[Mechanical Translation and Computational Linguistics, vol.8, nos.3 and 4, June and October 1965]

The Nature of Affixing in Written English *†

by H. L. Resnikoff and J. L. Dolby††, Lockheed Missiles & Space Company, Palo Alto, California

Any algorithmic study of written English must sooner or later face the problem of unscrambling English affixes. The role of affixes is crucial in the study of word-breaking practice. In the automatic determination of the parts of speech (a central feature of automatic syntactic analysis), the suppressing action of affixes must be understood in detail. In the determination of English citation forms, complete lists of affixes are necessary. The inflection of English verbs is tied up with the existence of suffixes.

Existing definitions of affixes suffer because they are neither computable nor in general agreement with one another, and none of them refers directly to written English. Existing lists of affixes vary widely in size and content, implying a lack of agreement as to what constitutes a complete listing of English affixes, or how one is to be obtained.

In this paper we show that there is a natural structural definition of English affixes, and that this definition can be implemented on existing word lists to provide exhaustive affix lists. In particular, the definition is applied to all the two-vowel string words in the Shorter Oxford Dictionary, and a complete list of the resulting affixes is provided. Some applications to problems of stress patterns, doubling rules in verb inflection, and the determination of the number of phonetic syllables corresponding to a written word are described.

Computational linguistics differs in at least three essential respects from traditional linguistics. Foremost among these is that computational linguistics deals almost entirely with written languages. Because of this restriction to strictly reproducible forms and because of its direct connection with computers, it is both possible and necessary to operate primarily with operational definitions that are capable of machine implementation. Finally, the same forces that require strict operational definitions also impose upon us the necessity of establishing procedures of extremely high precision and accuracy. In a word, 80% is not nearly good enough for machine operation, 98% might pass, and it is fairly clear that programs will have to operate at well above the 99% level of accuracy if they are to attain any degree of general use. The attainment of such precision, and the proof that such precision has been obtained in a particular case, may well be considered primary problems in this area.

If such precision is eventually to be obtained in the solution of such sweeping problems as machine translation, abstracting, indexing and the like, it must first be obtained on more mundane levels: at the sentence level and at the word level. Our own efforts have been restricted primarily to the treatment of words: to the determination of highly accurate algorithms for finding properties of words, and to the development of measures that allow us to determine when an algorithm has reached a desired level of accuracy. In so doing we have found it convenient to group the words of written English into a linear ordering according to the number of vowel strings contained in the word. Our study of the one-vowel string or CVC words is reported with some thoroughness in reference 1. There we established the conventions, which will also be adhered to throughout this paper, that the letters A, E, I, O, U, and Y are vowels but that E in final position is a consonant, and that words that begin or end with a vowel are augmented by the addition of a symbol called the *blank consonant,* so that all words can be considered as beginning and ending with a consonant. For example, according to these conventions, the words A, AT, BAT, BATE are all of the form CVC (where, as usual, C denotes a string of consonants, and V denotes a string of vowels). In this article we discuss our study of the two-vowel string, or CVCVC, words. Although much of the essential structure found in the CVC words is carried over, we find (quite naturally) that there is a new feature in the CVCVC words: almost all of them contain either a prefix or a suffix. It is therefore necessary to establish an operational definition of affixes.

It seems appropriate to describe briefly some of the previous work related to affixes. Although this discussion does not pretend to be complete, we do think that

* This paper was presented at the Bloomington meeting of the A.M.T.C.L., July, 1964, in a slightly different form.
† This work supported by the Office of Naval Research and the Independent Research Program, Lockheed Missiles & Space Company, Palo Alto, California.
†† Mr. Resnikoff is presently at the Institute for Advanced Study, Princeton, New Jersey, and Mr. Dolby is with C-E-I-R, Inc., Los Altos, California.

the major lines of development are covered. In Perry's extraction[2] from Johnson's dictionary, published in 1805, the word 'affix' is defined as follows: "some letter, syllable, or particle joined to the end of a word." 'Prefix' is defined as "some particle put before a word to vary its signification." The word 'suffix' is not given. The 1836 edition of *Walker Remodelled,*[3] edited by Smart, defines 'suffix' as a "letter or syllable added to a word," while the definitions of 'affix' and 'prefix' agree substantially with Johnson. The Oxford English Dictionary[4] draws its definition from Haldeman's *Affixes to English Words,*[5] published in 1865. He states: "*Affixes* are additions to roots, stems, and words, serving to modify their meaning and use. They are of two kinds, *prefixes,* those at the beginning, and *suffixes,* those at the end of the word-bases to which they are affixed." The terms have been fixed with essentially the same signification since Haldeman's time.

