[*Mechanical Translation and Computational Linguistics*, vol.10, nos.3/4, September and December 1967]

Automatic Determination of Parts of Speech of English Words

by Lois L. Earl,* Lockheed Palo Alto Research Laboratory, Palo Alto, California

The classifying of words according to syntactic usage is basic to language handling; this paper describes an algorithm for automatically classifying words according to thirteen commonly used parts of speech: noun, adjective, verb, past verb, adverb, preposition, conjunction, pronoun, interjection, present participle, past participle, auxiliary verb, and plural or collective noun. The algorithm was derived by a computerized study of the words in The Shorter Oxford English Dictionary. *In its operation it utilizes a prepared dictionary of around nine hundred words to assign parts of speech to special or exceptional words. Other words are split into affix and kernel parts and assigned a part of speech on the basis of the part-of-speech implications of the affixes and the length of the remaining kernel. An accuracy of 95 per cent is achieved from the point of view of inclusive part of speech, where inclusive part of speech is defined as that string which contains all the parts of speech attributed to the word by the dictionary but which may also contain one or two more parts of speech.*

Introduction

This paper describes the development and details of a procedure for automatically assigning part-of-speech characteristics to English words, largely from graphemic considerations. The development of the algorithm began with the observation of Dolby and Resnikoff[1] that the parts of speech associated with one-syllable words are frequently noun (or noun and adjective) and verb, while the parts of speech associated with multisyllable words are usually noun and adjective only. Development of a working part-of-speech algorithm required the study of exceptions to this general rule so that analytical subrules and exception lists sufficient to identify automatically all such exceptions could be derived. Two analyses were utilized for the isolation and study of exceptions: (1) Exhaustive sorts of a 73,582-word dictionary on magnetic tape were used to separate words consistent with the general rule from those words that were not and to classify them. (2) Computer analysis of possible part-of-speech implications of affixes was carried out on the same dictionary. The algorithm developed utilizes a prepared dictionary of around nine hundred words and an affix list of less than two hundred entries.

Parts of Speech Assigned and Their Abbreviations

The tape dictionary used for both analyses contained 73,582 words, with part-of-speech and word-status information from *The Shorter Oxford English Dictionary* (SOX)[2] and *Webster's Third New International Dictionary* (MW3).[3] The tape dictionary is reliable in most respects, since it was made from punched cards transcribed directly from the dictionaries, verified by different personnel, and spot-checked periodically during the process. Nevertheless, errors did occur, particularly in the recording of part-of-speech information which was not always understood by the keypunchers. The parts of speech recorded are as follows:

Noun	N	Adverb	AV	Pronoun	PN
Adjective	AJ	Preposition	PR	Interjection	IJ
Verb	VB	Conjunction	CJ	Past verb	PV

In addition, the category "other" (OT) was used whenever the dictionary gave some part of speech other than the nine listed above. Participles, numerals, articles, and collective nouns mainly comprise OT.

The algorithm was designed to assign these same nine parts of speech (excluding OT) with the addition of four more which were unfortunately subsumed under OT: present participle (PA), past participle (PP), auxiliary verb (AX), and plural or collective noun (NP). The category "noun" was changed to the category "noun-or-adjective" (NA) on the grounds that nearly all nouns can act as adjectives under some circumstances. Thus, although the algorithm attempts to distinguish words usable only as adjectives from those usable either as nouns or adjectives, it does not try to distinguish words usable only as nouns from those usable as either nouns or adjectives. Collective nouns will be assigned the string NA and NP to show possible use with either singular or plural verbs. Al-

*I wish to thank J. L. Dolby and H. L. Resnikoff, who have acted as consultants on Office of Naval Research contract Nonr 4440(00), which supported this research.

though a dictionary may show additional or fewer parts of speech for participial forms, their use (or lack of use) as nouns, adjectives, or verbs was considered implicit in the participle assignment, and no attempt was made to further partition the categories PA or PP. Thus, present participles are implicitly possible nouns, adjectives, or in a verb phrase, and past participles are implicitly adjectives, past verbs, or in a verb phrase. An attempt was made to identify participles which have any other special usages and to identify irregular past tense and past participial forms.

Like a dictionary, the algorithm is designed to indicate all the possible parts of speech for a word. That is, a part-of-speech string is assigned to each word, represented here by writing the part-of-speech abbreviations contiguously. For example, a word assigned the part-of-speech string AJ VB is a word that can act as an adjective or as a verb.

Design Plan

As a starting point in the design of a part-of-speech algorithm, three basic rules were postulated:

Rule A: The part-of-speech string associated with a word containing only one vowel string in its kernel will be NA VB, where a kernel will be defined as a word stripped of its affixes. Similarly, the part-of-speech string associated with words with multivowel string kernels will be NA.

Rule B: The part-of-speech string associated with a word ending in *ed* will be PP, and with a word ending in *ing* will be PA. All PP will also be considered PV. An NA classification will be changed to NP for all words ending in single *s*.

Rule C: The part-of-speech string associated with a word ending in *ly* will be AJ AV.

Rule A is basically a refinement of the original Dolby-Resnikoff[l] hypothesis and depends on the Dolby-Resnikoff definition of a legal vowel string. This rule also depends on the existence of an operational definition of affixes.[4,5] Rules B and C are a recognition of the most consistently used and meaningful suffixes of English.

A goal of 95 per cent accuracy was set for the algorithm. To reach that goal, three steps were decided upon:

Task 1: Tabulation of the exceptions to Rules B and C.

Task 2: Tabulation of special-purpose words, with part-of-speech PR, CJ, PN, or IJ, which are not covered by Rules A, B, or C.

Task 3: Modification of Rule A as much as necessary to achieve 95 per cent accuracy, using a study of affixes, or a tabulation of exceptions, or both, as a means to this end.

The first two tasks could be accomplished by sorting the dictionary on magnetic tape, as mentioned in the Introduction, although it may be of interest that not all of the necessary data handling could be accomplished with a generalized sort routine. The 7094 SORT was used in conjunction with special-purpose routines. The implementation of Tasks 1 and 2 is described in this paper; then the implementation of Task 3, which is more involved, is summarized with references for those who wish to pursue the details.

Dictionary Studies

TASK 1: EXCEPTIONS TO RULES B AND C

According to Rule B, all words ending in *ed, ing,* or single *s* should be categorized OT, for participle or noun-plural. All words violating this rule were listed and examined. Because many obscure and specialized words are listed in the dictionaries, it was decided that only words in standard usage would be included in exception lists. This reduced the list of Rule B exceptions somewhat, and further reduction was accomplished by removing the words ending in *as, is, ous,* and *us* whose part of speech would be properly inferred from these suffixes (see Task 3). Fortunately, many words ending in *ing* which are not participles could be removed because their actual parts of speech (usually NA, as for *pudding)* are subsumed under the participle heading. Classifying them as present participles is correct from the point of view of an "inclusive" part-of-speech string because present participles can be used as nouns or adjectives. (By an "inclusive" part-of-speech string is meant that string which is sure to contain all the parts of speech attributed to the word by either dictionary, but which may also contain one more or, rarely, two more parts of speech. Since use of inclusive part of speech becomes necessary in Task 3, its justification will be considered when Task 3 is discussed.) Similarly, words ending in *ed* which are not marked OT but are marked either AJ or VP are correctly classified past participle, from an inclusive viewpoint. All remaining *ed* and *ing* words, generally NA *ed* words and VB or AV *ing* words, are given in Table 1 along with the s-ending exception words. There are 104 words in this table, which is an exhaustive list.

