A History of the English Language

A History of the English Language

A History of the English Language

Garland Cannon

HARCOURT BRACE JOVANOVICH, INC.

New York Chicago San Francisco Atlanta Dallas

GARLAND CANNON is Professor of English and Linguistics at Texas A&M University. He has been on the graduate faculty of Columbia, the City University of New York, and the University of Michigan. His major books are *Oriental Jones* (London, 1964) and a two-volume edition of *The Letters of Sir William Jones* (Oxford, England, 1970), the latter having been selected by the London *Sunday Telegraph* as a Book of the Year. He has twice been to England on a grant from the American Philosophical Society.

PRINTED IN THE UNITED STATES OF AMERICA

ISBN 0-15-312356-7

ACKNOWLEDGMENTS: For permission to reprint copyrighted material, grateful acknowledgment is made to the following:

HARROW OBSERVER AND GAZETTE: Excerpt from the *Harrow Observer and Gazette,* Harrow, England, October 15, 1964.
NATIONAL COUNCIL OF TEACHERS OF ENGLISH and RAVEN I. McDAVID: Excerpt from "A Checklist of Significant Features for Discriminating Social Dialects" by Raven I. McDavid, from *Dimensions of Dialect,* copyright © 1967 by the National Council of Teachers of English.
PRENTICE-HALL, INC., ENGLEWOOD CLIFFS, N.J.: From *The Development of Modern English,* 2nd ed., by Stuart Robertson and Frederic G. Cassidy, © 1954.
SATURDAY REVIEW, MARIO PEI and DOUG ANDERSON: "How Did Language Begin?" by Mario Pei, illustrated by Doug Anderson, from *Saturday Review,* September 9, 1967, copyright 1967 Saturday Review, Inc.
THE UNIVERSITY OF MICHIGAN PRESS: The Map "The Speech Areas of the Eastern States" from *A Word Geography of the Eastern United States* by Hans Kurath, copyright 1949 by the University of Michigan.

TO JENNIFER AND WILLIAM

Contents

Chapter One *Early Man and Language* 1

man's acquisition of language
early man
writing systems
definition of language

Chapter Two *Comparative Linguistics: Dialects, Reconstruction, and Families* 8

divergence into dialects
causes of dialectal divergence
language families
the reconstruction process
other language families
divisions of linguistics
language and basic concepts
word order
tonal and nontonal languages
sounds and meaning

Chapter Three *Indo-European and Its Branches* 24

Indo-European migration
branches of Indo-European
the centum branches
the satem branches
the Germanic languages
Grimm's law
Verner's law

Chapter Four *The Generative-Transformational Process* 39

universal and particular grammar
deep-structure rules
insertion of words
transformational rules
summary

Chapter Five *Old English (449–1100)* 48

beginnings of English
the Anglo-Saxon heptarchy
Celtic loanwords
Latin loanwords
Scandinavian loanwords
Old English literature
Old English phonology
an Old English text
analysis of an Old English deep structure
lexical rules
transformational rules
Old English nouns
vocalic change
dental addition
Old English phonological rules

Chapter Six *Middle English (1100–1500)* 93

Norman invasion and history to 1250
loanwords: Latin, French, and Scandinavian
the Ormulum
history: 1250–1350, 1350–1500
four major dialects
Chaucer's Prologue
Wycliffe Bible
deep structure
modals and auxiliaries

lexicon
transformations
verbs
other transformations
phonology

Chapter Seven *Modern English (1500–1900)* 125

Early Modern (1500–1660): history, education, books
grammars, translations, language-teaching manuals, spellers, dictionaries
loanwords: Latin, French, Italian, Greek, Scandinavian, Celtic, Germanic, Indian, Arabic, Spanish, and others
structure of Early Modern English
deep structure
modals
progressive passive
lexicon
prepositions
transformations
nouns, determiners, personal pronouns, adjectives, verbs
"do" transformation
two other transformations
phonology
Great Vowel Shift
Authoritarian English, 1660–1800
appeal to authority
academies
grammars
Bishop Lowth, Dr. Johnson, Noah Webster
Late Modern, 1800–1900
Americanisms
American-British spelling

Chapter Eight *Present-Day English (after 1900)* 171

linguistic nationalism
British-American lexical differences
a British sports story
lexical differences
prepositional differences
transformational differences
phonological differences
dialects
American regional dialects: Northern, Midland, Southern
American social dialects
Negro speech
nonstandard speech
bilingual speech
necessary lexical variations
narrowing of dialectal differences
generalized Present-Day English
deep structure, lexicon, transformations, phonology

Chapter Nine *Language Attitudes* 201

twelve generalizations

Appendix

A. Phonological Symbols 223
B. Deep-Structure and Transformational Rules of Present-Day English 224
C. Glossary 228
D. "How Did Language Begin?" by Mario Pei 237

Figures

I. *Cardinal Numbers in Eight Languages* 14
II. *Derivational Chart of Selected Indo-European Languages* 32
III. *Indo-European Languages in Europe and Asia Minor Today* 37
IV. *Flow-Chart of the Generative-Transformational Process* 45
V. *Anglo-Saxon Heptarchy of the Seventh Century* 50
VI. *The Speech Areas of the Eastern United States* 180

Texts

A. *Late West Saxon Translation of St. Mark, iv. 1–10* 63
Literal Translation 64
Dictionary 64

B. *From the Preface to the* Ormulum *with a Translation* 100

C. *From the Prologue to the* Canterbury Tales 104

D. *Passage from the Bible in Old English* 111

E. *Passage from the Wycliffe Bible* 111

F. *Passage from the King James Version* 140

G. *A Sports Story from the* Harrow Observer and Gazette 173

Chapter One

Early Man and Language

How did man acquire language? Some religions contain metaphorical or supernatural explanations. According to the Old Testament, Adam was created and given language by God, who formed other living creatures and brought them to the Garden of Eden for naming by the first man. Adam's descendants were scattered by God into tribes speaking mutually unintelligible dialects (which we can call languages) because they presumptuously began to construct the Tower of Babel. Thus the original condition of one language and one speech for the whole earth was changed. The Koran contains a story much like the Judeo-Christian one, in which God created the first man and taught him the names of things. Hindu legend also gives language a heavenly origin. Brahma presented it as a gift to the gods, to the demons, and finally to men. The "laws" of language were then laid down by the ancient sages, with Sanskrit as the most perfect form. Other religions and cultures have other stories, often beautifully told.

Man's Acquisition of Language

There are some pseudo-linguistic theories about how man first acquired speech. These are really only idle speculation, with amusing names. According to the bow-wow theory, man imitated the sounds in nature and thereby began to talk. The wolf's howl sounded like "owoo-owoo," so that creature was called a *wolf.* Boys who make similar sounds upon observing a pretty girl are sometimes called by that name today. The difficulty is that the

theorist was probably an English speaker. He hears a dog's bark, for example, as "bow-wow," as opposed to the Spaniard's hearing of "gnau" or the Chinese's "wong-wong."

The ouch-ouch (or pooh-pooh) theory suggests that man spoke his first words as expressions of surprise or pain and the like. Perhaps some early man cried out "Ouch!" upon banging himself with his stone hammer, or "Pooh" when a burly friend made signs to indicate that he had just killed a thousand tigers and should be made king. Again, pain and emotions are expressed in different ways, according to the language. What sounds does a Parisian make when he spies a beautiful *fille* walking along the Champs-Élysées?

These and other speculative explanations don't give our ancestors much dignity or reveal anything about the acquisition of speech. The reluctant conclusion is that scholars will probably never learn exactly how language began because the event was too long ago. A communications corporation neatly resolved the dilemma at its New York World's Fair exhibit in 1964–65 by showing man needing speech in one diorama. Then in the next he has it as one of his most essential possessions. *Homo sapiens* needed language; ergo, it came into being.

Did language begin with a single man who thus became the progenitor of the three to six thousand tongues today? Or were several languages developed simultaneously? Either theory would be highly speculative. A few scholars maintain that several hominoid races evolved independently in varied locales and later came together in a single species. If so, the laws of probability would suggest a greater likelihood of several "first" languages. However, such a parallel origin of races contradicts the usual biological processes of development and evolution of one species from a single area. Most scholars think that man arose in a single area and then migrated around the world to become ultimately differentiated into the three or four basic races, which have become increasingly blurred over the millennia.

The only defensible answer, in short, is that the identity of man's common ancestral stock, much less of any first language, is simply unknowable at present. Many linguists doubt that there was some intermediate stage, when man didn't quite have speech but only a simpler system of sounds which later evolved into human speech. Of course, we can't prove that some primitive forerunner of today's complex system did not once exist; nor can

we prove that man didn't have language at all and then had it complete in the next moment. As for the two or more men who first used language, even if we could miraculously observe that important event, we can't be certain that we would understand the process of acquisition. The answer might lie hidden in their minds and sensibilities. These areas still cannot be directly observed, even in the twentieth century, for we lack complete physiological knowledge of the human brain.

How long has man had language? Fossil discoveries in South African rocks show that life existed on earth 3.4 billion years ago. Yet man is the only species that talks. He is *Homo loquens*. Archeologists have not been able to designate the exact time when the human species emerged from its nonhuman ancestors. Pitifully few artifacts are available to archeologists when they attempt to reconstruct periods prior to extant writing. (The oldest written records we possess are less than six thousand years old.) Certainly man was speaking before he engaged in writing, which is a way of recording his speech in the more permanent form of visible marks.

In recent years Louis S. B. Leakey has reconstructed the *Kenyapithecus,* who lived about fourteen million years ago in central Africa, and the *Homo habilis,* a primitive ancestor of the human species who lived almost two million years ago in the same area. We know little about these early ancestors, but we suspect that neither had speech.