This last definition is sufficiently general to account for the facts, but it is open to question just because of its generality, in that it permits too great a variation in the interpretation of the terms 'roots' and 'stems', and also because it is noneffective, in that it does not attempt to indicate how "modified meaning" and "use" are to be determined. The essence of the problem of the definition of 'affix' lies here. It is not too hard to construct a sufficiently broad and inclusive definition; the construction of an effective definition is another matter.

In his monumental grammar of the English language, Jespersen[8] devoted 44 pages of Volume VI to affixes, but never defined the basic terms. Contemporary linguists seem to be more aware of the need for and usefulness of accurate and adequate definitions, but affixes do not seem to be the center of interest. For example, Gleason[7] states that a definition of 'affix' would be immensely complex in general, but that it is feasible for one specific language. He proceeds to give some examples of English affixes, but makes no attempt explicitly to define the class. Bloomfield[8] recognizes the importance of the affixing and compounding processes, and gives a clear but noneffective definition. He states that "the bound forms which in secondary derivation are added to the underlying forms are called 'affixes'."

Part of the difficulty that these attempts at definition encounter is that there are really two problems to be faced. Although this is rather evident, no one seems to have taken the trouble explicitly to differentiate them, and this has resulted in a certain confusion. It is one question to ask whether a particular letter sequence is an affixing sequence, and quite another to ask whether it is an affix in a particular word. Bloomfield's definition, for example, does not logically permit one to consider affixes independent of the words in which they are bound; one cannot say that 're-' is a prefix, for in 'return' it is, while in 'receive' (at least by Bloomfield's illustration), it is not. Therefore, strict observance of Bloomfield's definition denies the possibility of even listing the affixes; the best that can be done is to list all words that contain affixes, and to indicate in each word which letter sequence is the bound form in secondary derivation.

Once the two questions are distinguished, it is possible to ask for the sequences that *can* occur as affixes, and to list these. We will distinguish the two questions by searching for those sequences that are affixes in *some* contexts (i.e., words), and we will call these sequences 'affixes'; the second question is then that of determining when an affix is an affix in a *particular* context (i.e., word).

Before proceeding further, we recall a definition from section 2 of reference 1. There a threshhold was established to eliminate words and other strings of letters with rare structural properties from the corpus of forms under consideration. The same criterion will be invoked in this paper: if a class of words or letter strings with a given property contains more than three (3) members, then the class will be called "admissible" with respect to the given property and the corpus. Thus, the set of CVC words that begin with the consonant string FN is not admissible, because there is only one word with this property (in the Shorter Oxford Dictionary): FNESE. The threshold level "three" appears to be the least number that leads to interesting results.

In order to obtain a procedure for finding affixes, we will make use of one of the main results of reference 1. There we found that certain consonant strings such as PL occur only in initial position in CVC words, certain strings such as NT occur only in final position, while some, such as T, occur in both positions. The initial and final consonant strings of the CVCVC forms turn out to be similar to sets found for the CVC forms. However, the internal consonant strings of the cvcvc forms include all possible admissible initial and admissible final C strings in CVC words (these are listed for reference in Table I), as well as some admissible strings not found in CVC words, such as NF (as in CONFINE), and this suggests a means for classifying the set of CVCVC words according to the behavior of the internal consonant string. We therefore consider four classes typified by the words:

I. DETER
II. REPLACE
III. RENTER
IV. CONFINE

These classes can be precisely defined as follows. Let 'B' denote the set of admissible initial consonant strings of cvc words, and 'E' denote the set of admissible final consonant strings of CVC words. Then a CVCVC word belongs to Class I if its internal consonant string belongs to both of the sets B and E, to Class II if its internal consonant string belongs to B but not E, to Class III if its internal consonant string belongs to E but not B, or the Class IV if its internal consonant string belongs to neither B nor E.