Just as there are *ed, ing,* and *s*-ending words which are exceptions to Rule B, there are also some participles, past tense verbs, and plural or collective nouns which are exceptions because they cannot be recognized from *s, ing,* or *ed* endings. When all such words were listed from the dictionary, there were 1,380 entries, a very long list, since the goal of automatic determination of part of speech presupposes as small a dictionary as possible. From the list of 1,380 words, all irregular participles and past tense verbs have been

TABLE 1

EXCEPTION WORDS ENDING IN "S," "ED," OR "ING"

Word	POS	Word	POS
abed	AV	interbed	NA VB
alas	IJ	is	VB AX
aliped	NA	johannes	NA
anything	NA AV PN	kermes	NA
atlas	NA VB	lampers	NA
backstairs	NA	lens	NA
bathos	NA	lobsided	NA
bating	PR	mars	NA
beestings	NA	milliped	NA
besides	CJ PN	minus	NA AV PR
biped	NA	misdeed	NA
blewits	NA	mohammed	NA
bob-sled	NA	molasses	NA
bonus	NA VB	monied	NA
bow-string	NA VB	moses	NA
bring	NA VB IJ	naked	NA
callus	NA VB	nothing	NA AV PN
canvas	NA VB	outsing	VB
caucus	NA VB	outspring	VB
census	NA VB	palmiped	NA
chaos	NA	pathos	NA
childbed	NA VB	perhaps	NA AV
chorus	NA VB	pinniped	NA
circus	NA VB	plus	NA VB AV PR
clivers	NA	proceed	NA VB
coal-gas	NA	quadruped	NA
debarras	VB	rebus	NA VB
debus	VB	rhinoceros	NA
disfoliaged	NA	rugged	AJ AV
dogged	AJ AV	sacred	AJ
during	PR	sanders	NA
embarrass	VB	sideling	AJ AV
embed	NA VB	soliped	NA
embus	VB	something	NA VB PN AV
exceed	VB	sos	NA
excepting	PR CJ	stocking	NA VB
fissiped	NA	succeed	VB
focus	NA VB	theirs	PN
good-natured	NA	thermos	NA
goose-wing	NA VB	this	NA VB AJ PN
gossipred	NA	thoroughbred	NA
half-breed	NA	tight-laced	NA
hamstring	NA VB	thus	NA AV
hatred	NA	unstring	VB
hocus	NA VB	various	AJ PN
hoicks	VB IJ	vartabed	NA
hotbed	NA	water-gas	NA
hundred	NA	watershed	NA
imbed	NA VB	whereas	NA CJ
inbreed	VB	whing	NA VB IJ
indoors	AV	worsted	NA

listed in Table 2 (145 words). The rest of the words (1,235) included numerals, obscure collective nouns (e.g., herb, scrub), words which become collective only when *s* is added (e.g., geriatric), and some errors in judgment by the keypuncher. From this heterogeneous group, sixty were selected as reasonably common collective nouns and were listed in Table 3. Since the list is subjective, it may have to be augmented from experience, but it is believed to be adequate to maintain the goal of 95 per cent accuracy.

TABLE 2

IRREGULAR PARTICIPLES AND PAST TENSE VERBS

Word	POS	Word	POS
beaten	PP	made	AJ PP
begotten	PV PP	meant	PP
bet	NA VB PP	met	NA PP
bidden	PV PP	might	NA PV PP
bitten	PP	misgotten	PV PP
blent	PP	molten	PP
blest	PP	mown	PV PP
blown	NA PP	ought	NA VB AV PN PP
bode	NA VB PP	paid	AJ PP
born	AJ PV PP	pent	NA PP
borne	AJ PV PP	put	NA VB PV PP
bought	AJ VB PP	quit	NA VB PP PV
bound	NA VB PP	rang	VP
bounden	PP	read	NA VB VP PP
bracken	NA PP	reft	NA PP
braiden	PP	rent	NA VB PP
brant	NA PP	riven	PP
broke	NA VB PP	run	NA VB PP
broken	NA PP	rung	NA PP
brought	PV PP	said	AJ PP
built	NA VB PP	saw	NA VB PP
burst	NA VB PV PP	sent	NA PP
came	NA VB PP	sewn	PV PP
caught	AJ VB PV PP	shaken	PP
chosen	NA PP	shapen	PP
clad	AJ VB PP	shaven	PP
cleft	NA PV PP	shod	AV PP
clove	PV PP	shone	PV PP
cloven	PV PP	shook	NA VB PP
clung	PP	shorn	NA PV PP
could	AX	shot	NA VB PP
crept	PP	shotten	PP
cut	NA VB PP	should	AJ AX
did	PV PP	shrunk	PP
done	NA PV PP	slew	NA VB PP
drawn	PP	slit	NA VB PP
drew	PV PP	slunk	NA PP
driven	PP	smelt	NA VB PP
drove	NA PV PP	sodden	AJ VB PV PP
drunk	NA PP	sought	PP
drunken	PP	spent	VB PP
felt	NA VB PV PP	split	NA VB PP
flew	NA PP	spoke	NA VB PP
flown	PP	spoken	PP
fought	PP	spread	NA VB PV PP
fraught	AJ PP	sprung	PP
frozen	PP	spun	PP
gilt	NA PP	stole	NA VB PV PP
given	PP	stolen	NA PP
gone	NA PP	stricken	PV PP
got	PP	strung	PP
graven	PV PP	stung	PP
grew	PV PP	sung	PP
grit	NA VB PP	sunk	AJ PP
ground	NA VB PV PP	sunken	PP
had	PV PP	swam	PP
held	PP	sworn	PP
hewn	PP	swollen	PP
hidden	PP	taught	AJ PV PP
hung	VB PV PP	thought	PV PP
knit	NA VB PP	threw	PP
known	NA PP	thrown	AJ PP
lay	NA VB PP	thrust	NA VB PV PP
left	NA VB PP	told	PP
lent	PP	torn	PV PP
let	NA VB PP	trodden	PP

TABLE 2–*Continued*

went	PP	wound	NA VB PV PP
were	PP	wove	NA PP
wet	NA VB AV PP	woven	PV PP
widen	PV PP	written	PV PP
woke	PV PP	wrought	AJ PP
worn	PP	wrung	PP
would	NA AX		

TABLE 3

IRREGULAR PLURAL OR COLLECTIVE NOUNS

Word	POS	Word	POS
apache	NA NP	marabou	NA NP
cattle	NA VB NP	maxima	NP
carp	NA VB NP	mice	NP
caribou	NA NP	milanese	NA NP
chinese	NA NP	men	NP
chinook	NP	pence	NP
cherubim	NA NP	people	NA VB NP
couple	NA VB NP	perch	NA VB NP
crane	NA VB NP	pike	NA VB NP
crustacea	NP	poultry	NA NP
cutlery	NA NP	regalia	NA NP
data	NP	rice	NA VB NP
dice	NA VB NP	roe	NA NP
dicta	NP	secreta	NA NP
fish	NA VB NP	seraphim	NA NP
foe	NA NP	sheet	NA VB NP
fulcra	NA NP	snipe	NA VB NP
game	NA VB NP	spawn	NA VB NP
geese	NP	sperm	NA NP
genera	NP	spoor	NA VB NP
grouse	NA VB NP	squid	NA VB NP
help	NA NP IJ VB	starfish	NA NP
hosiery	NA NP	steer	NA VB NP
ice	NA VB NP	strata	NP
ingesta	NP	swine	NP
irish	NA AJ NP	tripe	NA NP
japanese	NA NP	tuna	NA NP
lice	NP	viscua	NP
like	NA VB AV PR CJ NP	women	NP
lynx	NA NP	young	NA NP

In investigating exceptions to Rule C, adverbs with additional parts of speech of PR, CJ, PV, IJ, PN, and OT were ignored in order to avoid duplication of words with those in lists compiled in Task 2. Within this limitation, all words were extracted from the dictionary which, though ending in *ly,* were not adverbs or, conversely, though not ending in *ly,* were adverbs. Contrary to expectations, there was a large number of such words (slightly over 1,500). Many of these words were judged rare, or rare in the usage in question (e.g., dog-fly as NA, or dash, pi, rife, smell, thistle as AV); others could be predicted by an extension of the affix lists, to be discussed later. In accordance with the philosophy of maintaining a relatively short exception list without sacrificing too much accuracy, this list of 1,500 words has been arbitrarily reduced to a list of 361 of the common words which are exceptions to Rule C, as shown in Table 4. In addition, there are many non-*ly* adverbs which occur in Table 5.