Early Man

The latest anthropological research suggests that man is no more than a million years old and may be only seven hundred fifty thousand. His language presumably can't be any older because if, as linguists argue, only man has language, then protoman couldn't have it, by definition. One of the early members of *Homo sapiens* was *Sinanthropus,* or Peking man, who lived about half a million years ago. We don't know whether he had language. Reconstruction of his skull shows a brain capacity of about 1,075 cc., as compared to modern man's brain size of approximately 1,400 cc. One large cave apparently served as the home of successive groups of *Sinanthropi* for many centuries, if not millennia. Thousands of quartz fragments of flaked choppers, scrapers, and hammering implements have been found there. Peking man used

fire to cook a variety of wildlife, and other evidence suggests a kind of social life. The careful way in which the skulls found in the cave were broken may lead to the gruesome speculation that he valued brain power, perhaps to the extent of eating his enemy's brain supposedly to gain the victim's intellectual powers. Modern cannibals, contrary to popular conceptions, do possess language. It is apparently as adequate for communication as anybody else's.

When we come to Neanderthal man, there is a tiny bit of indirect information that he may have had language. Having a brain capacity of about 1,450 cc., he made lovely tools about one hundred thousand years ago in what is today the Le Moustier area of France and elsewhere. To create the stone knives and mineral pigments that were part of his culture, he presumably needed a method of communication more precise than frowns or smiles, and more subtle than a threatening ax or proffered saber-toothed tiger robe. As he also ceremoniously buried his dead, we suppose that he had rituals requiring speech. He lived until about fifty thousand years ago. Occasionally today we see Europeans whose strong brow ridges and sunken eye sockets suggest a distant relationship to Neanderthal man. The ugly (and probably untrue) rumor that he was bullnecked and walked with a stoop may have led to our dissociation from this intelligent being whose remains have been found in virtually every place where human fossils have been excavated.

We may not need to advance further than Cro-Magnon man in speculating about an early talking man. Cro-Magnon man had a brain capacity of up to 1,660 cc. and left delicate cave paintings near Aurignac, France, about forty thousand years ago. It's true that little is known about the relation of brain capacity to speech, but his artistic bone implements and sculptures constitute further indirect evidence that he normally communicated by means other than grunts and menacing gestures.

Anthropologists continue to shed light on early man, just as linguists continue to investigate the nature and properties of language. Not inconceivably, studies which are only now beginning on the way a baby learns and uses language may provide some indirect indication of the way man may have originally acquired speech. Biologists have discovered that the higher animals go through a series of changes from the embryo to their final form. At each stage the form resembles the adult stage of some lower animal. Certainly children are speaking by the age of two years

and are in command of the grammatical system of their native language by the time they are six. They've had no formal instruction in the matter, and parents' "baby talk" is often what they've mainly heard. So if enough children are closely studied during the critical early years, scholars may be able to construct a kind of acquisition model. This might shed some indirect light on how man began to speak, and should tell us more about how babies use certain innate capacities to master the complex human achievement of language.

Writing Systems

At the moment all serious linguistic evidence about the past must derive from history, or preserved writing. The earliest inscription discovered to date is on clay tablets, inscribed in the Sumerian language in cuneiform (wedge-shaped characters) about 3500 B.C. These were discovered in the ruins of Nineveh, capital of the Assyrian empire, a civilization that later developed in the same area where Sumerian had been spoken. Scholars' translations have provided some indication of the speech represented by the wedges. The characters apparently originated from pictures. For example, the symbol for star is something like an asterisk. Some characters also stood for the sounds associated with the object pictured.

Cuneiform would seem to have influenced another early set of written records, the Egyptian hieroglyphics of the late fourth millennium B.C. As many of the hieroglyphs pictorialize objects, the system is often called ideographic. It was not deciphered until one of Napoleon's soldiers discovered the famous Rosetta stone, containing two languages, one of which was Greek. The Greek translation of the text, which commemorated Ptolemy V's accession to the Egyptian throne in 197 or 196 B.C., permitted decipherment of this advanced writing system.

The first people to use writing to express the single sounds of a language were the Greeks, who developed a system of vowel representation. Our Roman alphabet came from the Greeks via the Romans. It is alphabetic; that is, each sound is supposedly represented by a particular letter. No letter pictorializes an object or an event, as was the case with the Sumerian and Egyptian writing. Still, just a moment's thought about the way we spell a word like *pneumonia* or *thatch*, or the way impressionists may paint a

perfectly common object, will make us prefer a living informant to either an alphabetic or an ideographic text. Thus if linguists could speak directly to the Indian who recorded the treaty with William Penn on a wampum belt, they could study many samples of his speech. The result would be a faster and more accurate description of his language, as opposed to a description based on decipherment and analysis of its written representation on the belt. Linguists will have to be content with written records, however, since only Methuselah is said to have lived 969 years. Since he didn't have a tape recorder, we can't have an actual corpus of his speech as some fortunate scholar a thousand years from now will have of Present-Day English. At least we hope that the time capsules of 1939 and 1964 will survive and remain clearly marked in Flushing Meadows, Queens.

That scholar of the future will certainly be interested in our language. It is, after all, a partial index to human nature itself, as well as a means of transmitting information from the speaker to the hearer. If we try to project what English might be like a millennium from now (a less speculative matter than any attempt to pierce the mists of deep prehistory), we can be fairly confident that it will be different from today's speech. Every language changes over a period of time; every one also has had basic qualities in common for thousands of years. There is no reason to expect these qualities to change rapidly or significantly. They help constitute a technical definition of the term *language* itself, which will be needed as we proceed through the history of our own tongue in succeeding chapters.

Definition of Language

Language is oral, as opposed to silent, semaphoric flags or the signs of mutes. It is both symbolic and systematic, so that random vehicular sounds on the freeway are excluded. It is arbitrary; that is, there is no necessary relationship between a given word and what it means. It is no more logical to call a certain four-legged animal a *dog*, than to call it a *dogge* or *dogue* or even a *cat*. Language is also recurrent. The animal is not called a *dog* today, a *cat* tomorrow, and a *mongoose* next week. Language is social and purely human. It is capable of transmitting information, misinformation, or even nonsense. It is adequate: man can make a linguistic response to any experience, if only to say, "I don't un-

derstand it at all." Overall, language is a noninstinctive method of communication, a particularly important one for *Homo sapiens*, although we all know that a silent smile may accomplish what a curt oral command will not. All these qualities are essential, as we will see.

Activities

1. Turn to the Appendix and read Mario Pei's "How Did Language Begin?" (*Saturday Review*, September 9, 1967, pp. 54–55). Is Pei serious or amusing? How is a dog's bark represented in some other languages which you may know? A cricket's chirp? After considering Pei's examples, try to defend the bow-wow theory.
2. Read Roy Chapman Andrews's *Meet Your Ancestors* (New York: New American Library, 1945). What other early men might have had language? See Ashley Montagu's *Man: His First Million Years* (New York: Viking Press, 1957) for other fascinating information about our forebears.

Chapter Two

Comparative Linguistics: Dialects, Reconstruction, and Families

Almost from babyhood, native speakers of English (or of any other tongue) differ considerably from one to another in the use of the overall language. Because we have no particular difficulty in understanding them, the differences between any two speakers are evidently superficial. The major syntactic rules (what we'll call deep structure) are the same for all native speakers of the language. The variation comes from the minor rules (surface structure, which will be described shortly).

We shouldn't be surprised that each individual's speech is somewhat different from that of every other English speaker. The human organism is both complex and unique, and each person's language and personality are his very own. The minute differences in the rules governing each person's surface structure help constitute his **idiolect.** The native speaker often isn't conscious of the differences, the individualization that distinguishes his speech from that of all other native speakers of the language. In fact, his idiolect contributes to his personality, his uniqueness. Many features of his speech may be shared by others, who can then be said to have the same dialect. Regardless of the geographical remoteness of any dialect from ours, we can still understand it. Although our ears pick up the particular variations, we should be tolerant enough to avoid trying to correct someone else's speech to make it conform to our own. After all, another's dialect and idiolect are adequate for his particular linguistic purpose.

Divergence into Dialects

Every language has at least one dialect. The dialectal variations can be conveniently grouped into three broad kinds, or components: syntactic, phonological, and semantic. The British say "in hospital" instead of "in the hospital," and "the government are" instead of "the government is." Americans immediately notice these syntactic differences, just as Englishmen note the American structures. Or consider the phonological difference between *burnt* and *burned.* Americans use terminal /d/ for the past tense of verb bases like *burn* and *dream;* the British use /t/. Of course, the past form of *kill* does not have /t/ in either dialect. In an oversimplified sense, the English spoken in the United States and much of Canada can be described as a collection of dialects loosely termed American English. The language in Kent, Cornwall, Yorkshire, and so on can be called British English. At least most Americans can quickly recognize an Englishman over the telephone, and vice versa. However, we shouldn't conclude that language is tightly and internally similar, either within the United States or within the British Isles. To prove the point, one need only ask for a spider in an Atlanta variety store, a frying pan in Boston, or possibly a skillet in New York. These are differences in vocabulary, in the semantic component.

Today there are from three to six thousand languages, considerably more than existed millennia ago. They presumably derived from earlier ones. Through a scholarly process called reconstruction, where known existing forms are used to reconstruct earlier and unattested forms perhaps of the parent language, each language can invariably be traced back to earlier stages. Reconstruction has definitively proved that Spanish and French have derived from Latin, so that general sources like encyclopedias record the fact without qualification. History explains why Spanish, rather than French or Portuguese, is the language of Puerto Ricans, Cubans, and Mexicans.