TABLE I.
ADMISSIBLE INITIAL CONSONANT STRINGS OF CVC WORDS

B	N	BL	GL	SH	TR	SCH
C	P	BR	GN	SK	TW	SCR
D	Q	CH	GR	SL	WH	SHU
F	R	CL	KN	SM	WR	SPH
G	S	CR	KR	SN		SPL
H	T	DR	PH	SP		SPR
J	V	DW	PI	SQ		STR
K	W	FL	PR	ST		THR
L	Z	FR	RH	SW		THW
M		GH	SC	TH		

ADMISSIBLE FINALCONSONANT STRINGS OF CVC WORDS
NOT ENDING WITH E

B	BB	MP	SH	GHT
C	CH	ND	SK	LCH
D	CK	NG	SM	LPH
F	CT	NK	SP	LTH
G	DD	NN	SS	MPH
H	FF	NT	ST	MPT
K	FT	NX	TH	NCH
L	GG	PH	TT	NTH
M	GH	PT	WD	NTZ
N	GN	RB	WK	RCH
P	LD	RC	WL	RSH
R	LF	RD	WN	RST
T	LK	RF	XT	RTH
W	LL	RK	ZZ	SCH
X	LM	RL		TCH
Z	LP	RM		
	LT	RN		
	MB	RP		
	MM	RR		
	MN	RT		

Note that S does not appear in this list because of the conventions used in reference 1.

From the affix point of view the problem is at its worst in the first case. Since any reasonable definition of 'affix' will recognize DE as a potential prefix and ER as a potential suffix we can decompose the word DETER in three possible ways:

1. as a prefixed form DE/TER
2. as a suffixed form DET/ER
3. as a 2-syllable kernel word DETER with no affixes at all.

This problem can only be resolved at the "affix in context" level. The collection of words belonging to Class I does not help us to formulate an operational definition of 'affix'.

The words in Class II, typified by REPLACE, have the property that the internal-consonant string is an admissible initial-consonant string. The words in Class III have the mirror image property that the internal-consonant string is an admissible final string, such as NT in RENTER.

There are two potential decompositions for words belonging to Class II and Class III, which are typified by the decompositions given below:

RE-PLACE
REP-LACE
and
RENT-ER
REN-TER.

From an operational point of view, PL is an admissible initial consonant string, so the first decomposition of REPLACE is reasonable. But, equally, the letter P is an admissible final consonant string, and L is an admissible initial consonant string, so the decomposition REP-LACE is equally conceivable. A similar argument applies to the Class III words. Note that we might choose to define the prefixing strings by requiring that the longest admissible initial consonant string be used to decompose words of Class II, but there is no evident reason to do so. Nonetheless, this idea is essentially correct, as we will see when we examine the Class IV words.

The Class IV words are distinguished by the property that the internal consonant string is neither an admissible initial- nor an admissible final-consonant string; for example, the string NF in CONFINE. Cursory observation appears to indicate that the internal consonant string C can always be written as a sequence C'C" of consonant strings such that C' is an admissible final consonant string of CVC words, and C" is an admissible initial consonant string of CVC words (and neither C' nor C" is blank). Thus NF can be written as N-F. It can of course happen that such a decomposition is possible in more than one way, but we are now concerned only with discovering whether there is always at least one such decomposition. If we examine the 22,568 cvcvc words in the Shorter Oxford Dictionary, we find that the internal consonant strings NCT, VR, and VV are the only ones that do not have a decomposition of the form C'C" as described above. These internal consonant strings occur in 21, 7, and 6 words respectively. Using the threshold criterion, since there are only three internal consonant strings that do not have decompositions of the form C'C", we delete the 34 words containing these strings from the corpus. Hence, every Class IV word in the (reduced) corpus has at least one decomposition of the required form.

It may be worth remarking that there are 180 two-letter, 180 three-letter, and 29 four-letter admissible internal consonant strings that *do* have at least one decomposition of the form C'C". Here, of course, an internal consonant string is admissible if there are more than three cvcvc words with this internal consonant string.

If a word CVC'C"VC has a unique decomposition point between C' and C", we will say that C'C" is a "mandatory decomposition point." For example, CONFINE has the mandatory decomposition CON-FINE. The CVCVC words with mandatory decomposition points can be used to generate a first list of affixes.

Let a two-vowel string word be given in the form CVC'C"VC, where the consonant string C'C" denotes the internal-consonant string of the word. Suppose a corpus K of CVCVC words is fixed. Then we define the class Cls(CVC'/C") to be a collection of all words in the fixed corpus of the form CVC'C"X, where X denotes an arbitrary string. Similarly, we define Cls (C'/C"VC) to be the collection of all words in the fixed corpus of the form YC'C"VC, where Y denotes an arbitrary string. With the aid of these sets, we make the following definitions:

Definition P1: Let P = CVC' be a fixed letter string, P is called a “strong prefix” if there exist two distinct classes, Cls(P/C_1") and Cls (P/C_2"), each of which contains more than three words, such that C'C_1" and C'C_2" are mandatory decomposition points.