TABLE 4

COMMON EXCEPTION WORDS TO RULE C

Word	POS	Word	POS
backhand	NA VB AV	hard	NA AV
bare-backed	AJ AV	hence	NA VB AV
bare-headed	AJ AV	here	AV
between-whiles	NA AV	home	NA AV
co-ally	NA	ill	NA VB AV
cock-sure	AJ AV	just	NA AV
counter-clockwise	AJ AV	keen	NA VB AV
counter-current	NA AV	large	NA VB AV
criss-cross	NA VB AV	last	NA AV
cross-country	NA AV	late	AJ AV
cross-grained	AJ AV	lax	NA VB AV
double-quick	NA VB AV	least	NA AV
end-ways	AJ AV	long	NA VB AV
edge-ways	NA	loose	NA VB AV
free-hand	NA AV	loud	AJ AV
god-damn	NA VB AV	low	AJ AV
half-and-half	NA AV	maybe	NA VB AV
half-way	NA AV	mean	NA AV
happy-go-lucky	NA AV	much	NA VB AV
harum-scarum	NA AV	needs	NA AV
now-a-days	NA AV	new	AV
off-hand	AJ AV	loud	NA AV
oft-times	AV	nope	NA AV
old-fashioned	NA AV	north	NA VB AV
over-hard	AJ AV	odd	NA AV
over-long	AJ AV	oft	AV
over-supply	NA VB	old	NA AV
point-blank	NA AV	part	NA VB AV
post-haste	NA AV	pat	NA VB AV
pot-belly	NA	prompt	NA VB AV
right-handed	AJ AV	queer	NA VB AV
rough-and-tumble	NA AV	quick	NA VB AV
second-class	NA AV	quite	NA AV
side-saddle	NA VB AV	real	NA AV
single-handed	AJ AV	right	NA VB AV
sky-high	AJ AV	sic	NA VB AV
so-and-so	NA AV	snug	NA VB AV
topsy-turvy	NA VB AV	soon	AJ AV
under-arm	AJ AV	sour	NA VB AV
up-country	NA AV	square	NA VB AV
up-grade	NA VB AV	straight	NA VB AV
up-stream	AJ AV	thence	AV
up-wind	NA VB AV	twice	NA AV
aft	NA AV	west	NA VB AV
back	NA VB AV	worse	NA AV
bad	NA AV	wrong	NA VB AV
blind	NA VB AV	yea	NA AV
broad	NA AV	yep	AV
cheap	NA AV	yes	NA VB AV
clean	NA VB AV	ablaze	AJ AV
damn	NA VB AV	adrift	AJ AV
double	NA VB AV	afield	AJ AV
east	NA VB AV	aground	AJ AV
faint	NA VB AV	ajar	AJ AV
fair	NA VB AV	akin	AJ AV
false	AJ AV	alias	NA AV
fine	NA VB AV	alike	AJ AV
flat	NA VB AV	alive	AJ AV
flush	NA VB AV	almost	AJ AV
forte	NA AV	alone	AJ AV
foul	NA VB AV	aloud	AV
free	AJ VB AV	always	AV
fresh	NA VB AV	amuk	NA AV
front	NA VB AV	andante	NA AV
full	NA VB AV	apart	AJ VB AV

TABLE 4—*Continued*

apiece	AV
aright	AV
askew	AJ AV
astray	AJ AV
away	AJ AV
awful	AJ AV
awhile	AV
broadcast	NA VB AV
broadside	NA VB AV
broadway	NA AV
complete	AJ VB AV
costly	AJ
counter	NA VB AV
curly	AJ
direct	NA VB AV
dirty	AJ VB AV
doily	NA
doubtless	AJ AV
earthly	AJ
even	NA VB AV
ever	AV
farther	AJ VB AV
farthest	AJ AV
further	AJ VB AV
furthest	AJ AV
galore	NA AV
gratis	AJ AV
gully	NA VB
hearty	NA AV
heaven	NA VB AV
herein	AV
hereof	AV
higher	NA AV
highest	NA AV
hilly	AJ
holly	NA
holy	NA
imply	VB
indeed	AV
indoor	AJ AV
indoors	AV
jelly	NA VB
july	NA
largo	NA AV
later	NA AV
latest	NA AV
lengthways	AV
lento	AJ AV
lesser	AJ AV
lily	NA
longways	NA AV
lower	NA VB AV
lowest	NA AV
matchless	AJ AV
measly	AJ
merry	NA AV
middling	NA AV
midstream	NA AV
mighty	NA AV
molly	NA
never	AV
nohow	AV
noways	AV
offshore	AJ AV
offside	NA AV
often	AV
open	NA VB AV
outboard	NA AV
outright	AJ AV
perchance	AV
perforce	NA AV
perhaps	NA AV
piano	NA AV
plenty	NA AV
pronto	AV
proper	NA AV
rally	NA VB
ready	NA VB AV
reckless	AJ AV
reply	NA VB
restless	NA AV
reverse	NA VB AV
sally	NA VB
scaly	AJ
seldom	AJ AV
sheepish	AJ AV
slantways	AV
slantwise	AJ AV
smelly	AJ
sooner	NA AV
speedy	AJ AV
starboard	NA VB AV
steadfast	AJ AV
steady	NA VB AV
sudden	NA AV
sully	VB
tally	NA VB
thereat	AV
thereof	AV
threefold	AJ AV
tidy	NA VB AV
topside	NA AV
twofold	NA AV
upright	NA VB AV
very	NA AV
vivace	AJ AV
weary	AJ VB AV
wellnigh	AV
whereat	AV
wherein	AV
whereof	AV
whereon	AV
wily	AJ
abundant	AJ AV
adagio	NA AV
aflutter	AJ AV
afterward	NA AV
afterwards	NA AV
aglitter	AJ AV
akimbo	AJ AV
alibi	NA VB AV
alongshore	NA AV
already	AV
amidships	AJ AV
anywhere	NA AV
apriori	NA AV
bareback	AJ AV
barefoot	AJ AV
butterfly	NA VB
careless	NA AV
cowardly	AJ VB
crescendo	NA VB AV
elsewhere	AV
evermore	AV
extempore	AJ AV
falsetto	NA AV
family	NA
forehand	NA AV

TABLE 4—*Continued*

foremost	AJ AV
forever	NA AV
forzando	AJ AV
furthermore	AV
henceforth	AV
hereabout	AV
hereafter	NA AV
hereby	AV
hitherto	AJ AV
homily	NA
however	AV
howsoever	AV
hundredfold	NA AV
impromptu	NA VB AV
inasmuch	AV
innuendo	NA VB AV
insomuch	AV
legato	NA AV
lentamente	AJ AV
lifelong	AJ AV
manyways	AV
miserly	AJ
nevermore	AV
ninefold	NA AV
outermost	AJ AV
overboard	AV
overhand	NA VB AV
overhead	NA AV
overland	NA AV
overnight	NA VB AV
overtime	NA VB AV
piecemeal	NA VB AV
sevenfold	AJ AV
sforzando	NA AV
sforzato	AJ AV
sideway	NA AV
sideways	AJ AV
sixtyfold	NA AV
somehow	AV
sometime	AJ AV
someway	AV
somewhere	NA AV
staccato	NA VB AV
straightaway	NA AV
thenceforth	AV
thereabout	AV
thereabouts	AV
thereafter	AV
thereby	AV
therefore	NA AV
thereto	AV
thereupon	AV
thousandfold	NA AV
twelvefold	AJ AV
unaware	AJ AV
underground	NA AV
underhand	NA VB AV
ungodly	AJ
unholy	NA
unruly	AJ
unsightly	AJ
unworldly	AJ
uppermost	AJ AV
upriver	AJ AV
verbatim	NA AV
whereabout	NA AV
whereby	AV
wherefore	NA AV
whereupon	AV
wholesale	NA VB AV
yesterday	NA AV
altogether	NA AV
beforehand	AJ AV
contrariwise	AJ AV
everyway	AV
everywhere	NA AV
fortissimo	NA AV
henceforward	AV
heretofore	NA AV
incognito	NA AV
malapropos	NA AV
melancholy	NA
moderato	AJ AV
monopoly	NA
nevertheless	AV
oftentimes	AV
pianissimo	NA AV
pizzicato	NA AV
prestissimo	NA AV
sometimes	AV
thenceforward	AV
unawares	AV
underhanded	AJ AV

TASK 2: TABULATION OF SPECIAL-PURPOSE WORDS WHICH ARE NOT COVERED BY RULES A, B, OR C

For Task 2, a subset of the dictionary was prepared containing all the words which: (1) have at least one standard meaning corresponding to a part of speech other than NA, VB, AJ, or AV (the parts of speech assigned by Rules A, B, C), (2) have all "irregular" entries removed (fragments, etc.), and (3) have all words ending in *ed, ing,* or *s* removed (the suffixes covered by Rule B). By extracting from this subset all words with standard meaning corresponding to a part of speech PR, CJ, IJ, PN, or OT, we should get an exhaustive list of those structural, special-purpose words which are so important in a mechanized handling of English.