Fortunately for modern scholarship, early Roman citizens were quite literate. They left posterity great numbers of manuscripts, unlike our generally nonliterate Germanic forebears. Covering about ten centuries, the Latin manuscripts reveal several dialects. Study of them as they changed over a long period of time reveals that some dialects became mutually unintelligible. Geographical isolation from other dialects played a signifi-

cant role. Of course, there were no television and radio on which a Roman could hear a Castilian or a Parisian. After a while, even the Romans were not speaking Latin. It changed into Italian, in the same way that the two related dialects became Spanish and French.

Extant manuscripts show the step-by-step process of change. Why was there change in the first place? The manuscripts can provide no direct answer. It's not especially illuminating to know that the French and the Spanish cannot really understand each other, but that their distant forebears could communicate because they spoke diverging dialects of Latin. Linguists have discovered the *what* of linguistic change, if not the *how,* from written French and Spanish texts. Linguists have also demonstrated changes like the development of *also* from Old English *eallswā.* But it would be more exciting to know the specific *why* of such change. Dialectologists, who systematically describe the dialects of one or more languages, are comprehensively investigating the varied English dialects in North America and England. Although many of their findings are still not published, analysis of the particular divergences has suggested the general causes of the overall change. Raven I. McDavid, Jr., the director of the Linguistic Atlas of the United States and Canada, lists a combination of seven forces. They can be partly verified by older records of human settlement, migrations, trade, education, and the like.[1]

Causes of Dialectal Divergence

First, an area's early population may include a large or influential group of people, whose speech may be emulated. Thus features of the speech of Ulster Scots in the English of Western Pennsylvania can be explained by the Ulsterites' early settlement there and elsewhere along the frontier.

Second, migration routes, like valleys and rivers, may later form dialect boundaries. Thus we shouldn't be surprised to find boundaries along the Upper Ohio Valley and the Virginia Piedmont. Old political and ecclesiastical lines, like the long-standing diocesan boundaries in England, can encourage dialectal variation. Physical conditions like a well-watered plain or valley help

[1]The following summary of the forces is from McDavid, in W. Nelson Francis's *The Structure of American English* (New York: Ronald Press, 1958), pp. 483–85.

determine migration routes and, later, dialect limits. Adjacent Persian areas in Afghanistan, for example, have seldom been in communication. Only about a mile apart by air, they are separated by the twenty-six-thousand-foot Hindu Kush.

A cultural center, a fifth force, can affect adjoining, less prestigious communities. The "King's English," or "Received Pronunciation," is used by upper-class Londoners. It has influenced the speech of many other Englishmen, whereas cockney, low in prestige, is not emulated. Florentine and Parisian speech have significantly affected Italian and French. People usually try to emulate the actions and speech of the leaders of their society, whose speech is the prestige, or standard, dialect.

The social structure of an area may also help shape a dialect. All social levels of the American South and South Midland have /y/ after the initial consonant in *due, news,* and *Tuesday.* In Pennsylvania some social levels have this usage. On other levels, *do* and *due* are homonyms, words that sound alike but have different meanings. Finally, a large body of new immigrants with a different linguistic-cultural background can introduce new words, pronunciations, and even structures. The Spanish were hardly immigrants into what is today the United States. Yet a Texan's casual use of words like *remuda, pesos, sombrero,* and *siesta* is partly explained by the juxtaposition of the two languages along the Rio Grande. Many New Yorkers, on the other hand, are unfamiliar with the word *remuda.* Surprisingly, a Texan places stress on the first syllable of *rodeo,* whereas the Californian stresses the middle syllable, as the Spanish do.

These seven forces and others intertwine to help explain why the people of Kent have a different dialect from those in the English Midlands, or why New York speech varies from that of South Carolina. It is true that the forces don't provide much specific information about a given language. We still don't know precisely why a language develops into one or more dialects that can be understood across boundaries for a time and that later become separate languages. A population's speech just seems to change as the people encounter new topography, flora and fauna, and changing times in general. New words are required so that new phenomena can be spoken about; the American colonists learned this when they first saw a skunk. The alteration is so slow that an individual sometimes never recognizes the differences in his own speech over the years.

Language Families

It was said earlier that French and Spanish speakers can't *really* understand each other. The qualification was deliberate. French and Spanish—along with Italian, Portuguese, and Rumanian—belong to the Romance language group. They have all developed out of Latin. French and Spanish have, therefore, retained enough originally Latin structure and vocabulary in common to permit faltering conversation. This is not the case with French and German. The Germanic languages are not descendants of Latin, but share an earlier common ancestor with the Romance languages. These two groups diverged from a common parent language called Indo-European so long ago that they are now mutually unintelligible. Still, it is generally easier for a German speaker to learn French than to learn Thai, which was never a member of the same language family.

It seems reasonable to speak of families. The Indo-European family includes those languages that developed from the protolanguage (a language from which another language or whole group is derived) that, according to most scholars, existed in Europe about 3000 B.C. As different dialects changed into separate languages out of which new dialects slowly emerged over the centuries, branches, or subfamilies, like Hellenic and Italic, grew out of the parent family.

By means of the family concept, the century-by-century development of particular tongues from others can be shown, and their individual relationships can be explained. Thus Greek and Latin are sisters. At one time, their predecessors were two mutually intelligible dialects spoken more or less side by side in a given broad area. The dialects were influenced by successive changes, much the way a stone thrown into a lake spreads waves across the surface. Each innovation introduced a new division. Eventually the two dialects were separated into related tongues, each with its own development. Similarly, linguists can quickly point out the direct ties between the nine Chinese languages, as well as their more distant relationship with Thai. They can do so because the Chinese languages were differentiated from the protolanguage of Sino-Tibetan over a long period.

A clear linguistic genealogy is comforting, somewhat like an attested family tree for people. Documented records prove Henry VIII to have been the son of Henry VII and the father

of Edward VI. This descent established their right to the throne, and history confirms their rule in that order. Obviously, language isn't a biological organism, where there are neatly recorded birth certificates to identify the father and mother. Language can't *do* anything; it is done to. It is changed as its successive speakers unconsciously change their speech over the centuries. Yet reconstruction of some protolanguages, including many from Indo-European, has provided sufficient documentation about the historical order of their derivative dialects so as to be equivalent to birth certificates.

The Reconstruction Process

What is the nature of such work? How can tongues that no longer exist be adequately reconstructed from faded manuscripts and weathered inscriptions to permit classification into families? The comparative method is used; that is, two or more things are observed and systematically catalogued. The more writings the linguist compares and the closer his inspection, the more dependable his tentative conclusions become. His corpora should be representative. Comprehensive analysis then verifies, alters, or explodes the experimental hypotheses.

The similarity of Sanskrit to Greek, Latin, Persian, and Germanic was observed by Sir William Jones and others in the late eighteenth century. When they assembled the numerals from one to ten in these languages, they began a study that still isn't quite completed. Figure I, which includes some examples transliterated into the Roman alphabet from other writing systems, shows the kind of data that these scholars compared. Just a glance will disclose so many similarities in sounds and their sequence that mere chance cannot explain them. Actually, all numerals on the list except the Thai ones belong to the same family, as developed from an originally common source. Once the variation between /t/ and /d/ in *two* and *dvā,* and *ten* and *dix* is explained, the words are seen to be much alike. The same is true for /f/ and /p/ in *fünf, pente,* and the other cognates of *five.* Cognates, incidentally, are words similar in sounds and structure, derived from a common ancestral language. A closer look reveals additional commonality, with the Thai forms always different, and further confirms the considerable nonrepresentativeness of English spell-

ing. In English *three*, for example, the cluster /θr/ is spelled with three letters. Actually, the cluster contains only the two consonants represented by the symbols within the slant lines.

Figure I:
Cardinal Numerals
In Eight Languages

English	German	French	Greek	Old Church Slavic	Persian	Sanskrit	Thai
one	ein	une	heis	jedinu	yak	eka	nỳng
two	zwei	deux	duo	duva	do	dvā	sŏng
three	drei	trois	treis	trije	se	trayas	săm
four	vier	quatre	tettares	chetyre	char	catvāras	sì
five	fünf	cinq	pente	pęti	panj	pañca	hâ
six	sechs	six	heks	shesti	shash	sat	hòg
seven	sieben	sept	hepta	sedmi	haft	sapta	chèd
eight	acht	huit	oktō	osmi	hasht	astā	pèd
nine	neun	neuf	ennea	devęti	no	nava	kâw
ten	zehn	dix	deka	desęti	da	dasha	sìb

Undeniably, the languages in Figure I might have borrowed the system of cardinal numbers from one to ten from some other tongue and actually be unrelated. There could have been a massive military conquest in which the conqueror's language replaced that of the natives. Chamorro, spoken on Guam and the other Marianas, was overwhelmed just that way by Spanish. However, Sanskrit and Old Persian are geographically and chronologically remote from English. It's hard to believe that a single language could ever have been so influential that its number system was borrowed by diverse tongues millennia and tens of thousands of miles apart. The rest of the cardinals bear out the impression given by the first ten. What are the cognates for *eleven* and *twelve* in the languages in Figure I? Cardinals in a Celtic tongue like Irish Gaelic or Welsh might profitably be added.

Besides studying base forms of words and their sounds, linguists investigate syntax and inflections. Inflections are modifications in the forms of words, principally by means of endings, which indicate grammatical relationships like person and number. What similarities must linguists have found in the following forms?