Definition S1: Let S = C"VC be a fixed letter string, S is called a “strong suffix” if there exist two distinct classes, Cls(C_1'/S) and Cls(C_2'/S), each of which contains more than three words, such that C'$_1$C″ and C_2'c" are mandatory decomposition points.

Definition A1: A letter string is called a “strong affix” if it is either a strong prefix or a strong suffix.

In the above definitions, all words are taken from the fixed corpus K of CVCVC words.

It is clear from the definitions that a two-vowel string affix, such as INTER, will not be found, for the corpus has been limited to CVCVC words, and the definition is phrased in terms of this corpus. However, the alterations in the definitions that will make them applicable to affixes containing an arbitrary number of vowel strings are quite straightforward, and will not be given here.

Definitions differing from the above only in that they require a different number of classes, containing a different number of words, to satisfy the given conditions, are reasonable on the surface, and so it is necessary to discuss the reason for requiring *two* classes, each containing more than *three* words. Application of the definition with these numeric requirements relaxed so that a class need contain only one word shows that minor structural irregularities of English lead to “affixes” that are unsatisfactory from an intuitive point of view, and are not found even in the most exhaustive affix lists. The "more than three" criterion is based on the identical procedure followed in reference 1. The requirement that at least two classes fulfill the defining conditions is more interesting. When this is relaxed, certain new letter strings satisfy the relaxed conditions. An example is FOR-; this string is usually considered to be a compounding unit. The example is typical of the new “affixes” produced by the relaxed definition. We take the view that the difference between affixes and compounding units is not one of kind, but one of degree: affixes are attached to more classes of words. One problem of 'affix' definition is to select the proper threshold for discriminating between affixes and compounding units. The requirement that there be at least two classes, as stated in the definitions above, leads to intuitively satisfactory affix lists, whereas requiring any larger number of classes would suppress certain well-known affixes.

Application of the definitions to the corpus K consisting of all of the cvcvc words listed in the Shorter Oxford Dictionary leads to the strong affixes given in Table II.

We give some of the details illustrating the application of the definitions to obtain the affixes listed in Table II. The strong suffix WARD occurs in the two admissible classes Cls(N/WARD) and Cls(R/WARD), each containing five words. The strong suffix -FUL appears in ten distinct admissible classes: Cls(D/FUL), Cls(SH/FUL), Cls(TH/FUL), Cls(RM/FUL), Cls(N/FUL), Cls(P/FUL), Cls(GHT/FUL), Cls(T/FUL), Cls(RT/FUL), and Cls(ST/FUL), containing 8, 6, 11, 4, 10, 5, 7, 5, 4, and 13, words respectively. The other strong affixes are found from similar determinations of their classes. See Table IV for the complete list of admissible classes for the determination of the strong suffixes.

From the definitions, it is clear that a strong prefix must end with a consonant, and a strong suffix must begin with a consonant. Hence, although the strong affixes given in Table II all seem to be reasonable intuitive affix candidates, the familiar vowel-ending prefixes and vowel-beginning suffixes are not accounted for.

TABLE II. STRONG AFFIXES

Strong Prefixes		*Strong Suffixes*	
AC-	IN-	-FUL	-LY
AD-	MIS-	-LAND	-LOCK
AL-	OUT-	-LER	-MAN
CON-	SUB-	-LESS	-MENT
DIS-	SUN-	-LET	-NESS
EN-	TRANS-	-LING	-WARD
EX-	UN-		

The definitions P1 and S1 can be extended to include the words belonging to Class II and Class III, and these will give the vowel-ending prefixes and the vowel-beginning suffixes. Because there is no mandatory decomposition for words belonging to these two classes, we cannot assert that the decompositions are invariably correct. For this reason, we refer to the affixes found from words belonging to Class II or Class III as “weak affixes.” The definition corresponding to Definition P1, for instance, is:

Definition P2: Let P = CV be a fixed-letter string, p is called a “weak prefix” if there exist two distinct classes Cls(P/C_1) and Cls(P/C_2), each of which contains more than three words, such that C_1 and C_2 are admissible initial strings. Here, C_1 and C_2 are the internal-consonant strings of the two-vowel string words comprised by the corpus K.