Table 5 shows the 253 function words so extracted.

TABLE 5

SPECIAL-FUNCTION WORDS

Word	POS	Word	POS
he	NA VB PN IJ	plunk	NA VB AV IJ
she	NA PN OT	whisk	NA VB AV IJ
the	AJ AV OT	all	NA AV PN OT
me	NA PN	fell	NA VB AV PV
a	NA VB PR OT	well	NA VB AV OT
dead	NA AV OT	till	NA VB PR CJ
mid	NA AV PR	still	NA VB AV CJ
bold	NA AV OT	him	NA PN
and	NA AV CJ	whom	PN
beyond	NA AV PR	from	PR
round	NA VB AV PR	cum	AJ AV PR
thud	NA VB AV IJ	than	PR CJ
whence	NA AV CJ	been	PV
since	AV PR CJ	then	NA AV CJ
once	NA AV CJ	when	NA AV CJ PN
bounce	NA VB AV IJ	in	NA VB AV PR
jee	NA AV IJ	lain	PV
strange	AJ AV IJ	on	NA AV PR
like	NA VB AV PR C OT	con	NA VB AV PR
		down	NA VB AV PR
while	NA VB CJ	who	NA PN IJ
vile	AJ AV OT	no	NA AV
same	NA AV PN	pro	NA AV PR
some	AJ AV PN OT	so	NA AV CJ PN
thine	PN OT	to	NA AV PR
mine	NA VB PN	sweep	NA VB AV
one	NA VB PN OT	plop	NA VB AV IJ
none	NA AV PN OT	pop	NA VB AV IJ
prone	NA AV OT	up	NA VB AV PR
woe	NA AV IJ OT	bar	NA VB PR
ere	NA AV PR CJ	dear	NA VB AV IJ
there	NA AV PN	near	NA VB AV PR
where	NA AV CJ PN	her	NA PN
maugre	AV PR	per	AJ PR
fore	NA AV PR IJ	or	NA PR CJ
more	NA AV PN	for	NA PR CJ
wise	NA VB AV OT	nor	CJ
whose	AJ PN OT	whirr	NA VB AV IJ
ante	NA VB AV PR	cross	NA VB AV PR
save	NA VB PR CJ	at	NA AV PR CJ PN
bove	AV PR	neat	NA AV OT
aye	NA AV IJ	great	NA AV OT
off	NA VB AV PR	that	NA AV CJ PN
if	NA CJ	what	NA AV CJ PN IJ OT
self	NA VB PN	wet	NA VB AV OT
of	NA PR OT	yet	AJ AV CJ
stag	NA VB AV OT	left	NA AV OT
dang	NA VB AV OT	light	NA VB AV OT
whang	NA VB AV OT	aught	NA AV PN
each	AJ AV PN	caught	AJ VB PA OT
which	AJ PN	ought	NA VB AV PN OT
rich	NA AV	it	NA PN
such	AJ AV PN	not	NA AV PR
nigh	NA VB AV PR	spot	NA VB AV OT
though	AV CJ	fast	NA VB AV IJ
through	NA AV PR	past	NA AV PR OT
plash	NA VB AV IJ	midst	NA AV PR
swash	NA VB AV IJ	best	NA VB AV OT
swish	NA VB AV IJ	lest	CJ
with	NA AV PR	most	NA AV PN OT
both	NA AV CJ PN	but	NA AV PR CJ PN
south	NA VB AV PR	out	NA VB AV PR IJ
crack	NA VB AV IJ	bout	NA VB AV PR
stock	NA VB AV OT	next	NA AV PR
rank	NA VB AV OT	you	NA PN OT

TABLE 5—*Continued*

how	NA AV CJ IJ	whither	AV CJ
now	NA AV CJ	other	NA AV PN
ay	NA AV IJ	after	NA AV PR CJ
by	NA VB AV PR	better	NA VB AV OT
why	NA AV IJ	whoever	PN
whizz	NA AV IJ	over	NA VB AV PR
ana	NA AV	atour	AV PR
supra	AV PR	across	AJ AV PR
contra	NA AV PR	unless	NA PR CJ
instead	AV OT	abaft	AV PR
abroad	AJ AV PR	outwrought	PA OT
amid	PR	albeit	CJ
inland	NA AV OT	howbeit	AV CJ
behind	NA AV PR	aslant	AJ AV PR
around	AJ AV PR	except	VB PR CJ
aboard	AV PR	athwart	AV PR
toward	AJ AV PR	amort	AV OT
astride	AV PR	amidst	NA AV PR
aside	NA AV PR	amongst	PR
beside	AV PR	against	PR
inside	NA AV PR	midmost	NA AV PR
outside	NA AV PR	adust	NA AV OT
unlike	NA AV PR CJ	about	AJ AV PR
before	NA AV PR CJ OT	throughout	AJ AV PR
because	AV CJ	without	NA AV PR
despite	NA PR	betwixt	AV PR
above	NA AV PR	adieu	NA AV IJ
himself	NA PN	below	NA AV PR
herself	PN	midway	NA AV PR
ourself	PN	bully	NA VB AV IJ
yourself	PN	only	AJ AV PR CJ
itself	PN	any	AJ AV PN OT
myself	PN	alongside	AV PR
along	AV PR	opposite	NA AV PR
endlong	AV PR	oneself	PN
among	PR	sidelong	AJ AV PR
anigh	AV PR	underneath	NA AV PR
although	CJ	wherewith	NA AV PN
enough	NA AV IJ	unequal	NA AV OT
awash	AJ AV OT	overall	NA AV OT
beneath	AJ AV PR	unbeknown	NA AV OT
argal	NA AV CJ	another	AJ PN
until	PR CJ	whichever	AJ PN
bidden	PV	whomever	PN
between	NA AV PR	whenever	AV CJ
amen	NA VB AV IJ	whensoever	AV CJ
certain	NA AV OT	whosoever	AJ PN
within	NA AV PR	wherever	AV CJ
upon	PR	whatever	AJ AV PN
ago	AJ AV	somewhat	NA AV PN
into	PR	unbethought	AV OT
presto	NA VB AV IJ	amidmost	AV PR
asleep	AJ AV OT	undermost	AJ AV OT
atop	AJ AV PR	anyhow	AV OT
anear	AV PR	anyway	AV OT
yonder	AJ AV PN OT	bimonthly	NA AV OT
under	NA AV PR	instantly	AV CJ OT
rather	AV OT	exenterate	AJ VB AV OT
whether	NA AV CJ	wherewithal	NA AV PN
either	AJ AV CJ PN	anybody	NA PN
neither	AJ AV CJ PN	everybody	PN
		immediately	AV CJ

The words are listed in groups according to number of syllables and are arranged alphabetically from the end of the word. Note that Table 1 lists the eighteen function words ending in *s* or *ing*. This list is otherwise

theoretically complete, but because of a misunderstanding by keypunchers in the original creation of the dictionary, some important pronouns were not so classified in the MW3 part-of-speech designations and are therefore missing from the list (I, your, his, we, them, our, us, their, they). Similarly, some important auxiliary verbs were not so classified in the SOX part-of-speech designations and are therefore missing (am, is, are, was, were, be, will). Also, the word as has been lost in the sorting process. No other significant omissions have been noted, but are possible, since checking of the tape dictionaries was not exhaustive. For the convenience of the reader, the words in Tables 1 through 5, plus the words given here, have been alphabetized and given in Table 6.

The parts of speech given in Tables 1 through 5 were taken from the tape dictionary and have not been verified in the dictionaries themselves. Particular care should be taken in the use of Table 2, which seems to have many errors in the omission or intrusion of the PV and PP codes.