	Fr. **vendre**	*It.* **vendere**	*Port.* **vender**	*Sp.* **vender**
singular	1 vends	vendo	vendo	vendo
	2 vends	vendi	vendes	vendes
	3 vend	vende	vende	vende
plural	1 vendons	vendiamo	vendemos	vendemos
	2 vendez	vendete	vendeis	vendéis
	3 vendent	vendeno	vendem	venden

The addition of similar suffixes to the base *vend-* in each of the six forms of the four Romance languages is so phonologically complex and internally systematic that the suffixes could hardly have been borrowed from an unrelated tongue. Even these limited data indicate another fact, that Spanish and Portuguese are the "closest" of the four. Why? They apparently came from a common immediate predecessor, a tongue called North Ibero-Romance. The striking inflectional resemblance isn't confined just to the present time or to the above verbs. It is repeated thousands of times in these and the other Romance languages, including the declensions of nouns and adjectives. A declension is an ordered list of all the inflectional endings of a given word base. Comparable sets of declensions occur in each of the branches of Germanic, Slavic, and so on.

The *vend-* examples are all from "living" tongues. There is an English cognate, borrowed from the French *vendre*. What is it? Of course, the word *sell* is used more commonly today. Unlike English, the other Germanic languages didn't come into close contact with French. Yet there are cognates for *sell* in Gothic (the East Germanic language of the Goths, now extinct) *saljan,* Icelandic *selja,* Norwegian *selge,* and Swedish *sälja.* As the five words seem much alike, we can begin to see how linguists have reconstructed the Indo-European base **sel-*, meaning "to take or grasp."[2] Without contemporary written (and certainly not oral) evidence, linguists have used written cognates of later languages to reconstruct the hypothetical earlier forms of Indo-European. Consider Greek *helein,* which doesn't look similar to *sell.* Actually, Greek is proved to be of a different line than the Germanic

[2] The asterisk is used throughout this book to indicate a reconstructed form or a nonstandard expression.

languages, once extensive numbers of cognates, the phonology, and syntax are studied. The initial Greek /h/ (which hereafter will be written as /h-/), contrasting with Germanic /s-/, poses the problem, as was seen with /t/ and /d/, and /f/ and /p/.

In Chapter Three we'll discover that all three pairs of sounds are quite consistent. They provide additional evidence that certain languages are Indo-European and others aren't, that some are Germanic and some aren't. Once the change of /s-/ to /h-/ is explained, *helein* is found to be related to **sel-*. What about *vendere, vender,* and the Rumanian *a vinde,* which don't sound or even look like the Germanic words above? We hardly need to check an etymological dictionary before deciding that none of these Romance words came from **sel-*. They had a different source, the Latin *vendere,* which apparently didn't derive from Indo-European but from some other source.

Just because a language is derived from Indo-European, we can't presume that all its vocabulary has come from that source. If so, there couldn't be coinings (new word creations) like *psychedelic* and *zap,* or borrowings from other tongues, two standard ways of addition to the English and other lexicons. The words *Allah, catamaran, igloo, lei,* and *tycoon* were borrowed from five languages, none Indo-European. What are the languages? Lexical borrowing is pleasant because we don't have to repay the loan later. Indeed, the Arabs and the Hawaiians are proud that some of their words have been borrowed. What do the following mean: *aloha, hula, lanai, luau, poi,* and *wahine?* They are now so naturalized that we don't have to jet to Honolulu to know their meanings.

Other Language Families

As an American territory and later a state, Hawaii has contributed dozens of words to English, partly because of the tourist and military activities there. Unfortunately, the number of islanders who speak Hawaiian as their primary language is fast diminishing. One of the borrowings, *lavalava,* is both Samoan and Hawaiian. So is *poi. Taro,* used in the Islands because poi is made from its root, is also Tahitian and Maori. *Tapa* is also Tahitian. Once we collect many cognates and discover their close phonological and syntactic resemblances, we can begin to describe the Malayo-Polynesian family. It includes Malay in the

west, Chamorro toward the north, and Hawaiian to the east. Some other Asian and Pacific families are Sino-Tibetan and Papuan. In the former are the Tibetan and Burmese languages and the Chinese tongues.

In the Middle East and North Africa we find the Hamitic and Semitic branches of the Afro-Asiatic family. Semitic includes the dead language Sumerian, as well as Arabic and Hebrew. The language in which verses from the Koran are chanted, Arabic is spoken natively by many Islamic peoples in Africa. The southern branch of Afro-Asiatic is Chatic, which includes Hausa, numbering more than six million speakers in northern Nigeria and Cameroons. There appear to be three other families on the continent: Chari-Nile, Central Saharan, and Niger-Congo. Although classification of African languages into families and branches therein is continuing, scholars have documented some eight hundred tongues south of the Sahara spoken by a total population of about one hundred fifty million people. About fifty have more than half a million speakers each. The Niger-Congo family, containing Ibo and Yoruba, has about sixty-seven million speakers, by far the largest number of any African language family. Swahili, sometimes taught in black studies programs, is primarily a trade tongue belonging to the widespread Bantu branch of Niger-Congo. Zulu and Xhosa, in the Union of South Africa, belong to this branch. Unauthentic, Grade B movies filmed in Africa often contain would-be dialogue in such African languages. The dialogue would not be understood by Zulus, since it is the screenwriter's version of what he thinks Western movie audiences think Zulu sounds like.

Approximately one hundred language families, scattered around the world, have been determined. They include Altaic (illustrated by Turkish), Finno-Ugric (Estonian), Mayan, and Athapaskan Indian (Navajo, Apache, and Tlingit). A fond but probably unrealizable dream has been to link the families backward in time. In the past there have been pseudo-scholarly assertions that Allah's language was Arabic, Jehovah's was Hebrew, and Brahma's was Sanskrit. Serious effort was made to link the Semitic branch of Afro-Asiatic with Indo-European, the family with the largest number of speakers, who are now spread over all continents. The effort was unsuccessful, principally because any possible common source is so remote in time that the most speculative reconstruction to date won't satisfy the needed rigorous

verification. All that can be proved at present, until other language families are reconstructed more fully, is the existence of numerous large ones. Some families are confined to a single continent; others can be found on every continent. Scholars have learned the most about the Indo-European languages. Their findings have made this book possible.

Divisions of Linguistics

The examples in this chapter from various dialects, languages, and families derive primarily from comparative study. Scholars logically call such study **comparative linguistics.** There are two other divisions of linguistics—**historical** (which must underpin any history of a single language) and **descriptive** (the analysis of a particular tongue at a particular time). Thus when we look at Old English shortly, the analysis will be descriptive. The term *synchronic* might be employed for linguistic study that doesn't use chronological evidence. **Dialectology** doesn't fit neatly into any of the divisions. Dialectologists often prepare an atlas, a publication consisting usually of a set of maps recording dialectal variations in syntax, lexicon, and phonology for a given area. Their work is usually more descriptive than anything else; but it can also be synchronic or historical or both. For this book, we need considerable dialectal information about British and American speech.

Linguistics is sometimes divided another way, into pure and applied work, like pure mathematics or applied chemistry. We're not concerned here with linguistic applications to the teaching of foreign languages. Those results can easily be observed in the tapes and methodology in the language laboratory. Nor is our concern with improved reading techniques or bilingual matters, as important as they are. Rather, our interest is in descriptive and historical information about English, which is valuable in its own right.

Before turning to a more detailed study of Indo-European and its branches, in final preparation for a survey of English as it has developed through the Germanic line to the present, we need to consider four more topics from comparative linguistics. The examples may help us understand why school friends whose native tongue is not English sometimes see things a little differently and may use our language in seemingly curious ways.

Language and Basic Concepts

The first topic is really a principle: different language groups reflect different ways of thinking. Man's view of basic concepts like person and number, color, time, space, and matter is partly conditioned by his language, and vice versa. In English the personal pronoun that goes with *has* is *he, she,* or *it.* Information about the sex is thus included in the choice of pronoun. We awkwardly inquire about a new baby by asking, "What is it?" The proud parents may respond, "It's a *he.*" The neighbor's dog must be referred to as *he* or *she,* depending upon the sex, not as *it* to the doting owner. *They* doesn't tell the gender but requires a plural verb. On the other hand, Persian has only *u* in the third-person singular, which can mean *he* or *she.* When the sex of animals is to be communicated, assuming that the particular word does not already indicate it (as in English *buck* or *doe* versus *deer*), the form *nar* (male) or *made* (female) is added before or after the word naming the animal. Navajo has no element for third person. Speakers simply indicate it by omitting the first-person and second-person elements. Thai has a particular first-person pronoun for the king. All Thais know it, but only the king would dare use it publicly.

Comparable differences can be found in conceptions of number. Some Malay dialects don't have a clearly defined system of number, which isn't a grammatical category for them. Samoan has singular, dual (for two people), and plural forms, like Old English. Concepts of color and time may also significantly differ from language to language. Because physicists have discovered the spectrum to be a continuum of light waves of varied lengths, we might falsely judge some Polynesian peoples to be both ignorant of physics and colorblind. They use the same word to describe the color of both the sky and the ocean, while we usually distinguish the two by *blue* and *green.* Actually, the divisions between colors are always arbitrary within a language. The names for colors in the world's languages preceded modern physics and so do not reflect modern scientific discoveries. Why does English, for example, have separate words for *orange* and *red,* rather than some item that we'll create as **ored?*

Some Americans expect people in other parts of the world to base their attitude toward time (like the necessity of being on time for an appointment, including not arriving early) upon a

stopwatch. Some American Indians make complex distinctions for past actions. Speakers of Wishram, a Chinook group in the Pacific Northwest mentioned by Lewis and Clark, have four separate past tenses. The speaker chooses the right one according to the remoteness of the action from the present. Navajo inflections represent shape, with the capacity to specify "roundness" and "visibility" in future time for a particular thing. American Indian languages have a rich variety of grammatical categories for time, space, and matter.