The definition of 'weak suffixes' involves a similar transcription of Definition S1, and we will therefore not give it here.

Application of these two definitions to the corpus K defined above leads to the weak affix lists given in Table III.

TABLE III. WEAK AFFIXES

Weak Prefixes	*Weak Suffixes*			
A-	-A	-ENT	-IS	
BE-	-AGE	-EON	-ISH	
CY-	-AH	-ER	-ITE	
DE-	-AL	-ET	-IVE	
E-	-AN	-EY	-O	
I-	-ANT	-IC	-OCK	
RE-	-AR	-IE	-ON	
	-ARD	-IER	-OR	
	-AT	-ILE	-OT	
	-ED	-IN	-OW	
	-EE	-INE	-UE	
	-EL	-ING	-UM	
	-EN	-ION	-URE	
			-US	

Although these affix lists appear quite reasonable, a more objective operational method is necessary if any degree of "proof" is to be claimed. This can be provided by examining various applications where it is known or suspected that affixation plays a dominant role, such as:

A. The determination of stress patterns

B. The determination of consonantal doubling rules in the inflection of English verbs

C. The determination of word-breaking rules as used in end-of-the-line practices in type composition

D. The determination of parts-of-speech assignments

E. The determination of the number of phonetic syllables corresponding to a written English word

In the first case, we have taken a random sample of 100 cvcvc words, each containing one affix from our lists, and found that in 95 of the words the syllable containing the affix was unstressed, thus providing some assurance that the affixes we have so identified are in fact affixes. A more complete sample is obviously needed for a precise estimate of the error rate of our procedures.

A more interesting check is provided by the verb-inflection problem. Here we can immediately determine the rather obvious algorithms needed for most of the words and put this together with a list of irregular forms for a working procedure, except for the presence of a number of verbs where it is necessary to double the final consonant in the preterite and participial forms. Without dwelling on the problem at length, we find that consonantal doubling never occurs when a suffix in context is present. Use of the present affix list enables us to reach an accuracy rate of 98.9% for our verb inflection algorithm, thus providing further evidence that we are not far off. Comparable figures are found in the word-breaking and part-of-speech problems.

TABLE IV.

ADMISSIBLE CLASSES OF THE FORM Cls (C'/C"VC) FOR THE DETERMINATION OF STRONG SUFFIXES. THE NUMBER OF WORDS IN EACH CLASS IS SHOWN. SUFFIXES ARE UNDERLINED.

-CA	Cls(C/CA) 6
-MA	Cls (G/MA) 10
-FOLD	Cls(N/FOLD) 6
-LAND	Cls(D/LAND) 4
	Cls(T/LAND) 4
-WARD	Cls(N/WARD) 5
	Cls(R/WARD) 5
-STONE	Cls(D/STONE) 4
-CATE	Cls(C/CATE) 4
-STATE	Cls(N/STATE) 4
-LING	Cls(D/LING) 10
	Cls(DD/LING) 4
	Cls(ND/LING) 8
	Cls(CK/LING) 9
	Cls(NK/LING) 4
	Cls(N/LING) 5
	Cls(T/LING) 15
	Cls(NT/LING) 6
	Cls(ST/LING) 4
-LOCK	Cls(D/LOCK) 4
	Cls(N/LOCK) 4
-FUL	Cls(D/FUL) 8
	Cls(SH/FUL) 6
	Cls(TH/FUL) 11
	Cls(RM/FUL) 4
	Cls(N/FUL) 10
	Cls(P/FULJ 5
	Cls(GHT/FUL) 7
	Cls(T/FUL) 5
	Cls(RT/FUL) 4
	Cls(ST/FUL) 13
-LY	Cls(D/LY) 12
	Cls(ND/LY) 8
	Cls(TH/LY) 6
	Cls (CK/LY) 7
	Cls (M/LY) 6
	Cls(N/LY) 9
	Cls(T/LY) 11
	Cls(GHT/LY) 10
	Cls(RT/LY) 5
	Cls(ST/LY) 15
-MAN	Cls (D/MAN) 10
	Cls (RD/MAN) 4
	Cls(G/MAN) 4
	Cls(CK/MAN) 5
	Cls(LL/MAN) 4
	Cls(P/MAN) 5
	Cls(T/MAN) 9
-LESS	Cls(D/LESS) 14
	Cls(ND/LESS) 10
	Cls(RD/LESS) 4
	Cls (TCH/LESS) 4
	Cls(TH/LESS) 6
	Cls(CK/LESS) 7
	Cls(M/LESS) 5
	Cls(RM/LESS) 6
	Cls(N/LESS) 17
	Cls(T/LESS) 14
	Cls(GTH/LESS) 7
	Cls(NT/LESS) 8
	Cls(RT/LESS) 4
	Cls(ST/LESS) 14
-NESS	Cls(D/NESS) 7
	Cls(LL/NESS) 7
	Cls(L/NESS) 4
	Cls(T/NESS) 11
	Cls(GHT/NESS) 4
-LET	Cls(M/LET) 7
	Cls(N/LET) 5
	Cls(NT/LET) 6
	Cls(RT/LET) 5
	Cls(T/LET) 4
-MENT	Cls(C/MENT)^
	Cls(SH/MENT) 4
	Cls(T/MENT) 4
-WAY	Cls(R/WAY) 5
-QUET	Cls(C/QUET) 5
-LER	Cls(CK/LER) 6
	Cls(ST/LER) 4
	Cls(TT/LER) 6