TABLE 6

EXCEPTION WORDS

a	amuk	been	butterfly	cut	faint	gratis
abaft	ana	beestings	by	cutlery	fair	graven
abed	and	before	callus	damn	false	great
ablaze	andante	beforehand	came	dang	falsetto	grew
aboard	anigh	begotten	canvas	data	family	grit
about	another	behind	careless	dead	farther	ground
above	ante	below	caribou	dear	farthest	grouse
abroad	any	beneath	carp	debarras	fast	gully
abundant	anybody	beside	cattle	debus	fell	had
across	anyhow	besides	caucus	despite	felt	half-and-half
adagio	anything	best	caught	dice	fine	half-breed
adieu	anyway	bet	census	dicta	fish	hamstring
adrift	anywhere	better	certain	did	fissiped	happy-go-lucky
afield	apache	between	chaos	direct	flat	hard
aflutter	apart	between-while	cheap	dirty	flew	harum-scarum
aft	apiece	betwixt	cherubim	disfoliaged	flown	has
after	apriori	beyond	childbed	dogged	flush	hatred
afterward	are	bidden	chinook	doily	focus	he
afterwards	aright	bimonthly	chorus	done	foe	hearty
against	around	biped	chosen	double	for	heaven
aglitter	as	bitten	circus	double-quick	fore	held
ago	aside	blent	clad	doubtless	forehand	help
aground	askew	blest	clean	down	foremost	hence
ajar	aslant	blewits	cleft	drawn	forever	henceforth
akimbo	asleep	blind	clivers	drew	forte	henceforward
akin	astray	blown	clove	driven	fortissimo	her
alas	astride	bobsled	cloven	drove	forzando	here
albeit	at	bode	clung	drunk	fought	hereabout
alias	athwart	bold	coal-gas	drunken	foul	hereafter
alibi	atlas	bonus	co-ally	during	fraught	hereby
alike	atop	born	cock-sure	each	free	herein
aliped	atour	borne	complete	earthly	free-hand	hereof
alive	aught	both	con	east	fresh	heretofore
all	awash	bought	contra	edge-ways	from	herself
almost	away	bounce	contrariwise	either	front	hewn
alone	awful	bound	costly	elsewhere	frozen	heynne
along	awhile	bounden	could	embarras	fulcra	hidden
alongshore	ay	bout	counter	embed	full	higher
alongside	aye	bow-string	counter-clockwise	embus	further	highest
aloud	back	bracken	counter-current	endlong	furthermore	hilly
already	backhand	brant	couple	end-ways	furthest	him
although	backstairs	bring	cowardly	enough	galore	himself
altogether	bad	broad	crack	ere	game	his
always	bar	broadcast	crane	even	geese	hitherto
am	bare-backed	broadside	crept	ever	genera	hocus
amen	bare-headed	broadway	crescendo	evermore	gilt	hoicks
amid	bareback	broke	criss-cross	everybody	given	holly
amidmost	barefoot	broken	cross	everyway	god-damn	holy
amidships	bathos	brought	cross-country	everywhere	gone	home
amidst	bating	built	cross-grained	exceed	good-natured	homily
among	be	bully	crustacea	except	goose-wing	hosiery
amongst	beaten	burst	cum	excepting	gossipred	hotbed
amort	because	but	curly	extempore	got	how

TABLE 6—*Continued*

howbeit
however
howsoever
hundred
hundredfold
hung
I
ice
if
ill
imbed
immediately
imply
impromptu
in
inasmuch
inbreed
incognito
indeed
indoor
indoors
ingesta
inland
innuendo
inside
insomuch
instantly
instead
interbed
into
irish
is
it
itself
japanese
jee
jelly
johannes
july
just
keen
kermes
knit
known
lain
lampers
large
largo
last
late
later
latest
lax
lay
least
left
legato
lengthways
lens
lent
lentamente
lento
lesser
lest
let
lice
lifelong
light
like
lily
lobsided
long
longways
loose
loud
low
lower
lowest
lynx
made
malapropos
manyways
marabou
mars
matchless
maxima
maybe
me
mean
meant
measly
melancholy
men
merry
met
mice
mid
middling
midmost
midst
midstream
midway
might
mighty
milanese
milliped
mine
minus
misdeed
miserly
misgotten
moderato
mohammed
molasses
molly
molten
monied
monopoly
more
moses
most
mown
much
myself
naked
near
neat
needs
neither
never
nevermore
nevertheless
new
next
nigh
ninefold
no
nohow
none
nope
nor
north
not
nothing
now
now-a-days
noways
odd
of
off
off-hand
offshore
offside
oft
oft-times
often
oftentimes
old
old-fashioned
on
once
one
oneself
only
open
opposite
or
other
ought
our
ourself
out
outboard
outermost
outright
outside
outsing
outspring
over
over-hard
over-long
over-supply
overall
overboard
overhand
overhead
overland
overnight
overtime
paid
palmiped
part
past
pat
pathos
pence
pent
people
per
perch
perchance
perforce
perhaps
pianissimo
piano
piecemeal
pike
pinniped
pizzicato
plenty
plop
plunk
plus
point-blank
pop
post-haste
pot-belly
poultry
prestissimo
presto
pro
proceed
prompt
prone
pronto
proper
put
quadruped
queer
quick
quit
quite
rally
rang
rank
rather
read
ready
real
rebus
reckless
reft
regalia
rent
reply
restless
reverse
rhinoceros
rice
rich
right
right-handed
riven
roe
rough-and-tumble
round
rugged
run
rung
sacred
said
sally
same
sanders
save
saw
scaly
second-class
secreta
seldom
self
sent
seraphim
sevenfold
sewn
sforzando
sforzato
shaken
shapen
shaven
she
sheep
sheepish
shod
shone
shook
shorn
shot
shotten
should
shrunk
sic
side-saddle
sideling
sidelong
sideways
since
single-handed
sky-high
slantways
slantwise
slew
slit
slunk
smelly
smelt
snipe
snug
so
so-and-so
sodden
soliped
some
somehow
something
sometime
sometimes
someway
somewhat
somewhere
soon
sooner
sos
sought
sour
south
spawn
speedy
spent
sperm
splash
split
spoke
spoken
spoor
spot
spread
sprung
spun
square
squid
staccato
stag
starboard
starfish
steadfast
steady
steer
still
stock
stocking
stole
stolen
straight
straightaway
strange
strata
stricken
strung
stung
succeed
such
sudden
sully
sung
sunk
sunken
supra
swam
swash
sweep
swine
swish
swollen
sworn
tally
taught
than
that
the
their
theirs
them
then
thence
thenceforth
thenceforward
there
thereabout
thereabouts
thereafter
thereat
thereby
therefore
thereof
thereto
thereupon
thermos
they
thine
this
thoroughbred
though
thought
threefold
threw
through
throughout
thrown
thrust
thud
thus
tidy
tight-laced
till
to
told
topside
topsy-turvy
torn
toward
tripe
trodden
tuna
twice
twofold
unaware
unawares
unbeknown
unbethought
under
under-arm
underground
underhand
underhanded
undermost
underneath
unequal
ungodly
unholy
unless
unlike
unruly
unsightly
unstring
until
unworldly
up
up-country
up-grade
up-stream
up-wind
upon
uppermost
upright
upriver
us
various
vartabed
verbatim
very
vile
viscua
vivace
was
water-gas
watershed
we
weary
well
wellnigh
went
were
west
wet
whang

TABLE 6–*Continued*

what	whereby	which	wholesale	with	wove	yesterday
whatever	wherefore	whichever	whom	within	woven	yet
when	wherein	while	whomever	without	written	yonder
whence	whereof	whing	whose	woe	wrong	you
whenever	whereon	whirr	whosoever	woke	wrote	young
whensoever	whereupon	whisk	why	worn	wrought	your
where	wherever	whither	widen	worse	wrung	yourself
whereabout	wherewith	whizz	will	worsted	yea	
whereas	wherewithal	who	wily	would	yep	
whereat	whether	whoever	wise	wound	yes	

TASK 3: MODIFICATION OF RULE A USING A STUDY OF AFFIXES

Rule A is based upon a general observation and is good for only a simple majority of words. The business of Task 3 is to discover if it is possible, by considering prefixes and suffixes, to convert this general rule to a more precise rule, adequate for 95 per cent of English words. As a first step, a formal and reproducible definition for affixes was developed, as is described in *The Nature of Affixing in Written English** and *Structural Definition of Affixes in Multisyllable Words.*[5] Then, the extent of correlation between affixes and part of speech was investigated, both for the formally defined affixes and for others listed in *Modern English* Usage.[6] This investigation is described in "Part-of-Speech Implications of Affixes"[7] but can be summarized here.