Word Order

The second topic contrasts languages with relatively strict word order, like English today, and those with numerous inflections and freer order, like Latin or Spanish. The Latin "Nolo patientiā tuā usque abuti" means "I don't wish continually to abuse your patience." A word-for-word translation is "I am unwilling patience your continually to abuse," a structure requiring much rearrangement to be understood in English today. Modern Spanish also has considerable leeway in word order. Either "La casa se ensució" or "Se ensució la casa" is acceptable. What about "Got dirty the house" in English?

A long time ago, Indo-European derived languages had a rather flexible syntactic order. Old English had a large number of inflections, as we'll see. The grammatical information communicated thereby permitted generally free syntactic order. "Sǣde hē heom" will illustrate for now. The *hē* signals third-person nominative singular, thus the subject of the sentence. *Heom* signals third-person dative plural. In modern terms, the dative can be described as the addition of the preposition *to* before the pronoun in the objective case. We couldn't make ourselves understood by saying "Said he them," because the loss of nominative and dative case has tightened the word order for modern speakers. What specific changes have occurred in this simple structure to give us "He said to them"?

Tonal and Nontonal Languages

The third topic contrasts the flexibility of English stress and pitch with the fixed pitch of tonal languages. "Batman zapped the

Joker" can have at least four different shades of meaning, according to location of the primary stress:

1. Batman did it (not Robin)
2. He was zapped (not patted)
3. The Joker was the victim (not Robin by mistake)
4. Ambiguous (when stress is put equally on all items).

Although English has four levels of pitch, variations therein can further alter the meaning. A rising pitch can convey a questioning of whether the zapping was done at all. Falling pitch indicates that the zapping was done as a routine action. A high pitch can convey excitement, as if the sentence is to be punctuated with an exclamation mark instead of a period or a question mark.

In Thai and the Chinese languages, such variation is not possible. A special syllable at the end of a sequence converts the whole thing to a question. In addition, each sequence of sounds in Thai has an intrinsic pitch, or tone. A different tone for the sequence makes it a different word, since there are five tones: ˇ (rising), ˆ (falling), ˊ (high), ˋ (low), and unmarked (middle). In Figure I we saw that *kâw* is "nine." *Kàw* means "to be old." With an aspirated consonant and a rising tone, the sequence *khǎw* means "he, she, they." It contrasts with *khâw,* "to enter." With a lengthened vowel, *khâ·w* is "rice," and *khǎ·w* "to be white." Thus a speaker careless of the tone and vowel length may say *"The boy is rice" instead of "The boy is nine."

The sequence *may* has three possible meanings: "to be new" (ˋ), "not" (ˆ), or an item indicating that the utterance just finished is a question (ˇ). Confusing two tones, an American housewife once told her amah, or female servant, to go get a snake (instead of a doctor) for her sick child. An American trying to learn Thai might call a mother a horse, or vice versa. Naturally Thais don't have this problem. They've learned the intrinsic tone of the particular sequence as a part of their native language, just as English speakers have learned never to say *"I went tomorrow."

Sounds and Meaning

Thais usually have a problem when speaking a nontonal tongue like English. "Two dollars" is likely to be singsong. Modeled upon "sŏng bàht" (the *baht* being their monetary unit), *two*

receives a rising tone, and the first syllable of *dollars* a low one. The singsong impression upon an English speaker illustrates the fourth topic, the occasional transfer of sounds across language barriers. No sound is meaningful in and of itself. Still, an intended meaning can be affected if a native sound is wrongly substituted in another language. Thai has no terminal /l/, so that a certain English borrowing is /hotɛn/. In asking for the *Hilton Hoten,* the Thai isn't maliciously misusing English. He is only substituting the Thai sound closest to the English terminal /l/.

We may have heard former servicemen talk about the passwords used on Japanese-held atolls in World War II. The words often contained medial *r,* /-r-/. If the person in the darkness beyond the outpost answered, "I'm a Maline," he was shot! Japanese does have /r/, as in *suru* (to do). However, the /l/ and /r/ don't contrast; they are two ways of pronouncing the same sound. The Japanese may use what sounds like an /r/ to English speakers in such words as *light* and thus be misunderstood as saying, "You right the fire." He may ask for a ticket to "Honoruru," though he will silently correct any rude English speaker who mimics him. That is, he hears the difference but does not make it in his own speech.

We can all think of similar examples from other languages. The Spanish confusion of the English pair *ch* (/c/) and *sh* (/š/) has sometimes dismayed American tourists in Mexico. They expected the *muchacho* to wat*ch* the car for one peso instead of wa*sh* it for fifteen. The Spanish substitution of /j/ for /y/ is equally familiar due to the now-canceled TV commercial about the "jellow" pages of the telephone book.

The principles involved in these four topics will so often apply to the changes in English described in succeeding chapters that we will summarize them. Different language groups reflect different ways of thinking. A tongue with comparatively numerous inflections is likely to have relatively free word order. Conversely, one with a fairly rigid word order usually has few inflections. English words, past and present, haven't carried fixed stress or pitch. An English word or sequence can be uttered with any one of four stresses and with any one of four pitches. The variations in stress and pitch significantly affect the meaning of the sequence. Finally, no single sound in a language is meaningful in and of itself, but can lead to miscommunication when wrongly substituted in another tongue.

Activities

1. Listen to a Britisher's speech. Then list the pronunciations, vocabulary, and word order that seem different from yours. If some members of your class have recently moved to your city, have them list some of the differences between their dialect and yours, while you note their differences. Compare the two lists. Do the differences in dialect ever impede understanding?
2. Why can language reconstruction at best tell us only the *what* and possibly the *how,* but not the *why* of linguistic change? If only one of the three questions could be answered, which answer would you like to know? Why?
3. Read the history and geography of your area. Was there a persuasive early element of population? Did an old migration route or political or ecclesiastical boundary cross your area? If so, how did they affect the dialects? Have immigrants into your area influenced the local dialect? Consider words for food, dress, and customs. What about cultural prestige?
4. Find at least five cognates in some of the Romance languages. Which words derive from Latin? Are there English cognates for any of them? If so, do the words *have* to be originally from Indo-European? The answer can be discovered by the history of the words *govern* and *preach* among many others.
5. Play a game in which various colored objects are shown to the class, and the color of each is written down by the students. Count the different color words that have been used. Do the boys or the girls have the larger number of color words? Which group has made finer distinctions between colors of almost the same wave lengths?

Chapter Three

Indo-European and Its Branches

Let's try to imagine the Indo-European language of about 3000 B.C. What was it like? Reconstruction shows considerable inflection. The noun bases had suffixes to provide information about the several cases, number, and gender. Verb bases were altered to inform the listener about person, tense, and mood. Word order was comparatively free, as we have already noted about languages with numerous suffixes. The lexicon contained some of the words from which the English names of biological functions have derived. *Breathe, die, eat, live, love, sleep,* and *walk* are examples. There were words for snow, winter, and common European flora and fauna, but none for Asian tropical flora and fauna. All in all, the language was adequate for the people's needs.

The location of the original homeland of the Indo-European speakers has been disputed by scholars. Nineteenth-century scholars believed that the homeland of these speakers was India, and gave them the name *Indo-European* rather than *Euro-Asian* or *Euro-Indian*. Twentieth-century linguists have generally held to an original European focal area. Recent archeological findings suggest that western Asia may have been the source of the migrations of the people we call Indo-European after all. Considering the difficulty of establishing the exact starting point in time and place of this highly mobile language group, let's hypothesize central Europe as their homeland for now, in order to help us get an idea of how the parent language evolved into various branches.

Indo-European Migration

About 3000 B.C., large numbers of speakers of Indo-European lived in central Europe, primarily from Lithuania to the Russian steppes. Some were already slowly moving in long, massive migrations toward northern Europe and Britain, the Mediterranean peninsulas, and Asia Minor. About 2000 B.C., these migrations led to the formation of new dialects that gained words through coining or borrowing from the languages with which the migrators came into contact. Sometimes a word replaced an Indo-European one; sometimes it filled a gap in the existing lexicon. Words were needed to name geographical features and flora and fauna not found in central Europe. When the startled migrators saw bamboo, banyans, crocodiles, and monkeys for the first time, they needed words for these things. If they didn't see rabbits and partridges for a time, their Indo-European words for these might not be used and might die. When they encountered the inhabitants of the new lands, they borrowed some of the inhabitants' vocabulary. They also "borrowed" the land and perhaps wives. Overall isolation from people they had left behind, both in geography and time, further assured eventual unintelligibility among certain Indo-European dialects.

Branches of Indo-European

Some of the groups of differentiating dialects, which we call branches, may well have died out without a trace, particularly when their speakers were nonliterate and left no inscriptions. Definite evidence remains, however, for at least eight generally unified branches of Indo-European. We can name the branches roughly geographically rather than chronologically, as long as we understand that they all didn't evolve at the same time. We find Celtic in western Europe, Germanic in the north and west, Italic on the Italian peninsula, Hellenic in Greece and the adjacent islands, Albanian, Balto-Slavic in the north and east, Armenian, and Indo-Iranian all the way from the Iranian Plateau to the Ganges. Some of the later migrators conquered their fellow speakers of Indo-European languages and took from them the lands previously wrested from the indigenous inhabitants.