The last problem has a double interest because it not only illustrates the role of affixation in written English, but also indicates that a remarkably close connection exists between written English and its spoken forms (In this respect, note also reference 10). It turns out that the trivial rule:

number of vowel strings equals number of phonetic syllables

is about 80% accurate. By introducing the affixes found in this paper it is possible to construct an elementary algorithm that has an accuracy of better than 94%. The problems that remain have to do primarily with internal "consonantal" ES, i.e., "silent" ES, and with compounding units that are not affixes. Problem E is discussed in reference 9.

In this paper we have been primarily concerned with offering an operational definition of 'affix of English', rather than with the detailed problems that arise in the application of the definition. However, we must add a word about some of these problems in order to place them in the proper perspective. First, because of the final E convention used in reference 1, the final letter string -LE is a consonant string, and is not obtainable as a strong suffix from the corpus of cvcvc words. But methods completely analogous to those used here will show that -LE is a strong suffix obtainable from the corpus of CVC words. Most of the details are contained in reference 1, where a complete list of cvc words ending with -LE is given. Although the final string -RE behaves like -LE in many ways, it turns out that -RE is not a strong suffix in the sense of that term as defined here.

Second, at least two important classes of affixes do not show up in the CVCVC words: the multivowel-string affixes such as INTER-, and the affixes that are appended only to other affixes, such as -OUS. The investigation of these affixes requires examination of the three-, four-, etc. vowel-string words. As an indication of the complexity of this problem, we recall that there are 20,762 three-vowel-string words, 10,293 four-vowel-string words, 2,770 five-vowel-string words, 393 six-, 30 seven-, and 4 eight-vowel-string words in the Shorter Oxford Dictionary. This gives a total of 89,656 internal consonant strings that must be examined and classified, compared with the 22,568 internal consonant strings examined for the present study of the two vowel string words.

Finally, we have discussed only the question of determining the affixing strings. The more delicate problem of deciding when an affix is *acting* as an affix in a particular word remains. For example, the weak prefix RE- acts as an affix in READJUST, but not in READING. We hope to report on these problems directly.

Received September 25, 1964

References

1. J. L. Dolby and H. L. Resnikoff, "On the structure of written English words," *Language 40* (1964) pp. 167-196.
2. William Perry, *The Synonymous, Etymological, and Pronouncing English Dictionary,* London, 1805.
3. Benjamin Humphrey Smart, *Walker Remodelled: a new Critical Pronouncing Dictionary,* London, 1836.
4. James A. H. Murray, et. al. (editors), *The Oxford English Dictionary,* Oxford, 1933.
5. Samuel Steman Haideman, *Affixes in their Origin and Application, Exhibiting the Etymological Structure of English Words,* Philadelphia, 1865.
6. Otto Jespersen, A *Modern English Grammar on Historical Principles,* Copenhagen, 1909, 1949.
7. H. A. Gleason, Jr., *An Introduction to Descriptive Linguistics,* revised edition, New York, 1961.
8. Leonard Bloomfield, *Language,* New York, 1933.
9. J. L. Dolby and H. L. Resnikoff, "Counting phonetic syllables—an exercise in written English," (to appear).
10. B. V. Bhimani, J. H. Dolby, and H. L. Resnikoff, "Acoustic phonetic transcription of written English," presented to the 68th meeting of the Acoustical Society of America, Austin, Texas, 1964.