All words with part of speech AV, PR, PN, NP, IJ, PA, PP, VP, and CJ can be automatically assigned part of speech by reference to the word lists in Tables 1 through 4, followed by application of Rules B and C for words not in these lists. "Part-of-Speech Implications of Affixes"[7] was therefore concerned only with words whose part-of-speech string contained the elements NA, AJ, and VB, which allows the five possible combinations VB, NA, AJ, NA-VB, AJ-VB. NA-AJ is considered equivalent to NA. Attempts to establish a 95 per cent correlation between the part-of-speech string of a word and its affixes failed. However, it was noted that the correlation was closer for four- to seven-syllable words than for two- to three-syllable words and that a very good correlation could be obtained for all words between an "inclusive" part-of-speech string and the affixes. Thus, in some cases determining the affixes and counting vowel strings lead to an absolute identification of the part of speech of a word, but in other cases identification is to a more inclusive set. For example, an NA or a VB may be classified as NA-VB, or an AJ may be classified as an NA. Such a classification is justifiable on the following grounds: (1) A primary use of part-of-speech information is in automatic syntactic analysis. It is the natural task of a syntactic analysis program to choose among several possible parts of speech, and it is easier to do so than to supply a missing part of speech. (2) Dictionaries are very reliable in the information explicitly given, but implications inferred from the absence of information are less reliable. Thus, the inclusive part-of-speech string assigned by the algorithm may in some cases be more correct than the more limited one assigned by a particular dictionary. In our experience with the SOX and MW3 dictionaries, we found many instances of non-agreement; usually one was more inclusive than the other.

In "Part-of-Speech Implications of Affixes,"[7] the results of the correlation study are given for seventy-two prefixes and eighty-seven suffixes. Implications are of the form NA or NA-VB, or VB or AJ. For example, the four s-ending suffixes mentioned in the discussion of Task 2 carry the following part of speech implications :

is	NA-VB	*as*	NA
ous	AJ	*us*	NA

For forty-one of the affixes, the part-of-speech implication changes with the length of the word, from NA-VB for two- and three-syllable words to NA for four- to eight-syllable words.

Later a correlation was made for other affixes which seemed to be likely candidates for reducing the exception lists by aiding in the identification of adverbs or in the identification of words ending in *ed* which are not past participles. Though not operationally defined, these affixes are of practical importance and are therefore listed here, with their part-of-speech implications:

Prefixes	POS	Suffixes	POS
north	NA AV	seed	NA
south	NA AV	weed	NA
west	NA AV	like	NA AV
a-	AJ AV	wise	AJ AV
		ward	NA AV
		wards	NA AV
		-fly	NA
		-bed	NA
		-deed	NA
		-feed	VB
		-tenths	NA

Testing and Evaluation

Rules A, B, and C, the exception lists, and the prefix and suffix implications reported in Reference 7 formed the basis of a part-of-speech algorithm, which has been programed on the IBM 7090 and is being implemented on the IBM 360/30. In the program, a word whose part of speech is to be determined is first checked against the exception lists, which yield a part-of-speech string for words which match. For all other words, the word is separated into kernel and affix parts, and the part-of-speech implication of the affixes is looked up and applied to the word. For any word without affixes or whose affixes do not have an implication, Rule A is applied to obtain the part-of-speech assignment. There are some complications involved in some of these steps, particularly in separating a word into kernel and affix parts and in assigning parts of speech on the basis of affixes. The logic used by the program for these steps is given in Figure 1.

To summarize the logic briefly, we can say that affixes are stripped from the word one at a time, with prefixes given a limited priority over suffixes other than *ed.* Thus, the word *exceptional* becomes first *ex-ceptional,* then ex-*ception*-al, and finally ex-*cep*-tion-al. The criterion by which an affix sequence was accepted was for most affixes the same as that given in Reference 7; simply stated, this means that the affix was accepted if the remaining kernel was a reasonable syllable or syllables, determined by examining the consonant and vowel strings. Some affixes were designated as transformational and were subject to additional constraints or modifications. For example, *s* is a suffix only at the end of a word and when not preceded by another s. The implications of the outermost affixes were used in assigning parts of speech, and the priority indicators were set to use suffix implications, if any, in preference to prefix implications, in accordance with the findings of Reference 7.

To test the algorithm, five hundred words were chosen at random from the tape dictionary, [2,3] and the parts of speech assigned by the algorithm were compared with those given in the dictionary. If dialectal, obsolete, archaic, and rare words causing errors are removed, and if program errors are corrected, results are as follows:

Category	No. of Words in Category
Assigned POS matches dictionary POS	271
Extra POS assigned	196
Missing POS	16
POS does not match at all—error	8
Total sample	491

This shows that 95.1 per cent of the words were assigned the correct inclusive part of speech and 55.2 per cent were assigned parts of speech exactly coinciding with those assigned by the dictionary. Thus, the goal of 95 per cent is just achieved.

It is interesting to consider how little the affix implications have improved the results for this sample. Taking the first 192 of the five hundred alphabetized words and applying the original Rules A, B, and C only, twenty words are shifted into the exact-match category and twenty-five words shifted from the exact-match category, for a net loss of five words, where two of these go into the error category. Six words are added to the words with missing part of speech, while two words are taken out of the category. Thus, the total loss is four more words into the missing category and two more words into the error category, or about a 3 per cent loss from the point of view of inclusive part of speech. Rule A, it will be remembered, requires the removal of affixes from the kernel of the word. If this kernelizing of the word is omitted, there is about a 13 per cent loss from the point of view of inclusive part of speech, indicating that the fact that a word is affixed is more important in predicting part of speech than what the affix is (the affixes *ing, ed, ly,* and s excepted). Nevertheless, using the implications of affixes is a refinement in an area where refinement is sorely needed.

It might be interesting at this point to evaluate the two original premises—that one-syllable words are largely noun-verb and that all other words are largely noun only.[1] Although the tape dictionary does not provide a syllable count, it does provide a count of the number of legitimate vowel strings; final *e* is not to be considered legitimate. To test the first premise, the standard one-vowel-string words in the tape dictionary were divided into two sections, those which were NA-VB (and only NA-VB) and those which were not (the OT category was ignored). There were 2,520 words in the NA-VB category and 1,925 words with more or fewer parts of speech than NA-VB. The 1,925-word list includes the 132 one-vowel-string members of the word-class with parts of speech PR, CJ, IJ, PN, and PV listed in Table 4. Discounting these 132 function words, then, the first premise is true for 2,520 out of 4,313 cases, or about 58 per cent. To get 95 per cent of the one-vowel-string words assigned as in the dictionary, most of the 1,793 non-NA-VB words would have to be in an exception dictionary. However, since most of these are NA, from the point of view of inclusive part of speech, the NA-VB rule for one-vowel-string words is quite good, giving results very close to those obtained in the five-hundred-word random sample of all words (55 per cent exactly matching dictionary, 95 per cent giving correct inclusive part of speech). Note that these statistics hold for one-vowel-string words and that the statistics for one-syllable words would differ somewhat.