Tocharian was probably a ninth branch; Hittite, a possible tenth. The two complicate an otherwise neat geographical pano-

rama. Hittite, the language of powerful biblical kings in Asia Minor beginning about 1600 B.C., may have been a development of a parallel line. If so, we might speak of Indo-Hittite. Their ancestors were already in Asia Minor before the mid-third millennium B.C., when their ancestors' cousins were still at home. Hittite cuneiform demonstrates a definite affinity with Indo-European. Tocharian, spoken in central Asia until its mysterious death about A.D. 1000, is preserved in Buddhist writings of about A.D. 600 found in Chinese Turkestan. Further research is needed to firmly establish Tocharian as an Indo-European derivative.

The problem posed by Hittite and Tocharian is their preservation of the Indo-European consonant /k/ before certain vowels. In other Indo-European languages in the east, /k/ became /s/, whereas in the west it remained unchanged. Until Tocharian was discovered, scholars had patly projected a **centum** protolanguage for the west (Latin "hundred," pronounced with /k-/, as in *Kent*), and a **satem** protolanguage for the east (Avestan "hundred," pronounced with /s-/). Avestan was an early member of the Indo-Iranian branch. The centum branches include Celtic, Germanic, Italic, and Hellenic. During World War I, scholars learned of Tocharian, a centum tongue greatly eastward of the known centum ones. Of course, the geographical "misplacement" of Tocharian might be explained by earlier migrations, but this may not be the answer.

The Centum Branches

Derivatives of Celtic were spoken in the area known today as Great Britain when the Germanic conquerors crossed the North Sea from the Continent. Indeed, long before the Germanic rise to influence, Celtic speakers occupied much of Europe. They reached as far east as Ankara and as far south as Spain and northern Italy. Celtic now seems to be the centum branch most likely to become extinct. Gaulish and Cornish have died, and Welsh and Munster Irish are seriously threatened by English. The last two possess rich literary works. On the whole, the Germanic languages slowly displaced the Celtic ones after 1000 B.C. Germanic will be discussed in some detail after the remaining branches are sketched.

Italic was much like Celtic for a time. Some scholars have even proposed a predecessor language, called Celto-Italic, out of

which the two emerged. Italic was carried down the Italian peninsula by massive migrations in the second millennium B.C. Oscan and Umbrian belonged to one dialectal line. Latin, which was to dominate and to exert vast literary and political influence, belonged to another line. Latin replaced the Etrurian inhabitants' language, Etruscan. Out of Latin came Spanish, which has the largest number of speakers of all the Italic-derived tongues. Portuguese, French, and Italian also descended from Latin. After Slavic intrusions cut off the peoples living in present-day Rumania from Rome, Rumanian, a Latin-derived tongue, gained many Slavic words.

Like Latin, Hellenic has exerted much influence on Europe, although the number of speakers of this branch has declined over the centuries. Hellenic was carried into Greece and the adjacent islands after 2000 B.C., replacing the non-Indo-European language spoken there. Mycenaean Greek tablets of about 1450 B.C. tell us much about early Greek. We know there were dialects like Aeolic, Doric, and Ionic-Attic. Modern Greek descended from the Attic dialect (of Athens) of the Age of Pericles. Greek philosophy, together with epics like the *Iliad* and the *Odyssey*, continue to influence Western man.

The Satem Branches

There probably wasn't a separate language derived early from Indo-European out of which developed the satem group of languages. The group is chiefly differentiated from the centum group by replacement of the palatal /k/ with /s/. The satem branches include Albanian, Balto-Slavic, Armenian, and Indo-Iranian. Albanian is spoken farther west than most of the satem tongues. Its early history is probably the least known of the eight branches, since the oldest Albanian record dates back only to the fourteenth century. Perhaps some "back" migration can explain its presence west of historically earlier Hellenic. Migrating groups of other Indo-European people seem to have crossed Greece en route to Asia Minor and other places that became the Asian satem areas. One of these might have later gone west instead of east.

To the north and east of Albania reside hundreds of millions of speakers of languages in the Balto-Slavic branch. Some scholars prefer to speak of separate Baltic and Slavic branches. Of the

Baltic descendants, the language of greatest interest to military historians is Old Prussian, which died about A.D. 1700. The Junker descendants (a military aristocracy) of Old Prussian speakers militarily influenced European events thereafter. The inhabitants of Latvia and Lithuania, forced to submit to Russia in 1940, still speak Baltic derivations despite pressures from Russian. Lithuanian has been especially studied. Its earliest record is a 1547 translation of a work by Martin Luther. In sounds and inflections, Lithuanian seems to have changed least of the Indo-European tongues. Over the centuries its speakers have remained in the same area. We might pause for a glance at five similar structures. The last three are clearly the most alike:

English: God has given teeth; God will give bread.
Latin: Deus dedit dentes; Deus dabit et panem.
Sanskrit: Devas adadāt datas; Devas dāt (*or* dadāt) api dhānās.
Lithuanian: Dievas dawe dantis; Dievas duos ir duonos.
Indo-European: *Deivos ededōt dn̥tn̥s; Deivos dedōt (*or* dōt) dhōnās.[3]

Lapland, Estonia, Finland, and Hungary lie between the areas where Baltic and Slavic languages are spoken. Their languages are descended from the Finno-Ugric family, which is unrelated to Indo-European. Of the Slavic tongues, Russian is probably the best known. A ninth-century biblical translation provides an early record of Old Church Slavic, still the official tongue of the Russian churches. Other Slavic languages are Bulgarian, Serbo-Croatian, Czech, and Polish. These continue to be influenced by the widely spoken Russian. It might be called Great Russian to distinguish it from White Russian (Byelorussian) and Little Russian (Ruthenian).

Armenian, another satem branch, can be traced back at least as early as a fifth-century biblical translation. Armenian predecessors conquered the inhabitants along the Black Sea about 750–500 B.C.; the Armenians themselves were later subjugated by the Persians and later again by the Turks. The language natu-

[3] From Paul Thieme, "The Indo-European Language," *Scientific American*, CXCIX (October 1958), 74. A dotted *n* means vocalization accompanying the consonant (Indo-European didn't have a sequence like **dntns*). There had to be some intervening vowels, but they haven't been reconstructed because linguists have no idea of what they were.

rally contains many Persian borrowings, just as English gained much French vocabulary during the Norman Conquest.

Indo-Iranian could be divided into two branches, except that the ancestors of Persian and Sanskrit speakers apparently migrated together through the Middle East over a long period of time. Many settled in Iran. The Iranians' descendants became followers of the prophet Zoroaster, who believed in a cosmic war between good and evil. Their scriptures were the *Zend-Avesta* of about 1000 B.C., in Avestan, which later died out. Old Persian, the other dialect of Iranian, is preserved in Darius's cuneiform of about 522–486 B.C., which is carved on a mountainside near Kirmanshah. Old Persian gave us Pahlavi, or Middle Persian, a later form of the language in which rich literature like the epic *Shahnamah* (Book of Kings) is preserved. In turn, Pahlavi led to Persian, which now may be diverging into Tehran Persian and Afghan Persian. Pashto and Kurdish are other Iranian descendants. When Muhammad's followers marched through the Middle East, the inhabitants of Iran became Islamic. We can't call the modern Persians a part of the Arab World, however, because they use Arabic only in reciting the Koran and other religious rituals.

Leaving Iran behind, the Indic members of the long migration eventually moved on through Afghanistan to the Indus River. They paused long enough to crush the cosmopolitan "High Indus" civilization, which boasted large cities like Harappa and Mohenjo-Daro about 2500–1700 B.C., en route to the Indian subcontinent. There they became known as Hindus (people who live along the Indus), in Hindustan (place of the Hindus). Their literary records, written in Sanskrit, are among the oldest of any Indo-European derivative that we possess. Their religious Vedas (from the IE base **wid,* "to know") were written before 1000 B.C. A wealth of Sanskrit literature is extant, including the epics *Ramayana* and *Mahabharata,* which are longer than the Greek and Latin epics. The oral dialect of Indic, the Prakrits, led to several modern derivatives. Hundreds of millions of Indians speak such derivatives as Hindi (India's national language), Urdu, Bengali, Marathi, Gujarati, and Panjabi. Some of the areas where a given Indic language is concentrated are now states, approved by the Indian Government after pressure from the inhabitants for a "language state." The interesting Gypsy tongue of Romany also comes from Indic.

The Germanic Languages

Let's return to the Germanic branch. It was one of the last branches to expand, and its speakers carried it primarily into Celtic areas. It initially caused scholars many headaches because the early speakers were nonliterate and left no written records. This generally unified group of dialects of the first millennium B.C. had to be reconstructed without recourse to records. A real complication was that, over many centuries, most Indo-European consonants had shifted to other sounds in the process of becoming Germanic. Successful reconstruction required a reconciling of the consonants in Germanic words descended from Indo-European with their original forms in non-Germanic cognates. The comparative method used in proving these shifts of sounds, as we'll see shortly, gave us Grimm's and Verner's laws.

Germanic differentiation from Indo-European resulted in groups of dialects in the east, north, and west. The cleavage is observable among some of their descendants today. East Germanic was easiest to reconstruct. The Goths, such a thorn in the Romans' side, were literate. They left posterity a fourth-century biblical translation, from which Gothic was principally analyzed. The language itself died in the Crimea around the sixteenth century.

North Germanic records date from fourth-century inscriptions composed in runes. These twenty-four angular letters tell us much about Old Norse, the early North Germanic tongue. Several very similar languages developed out of it. Indeed, speakers of Norwegian, Swedish, and Danish are able to understand each other. These three languages might even be called dialects, particularly the first two. Another descendant of Old Norse, Icelandic has a fine literature in the form of heroic poems, or Eddas, dating from about the second half of the thirteenth century.

