemes /s...ŋ/ constitute a discontinuous morpheme. In the Hebrew examples cited in §4.13 the phonemes of the radical CCC are usually discontinuous. In all true instances of infixation the two parts of the morpheme in which the infix appears together make up a discontinuous morpheme, e.g. the sequence vi...c in Latin vincō (with infix -n-).

Sometimes we find separable units which occur only with one another. In Kissi (a 'Sudanic' language of French Guinea) there are two interrogative particles yɛ̀ and nɛ̄. In any interrogative expression the first of these occurs immediately after the verb and the second occurs in sentence-final position, e.g. vɛ̀ yá gbɛ́ngbɛ́lá yɛ̀ ndū sísà nɛ̄, literally 'how I see interrogative-particle him now interrogative-particle'. The forms yɛ̀ and nɛ̄ constitute a discontinuous morpheme yɛ̀ ... nɛ̄.[49]

4.3 Morphemes may be classified, finally, by their position with respect to other morphemes as (1) additive, (2) replacive, (3) additive and replacive, and (4) subtractive.

4.31 Additive linear morphemes are stems and affixes. The affixes may be prefixes, infixes, or suffixes. Such morphemes are so common in so many languages that no illustrations are needed. Additive supralinear morphemes (consisting solely of suprasegmental phonemes) are intonational units and tonal morphemes.

Additive morphemes of this last variety are not common, but the Ngbaka data (§4.12) are not unique. In Kissi, polysyllabic verb stems have no independent basic tone, but are accompanied by four sets of tone patterns indicating mode-aspect and number-person categories. With plural subject, the typical verb baŋa 'redeem' has the following tones: bàŋá completive, bàŋà incompletive, báŋà negative, and bāŋā interrogative. The tones on the verbs are in no way associated with basic morphological classes of verbs, and hence are genuine additive morphemes.

4.32 Replacive morphemes, like additive morphemes, may consist of segmental or of suprasegmental phonemes; we have already noted a number of examples (§3.04, §4.11). A rather rare type of replacement is represented by the English series bath : bathe, sheath : sheathe, wreath : wreathe, teeth : teethe, safe : save, strife : strive, thief : thieve, grief : grieve, half : halve, shelf : shelve, serf : serve, advice : advise, house /haws/ : house /hawz/, etc. In each pair, the noun has a voiceless continuant, the verb a voiced continuant. If we agree to derive the verbs from the nouns,[50] we set up three specific replacive elements: /ð̌ ← θ/, /v ← f/, and /z ← s/; but since these three elements exhibit a phonetic-semantic resemblance to each other, and since their occurrence is phonologically conditioned, we combine them into a single replacive morpheme.

In Shilluk (a Nilotic language of the Anglo-Egyptian Sudan), there are four principal ways of forming plural nouns from singular bases in phrase-final position: by syllabic replacives, by lengthening of the stem vowel (a type of additive morpheme), by the suffixation of -i, and by tonal replacives. No one of these types is sufficiently predominant to be called the basic alternant; and there are even certain combinations of two of the types in one formation. The tonal replacives occur in a definite pattern; compare the following singulars and plurals: wāt 'house' : pl. wàt, ógāṭ (where o- is a prefix) 'cloth' : pl. ógàṭ, yíṭ 'ear' : pl. yīṭ, tík 'chin' : pl. tīk, tóŋ 'spear' : pl. tōŋ. The plural stems differ from the singular through the replacement of a mid tone by a low tone, or of a high tone by a mid tone: in either case, the replacement of a higher tone by a lower one. Both types of replacement may be combined in a single morpheme.

In Mongbandi (another 'Sudanic' language of the northwest Congo), verbs have basic tonal patterns. Verb stems, both monosyllabic and dissyllabic, belong to various form-classes according to the fundamental tone that accompanies them with a singular subject and in the completive aspect. With a plural subject these form-classes are not distinguished: all monosyllabic verbs in the completive aspect have a high tone, all dissyllabic verbs have mid-high tone.

[49] The situation with respect to French ne ... pas is not analogous. The element ne occurs without pas (in ne ... rien, ne ... point, ne ... que, etc.), and pas occurs without ne (in pas du tout, pourquoi pas, etc.).

[50] The direction of derivation depends primarily on semantic criteria; but there are parallel formal patterns to guide the choice.

Accordingly, the morpheme that means plural number and completive aspect consists of the replacement of the basic tone (whatever it may be) by a high or a mid-high tone.

The shift of stress in related nouns and verbs in English (impact, import, insult, insert, discourse, rebel, protest, etc.) is also a type of replacive. The morpheme in this instance is /V̌...V ← V...V̌/, where V stands for any syllabic. The stress of the underlying verb is not here regarded as a morpheme by itself; what is morphemic is rather the replacement of a stress on the second syllable by a stress on the first.

4.33 Morphemes may also consist of a combination of additive and replacive elements. As examples I may refer to the morphemes cited in §4.13, which are composed of segmental and suprasegmental phonemes not simultaneous in their occurrence. Thus, the Turu final -è (an additive element) occurs in combination with a particular tonal pattern on the preceding syllables (a replacive element).

4.34 Subtractive morphemes are illustrated by the masculine forms of certain French adjectives.[51] In the series plat /pla/ : platte /plat/ 'flat', laid /lɛ/ : laide /lɛd/ 'ugly', long /lɔ̃/ : longue /lɔ̃g/ 'long', soûl /su/ : soûle /sul/ 'drunk', the subtractive element (the final consonant of the feminine forms) has different phonological shapes—/t, d, g, l/; but at least it is always final. Accordingly, in spite of the diverse phonemic make-up of the various allomorphs, we may still combine them all into one morpheme on the basis of their semantic distinctiveness and the phonological determination of their distribution.

[51] Bloomfield, op. cit. 217. William L. Wonderly has proposed in discussion that these French forms might be most economically handled as morphophonemic. Cf. George L. Trager, The verb morphology of spoken French, Lang. 20.131–41 (1940).

[[If there can be any such thing as 'common sense' in descriptive linguistics, Nida's work is where we may expect to find it, because he carries a responsibility both to perform and to guide work that is adequate to practical purposes and will also stand up under theoretical criticism. Because I lack experience to match his, I must content myself here with referring to the general comments appended to Bloch's paper, p. 254.]]