The second premise has not been directly tested, but may be inferred from the five-hundred-word

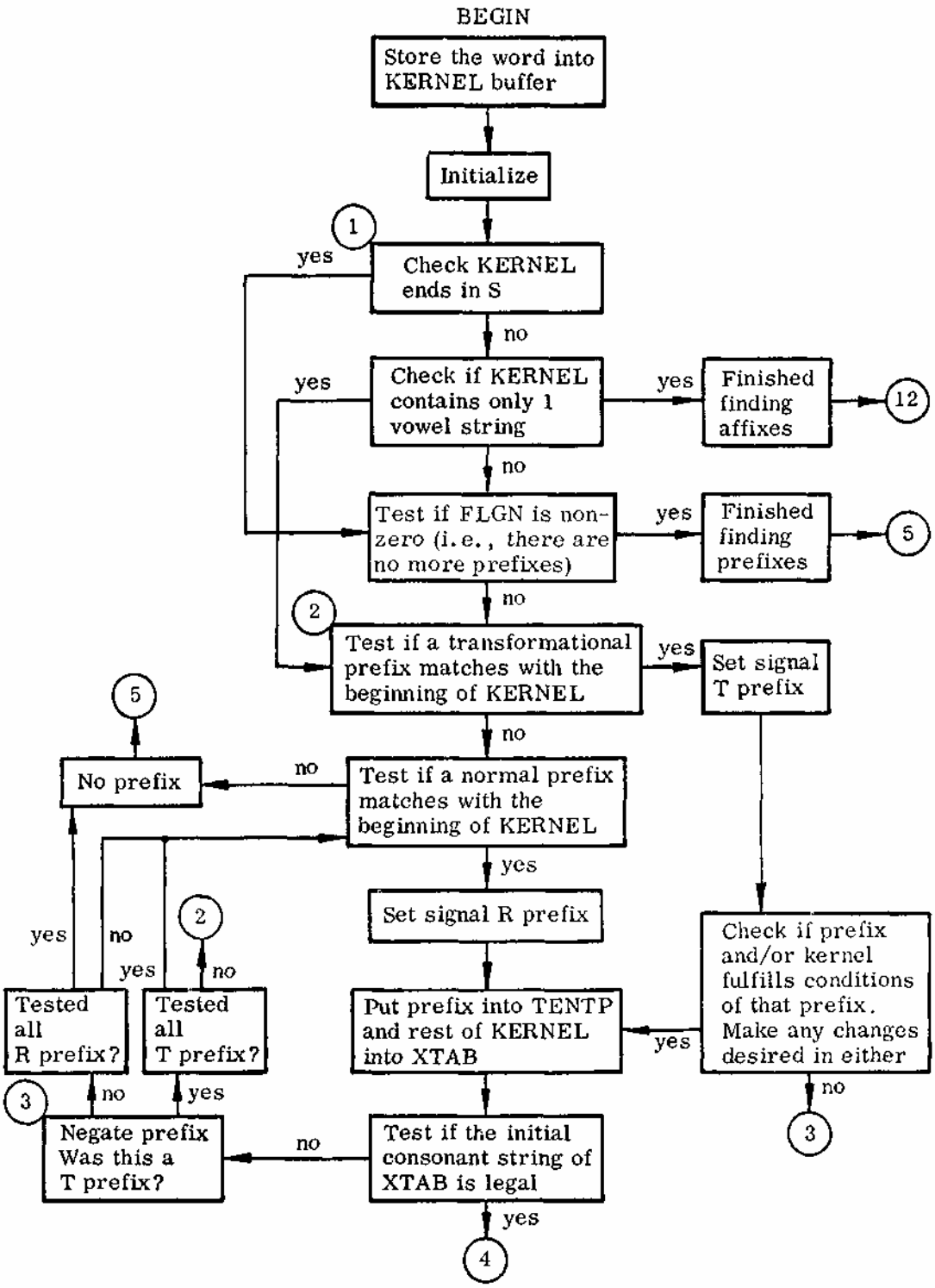

FIG. 1.—Search-for-affixes flow diagram

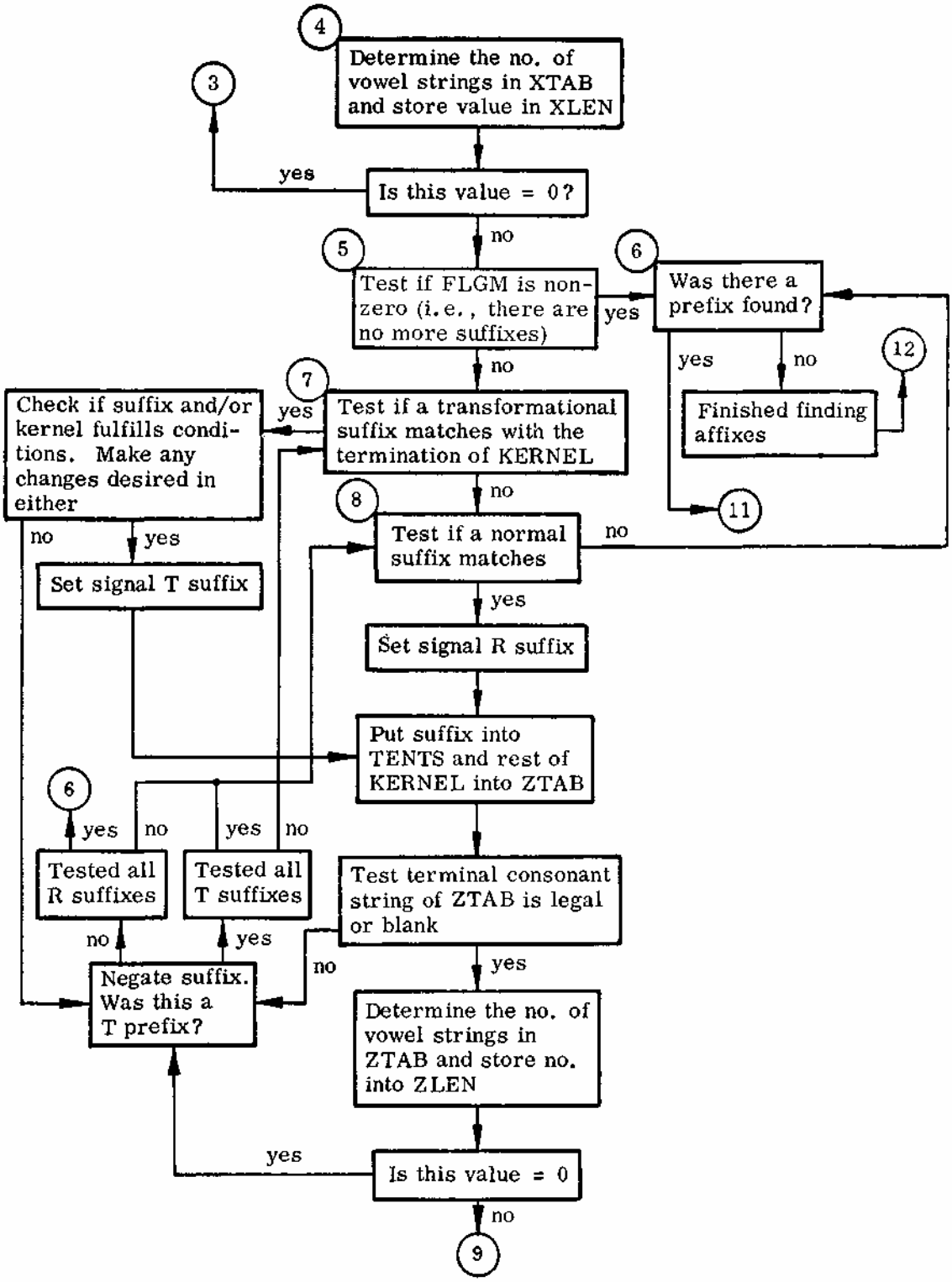

FIG. 1.—*Continued*

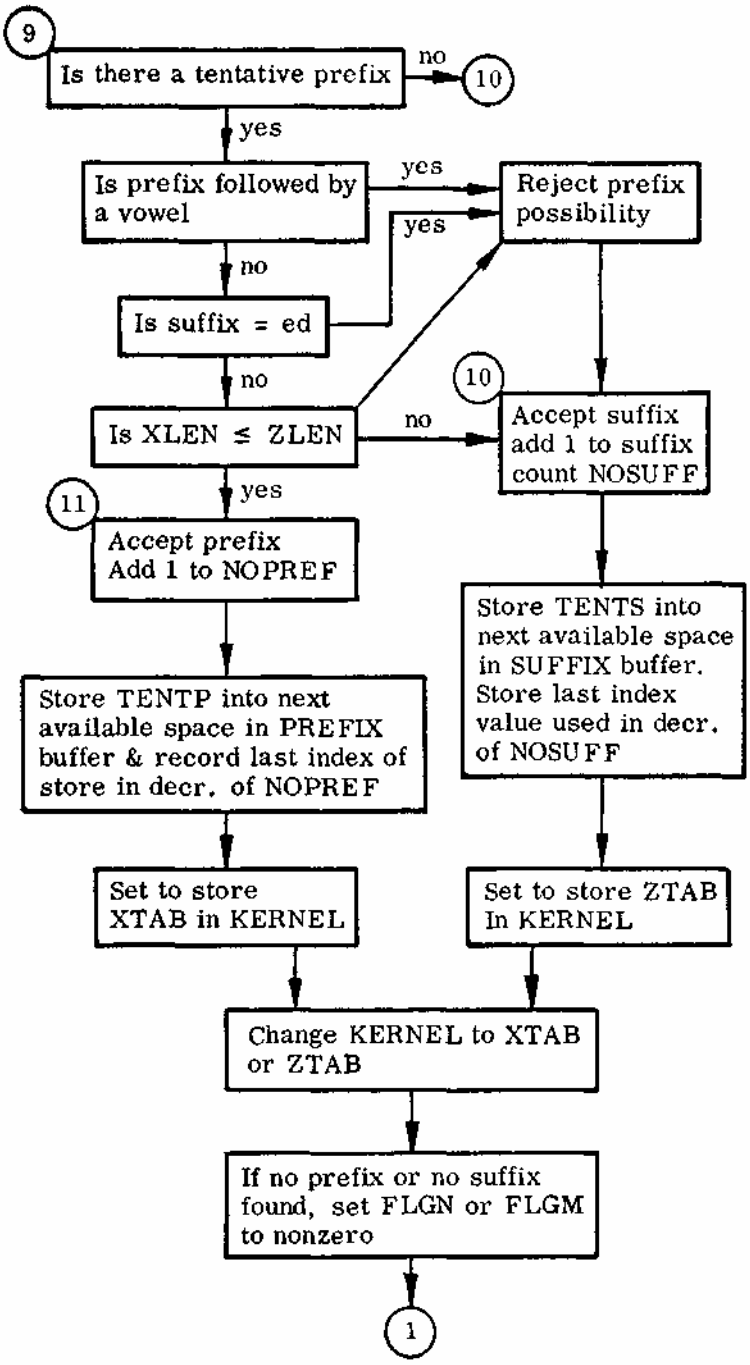

FIG. 1.—*Continued*

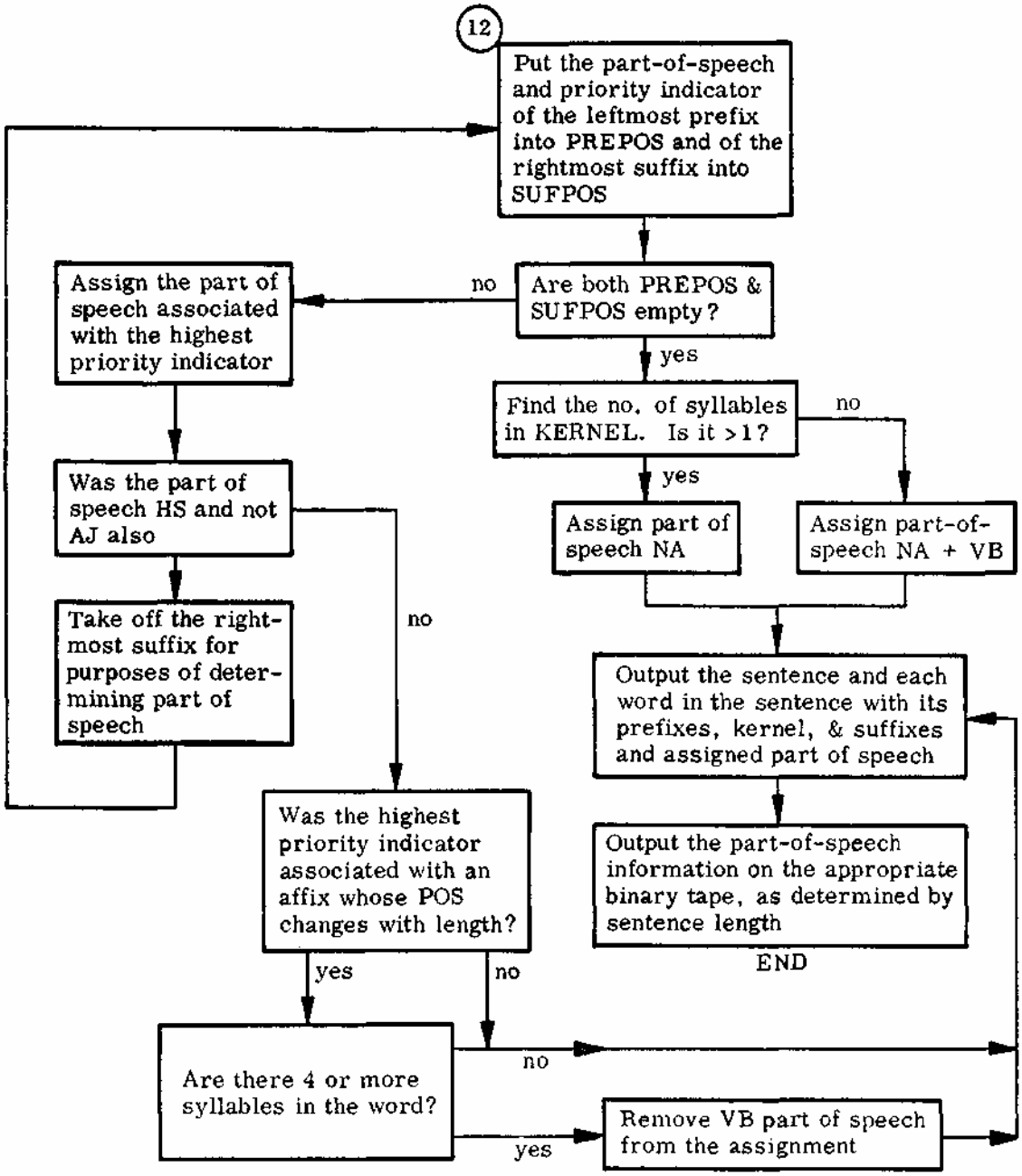

FIG. 1.—*Continued*

random sample, since we have just proved that the one-syllable words (there are forty-six in the sample) do not affect the results substantially. In its general form the second premise is accurate about 70 per cent of the time, as is reported in Reference 1. In its modified form, as stated in Rule A and tested by our five-hundred-word sample, it is accurate for only about 55-60 per cent of the cases, but is good for about 90-95 per cent of the cases from the point of view of inclusive part of speech, with something less than 5 per cent variation, depending on whether part of speech implications of affixes are used.

Summary

The net result of the part-of-speech studies is an algorithm which, used in conjunction with a dictionary of less than one thousand words and an affix list of less than two hundred, gives a correct "inclusive" part of speech for 95 per cent of a five-hundred-word random sample and which should do better on textual material. The dictionary is derived from an exhaustive compilation of words which the algorithm is not capable of handling. Such words are adverbs, function words, participles, or collective nouns