West Germanic has been the most politically important of the three Germanic lines, although there was apparently considerable communication among all three. Geography played an important role. High German developed in mountainous southern Germany, and Low German in the northern lowlands. Old High German is usually dated 600–1100. Middle High German, from which Yiddish is derived, is dated 1100–1500. Modern High German goes from 1500 to the present. Such traditional terminology is unsatisfactory. It is employed here only because the *Old,*

Middle, and *Modern* designations are still applied to English. English and other Germanic languages have changed considerably since 1500, so that the term *Modern* encompasses too much time. Some American linguists call the contemporary language *Present-Day English.* We'll use that designation, although later linguists will have to worry about a name for tomorrow's English, which will be "present-day" to its speakers. English came from the lowland variety of West Germanic, which spread to the English Channel. Dutch, Frisian, and Plattdeutsch also derived from Low German, as can be seen in Figure II (page 32). Plattdeutsch as a Germanic language has about the same status that cockney has in English. Flemish is a modern Dutch dialect.

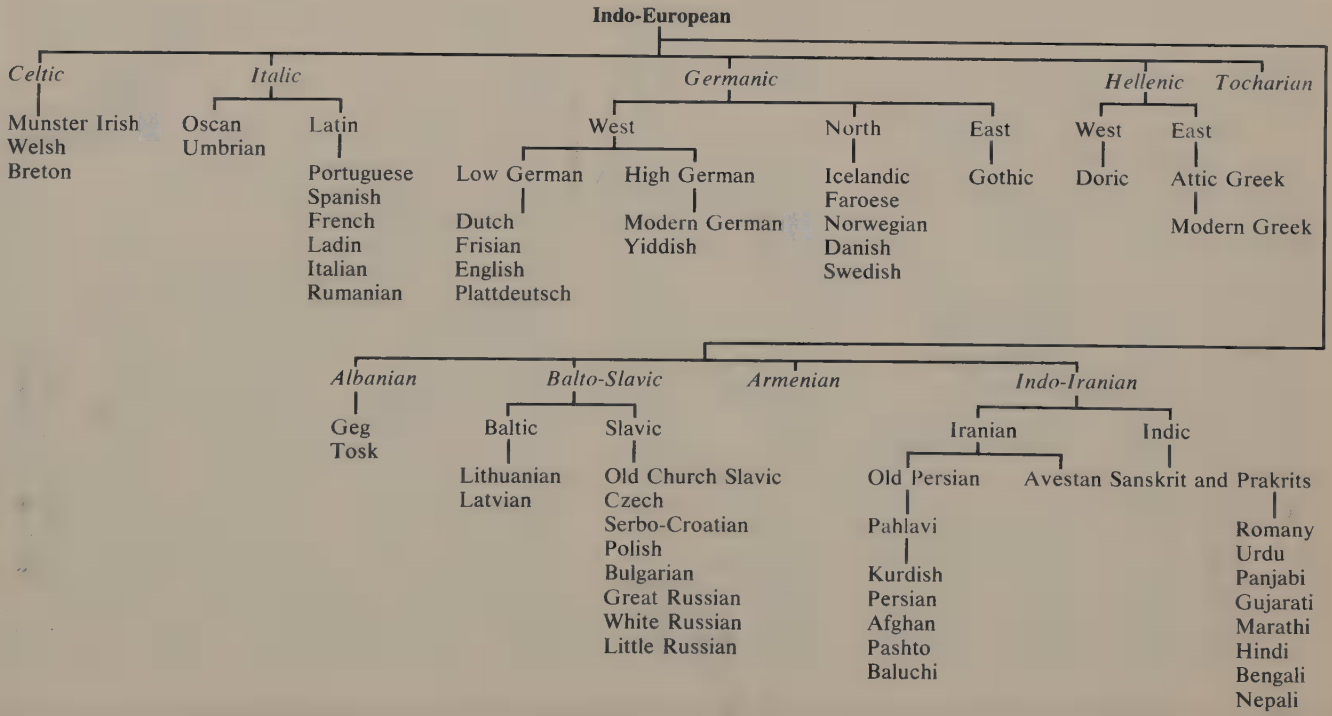
Grimm's Law

As has been said, differentiation toward this set of three geographical Germanic dialects included a shift in almost all Indo-European consonants. In Celtic and in some of the other Indo-European branches, the consonants remained relatively unchanged over the intervening centuries. The Germanic shifts were consistent. *Pente* and *fünf,* and *deux* and *two* in Figure I are examples. The two pairs illustrate the change of /p/ to /f/ and of /d/ to /t/ by the time that Germanic had developed. Jakob Grimm, the great fairy-tale collector, gave his attention to the overall problem in 1822. His explanation is called Grimm's law, which is partly built upon Rasmus Rask's discoveries in 1818.

To appreciate the systematic quality of the consonant shift, we need a brief description of the human vocal tract, or air passage. It is possible to produce four different degrees of narrowing in the air passage.[4] Each degree produces a different set of **phonemes,** or unitary sounds. **Stops** are produced when two opposed parts of the passage make contact, completely damming the breath stream. The lips can obstruct the air to produce the pair of labial stops, /p-b/. The same can be done with the tongue-apex against the upper teeth or gums, to produce the dental stops, /t-d/. When the back of the tongue makes contact with the front of the palate, /c-j/ are formed. Contact of the back of the tongue against the back of the palate, velum, or uvula produces /k-g/. By

[4] From Morris Halle, "On the Bases of Phonology," in Jerry A. Fodor and Jerrold J. Katz's *The Structure of Language* (Englewood Cliffs, N.J.: Prentice-Hall, 1964), pp. 326–28.

Figure II:
Derivational Chart of
Selected Indo-European Languages



feeling the vibration in our Adam's apple, we can determine that a sound is voiced. There is a vibration when we say /b, d, j, g/, but none with their voiceless counterparts, /p, t, c, k/. Indo-European had /p-b, t-d, k-g/ as well as three voiced, aspirated stops, /bh, dh, gh/. Aspiration means producing a sound with a puff of suddenly released air.

The second degree of narrowing, causing near-damming of the air passage, produces **fricatives.** The opposed parts of the passage don't quite touch, but the narrowing causes turbulence or a rubbing sound. Air comes through the mouth under pressure. There can be at least five pairs of voiceless-voiced fricatives, /f-v, θ-ð, s-z, š-ž, x-ɤ/. Pronunciation of /f-v, s-z/ is known from ordinary spelling. The /θ/ can be illustrated by the initial phoneme in the word *thistle,* misleadingly represented by two letters. The /ð/ can be heard in *the,* if we don't pronounce the word with /d/! The /š/ is found in *shack,* again poorly represented by two letters. *Collision* contains /ž/. The velar /x/ can be illustrated terminally by the German interjection *ach* or by music lovers' pronunciation of Johann Sebastian's last name. The same production, when voiced, gives /ɤ/.

The third degree—without police brutality—involves much less obstruction of the air passage. **Glides** like /y, h, w/ are illustrated in the initial sound of *yet, hot,* and *wet,* respectively. The last degree is non-narrowing. When the vocal tract is open, **vowels** are formed. These contrast with all the sounds produced by the other degrees. (We'll discuss the vowels later.) Thus the degree of narrowing of the air passage creates a kind of continuum. It extends from complete damming in stops, to complete openness in vowels. Somewhat outside this continuum are the **nasals,** sounds produced when the air goes up through the nasal cavity and out the nose. If the lips make contact, we have /m/. Behind the lips, there are the alveolar /n/ and the palatal /ŋ/, as in *bin* and *bing,* respectively. Perhaps no language has all the stops and fricatives produced by the first two degrees, but it may be helpful to chart them anyway. The phonemes /p-b/ are labial, moving to velar ones like /k-g/ on the far right:

stops	p-b			t-d	c-j	k-g
fricatives		f-v	θ-ð	s-z	š-ž	x-ɤ

Development of the Germanic dialects of about the first mil-

lennium B.C. resulted in several sets of changes in the Indo-European phonological component. The aspirated, voiced stops /bh, dh, gh/ eventually changed to the voiced stops /b, d, g/. As a result, we can contrast Sanskrit with English: *bharami (bear), bhrātar (brother), dhur (door),* and *dhub (deep).* Indo-European had **ghostis (guest).*

Indo-European voiced stops shifted to voiceless fricatives except when preceded by /s/. The /p/ changed to /f/, the closest fricative, as can be seen on our chart. The same process occurred with /k/ to /x/. The /t/ didn't change to /s/, the closest fricative, because of certain intrinsic features of /s/, which persisted unchanged. Nor did it change to /š/ because of equivalent features. Thus /t/ became /θ/, the closest voiceless fricative. Numbers of examples come readily to mind: Latin *piscis (fish),* Greek *pyr (fire),* and Latin *pedem* (Danish *fod,* English *foot*). To illustrate the rule **t → 0,** we can note Latin *tres* and *tonare.* What are the English cognates? For the **k → x** shift, Latin is again helpful: *quid (what), octo (eight),* and *centum (hundred).* In *hundred,* initial /k/ has become /h/. This is true too in *cordem (heart), cornu (horn),* and *carpēre (harvest).*

Close study of these examples leads to a harrowing conclusion. Today's English can have no voiceless stops, for we've just seen all the little characters shift to voiceless fricatives! Actually, a third shift occurred after the voiceless stops had become fricatives. Again, a whole set of sounds was affected. Voiced stops became the temporarily lost /p, t, k/ in the dialects developing toward Germanic:

b ⟶ p:	cannabis-hemp, turba-thorp, baite-pad, lab-lip
d ⟶ t:	dentem-tooth, decem-ten, deux-two
g ⟶ k:	granum-corn, ager-acre, genu-knee