not recognized by the program or, conversely, words so classified which should not be. The number of words in the exhaustive list is 3,163, of which less than one-third were selected for the dictionary. However, all of the function words with parts of speech other than NA, AJ, VB, or AV have been included, as have all of the irregular past verbs and past participles and the more commonly used adverbs and collective nouns. The omitted words are mainly less common adverbs and collective nouns, and they comprise only about 3 per cent of the total 73,582-word dictionary.

Received March 6, 1967

References

1. Dolby, J., and Resnikoff, H. "On the Structure of Written English Words," *Language,* Vol. 40 (April-June, 1964), p. 2.
2. *The Shorter Oxford English Dictionary on Historical Principles* (3d ed., revised with addenda; Oxford: Clarendon Press, 1959).
3. *Webster's Third New International Dictionary of the English Language* (Springfield, Mass.: G. C. Merriam Co., 1961).
4. Resnikoff, H., and Dolby, J. "The Nature of Affixing in Written English," *Mechanical Translation,* Vol. 8 (June and October, 1965), pp. 84-89.
5. Earl, L. L. "Structural Definition of Affixes in Multisyllable Words," *Mechanical Translation,* Vol. 9 (June, 1966), pp. 34-43.
6. Fowler, H. W. A *Dictionary of Modern English Usage,* rev. and ed. Sir Ernest Gowers (2d ed.; New York: Oxford University Press, 1965).
7. Earl, L. L. "Part-of-Speech Implications of Affixes," *Mechanical Translation,* Vol. 9 (June, 1966), pp. 38-43.