Another example of the third shift, the change of /b/ to /p/, explains *lip* as opposed to Latin *labium.* If this had not occurred, English would have been *p*-less. Let's suppose that the second set of shifts could have operated a second time. Then *lip* would have become **lif, two* something like **dwo,* and *knee* a form like **hnee.* However, the shifts occurred in order and only one time each, so that there could not be a circular correction. The third shift operated on words already changed in the second shift. Inspection of examples suggests this sequence:

IE *ped- → *fod → foot (**p** ⟶ **f**, then **d** ⟶ **t**)
IE *kerd- → *hord → heart (**k** ⟶ **h**, then **d** ⟶ **t**)

Verner's Law

Grimm himself observed numerous exceptions to the otherwise regular change of Indo-European stops to certain other consonants. Using the comparative method, he had discovered the three sets of changes. Voiced stops lost their aspiration, voiceless stops changed to voiceless fricatives, and then voiced stops became voiceless. The exceptions arose in words of more than one syllable such as those seen in Figure I. Consider Latin *septēm* and *seven,* or Sanskrit *pitār* and *father.* The rule **p** → **f** should have given something like **sefen.* The rule **t** → **θ** should have given *father* with the medial voiceless /θ/, rather than the voiced /ð/ that we actually have. In reality, as the stress in Germanic moved to the initial syllable in words of two or more syllables, except when the first one was a prefix, there was a shift in some consonants. (But note some Americans' pronunciation of *résource, résearch, cément, économics, fínance!*) The rule explaining this shift was discovered by Karl Verner in 1875. He showed that all voiceless fricatives occurring between vowels became voiced unless the preceding vowel was stressed:

medial **f** ⟶ **v:** septēm-séven, carpēre-hárvest
medial **θ** ⟶ **ð:** pitār-fáther
medial **s** ⟶ **z** ⟶ **r:** IE *wes → Germanic *wæz → Old English wǣron → were
medial **x** ⟶ **ɣ** ⟶ **g:** IE *dherāgh → Old English dragan → draw

Grimm's and Verner's laws are considerably more complicated than this brief presentation may imply. Yet the changes are basically simple and systematic, operating in sets and in the order given. The modern word *brother* was probably developed in the following simplified four steps:

IE	**bh → b**	**t → θ**	**θ → ð**	**stress shift**	
*bhratēr-	*brotór	*broθór	*broðór	*bróðor	brother

The last two steps illustrate Verner's law. Given **pater* as the

Indo-European original, how might *father* have developed? With the completion of these sound shifts, some other phonological changes (including the stress shift to the initial syllable), and certain lexical and syntactic changes, the Germanic dialects were finally differentiated from the other Indo-European dialects. Out of Low German came Old English. In turn, Old English itself developed several dialects.

Activities

1. List words naming some of the topographical features and flora and fauna in your area. Do these phenomena occur elsewhere in the United States? Did they occur in the Indo-European homeland? Check a good dictionary to determine the etymological sources of the words. Do any of the sources surprise you? What are the sources of *bamboo, banyan, crocodile, monkey, rabbit,* and *partridge?*
2. Which one of the eight (or ten) Indo-European branches figures significantly in the Bible? How can the geographical "misplacement" of centum Tocharian and satem Albanian be accounted for? Using colored pencils to locate the Indo-European branches, copy the contemporary language map of Europe and Asia Minor in Figure III.
3. What are the five official United Nations languages? How did political conditions influence selection of these? Why do you suppose Japanese and German aren't included? How many of the official languages are Indo-European derivatives? What is the chief language of diplomacy today?
4. Words are often borrowed from the language of areas that are politically or culturally influential. Why would you expect large-scale Russian borrowing into Bulgarian, Czech, Polish, and Rumanian? Why is Rumanian a special linguistic case? What languages would you expect English to be influencing today?
5. Why did the Germanic languages cause headaches for scholars? How is the word *literate* used in this book, apart from other meanings that you know? Why might Norwegian and Swedish be called dialects rather than separate languages? Why has West Germanic been the most politically important Germanic line?

Figure III:
Indo-European Languages in Europe and Asia Minor Today

6. One possible designation for language periods is by centuries. Thus yesterday's English could be Nineteenth-Century English; today's, Twentieth-Century English; and tomorrow's, Twenty-first–Century English. What are some objections to such chronological designations? Why preserve names like *Modern English* in this book when they admittedly create problems?
7. Let's do some reconstructing by testing ourselves on the four sets of phonological changes learned in this chapter. What should the English forms be, given the following sets of data:

   a. *IE* *pleu-, *Rumanian* pluti, *German* fluten
   b. *IE* *edont-, *Italian* dente, *Greek* odont-, *Danish* tänd
   c. *IE* *kel-, *Latin* collis, *Middle Dutch* hille
   d. *Latin* bursa, *Greek* byrsa, *Old Norse* pungr
   e. *IE* *domā, *Greek* damāein, *German* zahm
   f. *IE* *gṇti-, *French* genre, *Old English* cynd
   g. *IE* *ped-, *Icelandic* fat, *German* Fass
   h. *IE* *tu, *Latin* tu, *Old High German* dū
   i. *IE* *bhāgos, *Latin* fāgus, *Greek* phēgos, *German* Buch

Chapter Four
The Generative-Transformational Process

As we saw in the first chapter, every language is a method of communication that is oral, symbolic, systematic, arbitrary, recurrent, social, purely human, adequate, and noninstinctive. In addition, babies don't seem to encounter difficulty in learning the language spoken by their mothers. They learn it even though what they've mainly heard is "baby talk" or just noise. They learn it whether the language belongs to the twentieth century or, by extension, was Indo-European. It's highly unlikely that every baby repeats the entire evolutionary process of *Homo sapiens* in acquiring speech. If so, some babies would presumably never succeed in learning their native tongue.

Universal and Particular Grammar

Each human baby apparently has a capacity, unique to man, for learning and using language regardless of his intelligence. However mysterious (because unobservable directly), this capacity dictates many crucial features of every tongue. Thus each language is a union of two sets of procedures and rules. Since one set is found in every language, it may be termed **universal grammar.** As we can't see its mechanisms directly, its existence is admittedly an hypothesis rather than a fact that we can prove for ourselves. Yet this hypothesis explains the deep-seated regularities shared by all tongues. Universal grammar permits and perhaps even requires the baby's easy, inevitable acquisition of his parents' speech, and in this sense, makes all languages much alike.

The other set of procedures and rules is unique to each individual tongue and so is called **particular grammar.** It makes one's speech English, for example, rather than something else. The differences among particular grammars may be substantial. If one hasn't learned the particular grammar of a language, he can't understand it. Nonetheless, the procedures and rules of universal grammar help insure that if a speaker of Indo-European were alive today, he would be able to learn any contemporary language. After all, the same rules of universal grammar occur in any contemporary tongue that the Indo-European speaker might learn as a second language.

The union of the two sets of grammar helps make original in structure almost everything that anyone might say. The ideas may not be original, but the sequences usually are. Probably few sentences in this book have ever been expressed before in exactly the same sequence of words. As there's no limit to the length of a single sentence, we can demonstrate the remarkable, unlimited creativity of English by expanding a simple structure:

The cobra bit Ali.
The cobra bit Ali and Muhammad.
The cobra bit Ali, Muhammad, and Ghulam.
The cobra bit Ali, Muhammad, Ghulam, and Fatima's sister-in-law who lives in Jerusalem . . . .

Instead of *The cobra,* the snake in the grass might have been *An adder,* which might have *struck and killed* the unfortunates. Hundreds of other nouns and noun phrases and verbs and verb phrases might have been used instead. The capacity for such substitution, or equivalency, further enhances man's linguistic individuality. Substitution seems to be a process common to all languages.

Every baby's mind would seem to have in latent form the needed language-learning mechanisms. Perhaps both physiological and psychological, these mechanisms enable him to quickly master the necessary universal and particular grammar of his mother's tongue. One of the first rules of universal grammar that a baby masters is one like **S** → **Sequence** + **Intonation.** That is, **S** (the notion of what constitutes a sentence) may be rewritten as the more specific constituents of **Sequence** and **Intonation.** These are still quite abstract and are later replaced by the actual words of the sentence, together with the stress, pitch, and juncture con-

stituting the intonation. This rule belongs to universal grammar because every language must have a sequence plus some kind of intonation. The speaker probably doesn't have conscious knowledge of this rule or of the other actual rules whereby the sequence finally becomes sounds recognizable as a sentence by other speakers of that tongue. But he uses the rules intuitively in producing his own sentences and in interpreting those said to him.

Suppose the first rule in the generation of a sentence is **S → Sequence + Intonation.** What's the next? Let's say that it's **Sequence → (Conj) NP + PP.** Here, **NP** means "Noun Phrase," **PP** means "Predicate Phrase," and the pair of parentheses means an optional conjunction. This deep-structure, or basic, rule can't be part of universal grammar, since in some languages the verb precedes the subject. The rule obviously belongs to the particular grammar of languages in which the noun comes first. Because the same deep-structure rules are presumably found in all the dialects of a given language, communication across dialectal boundaries is made possible, as was earlier observed. Within universal grammar there is a generalized rewriting procedure like the one just used above, which is used in particular grammar too. The procedure begins with an abstract notion and concludes with a real sentence in an incredibly brief time in terms of the human being's actual production of that sentence.

Deep-Structure Rules

The rules of a particular grammar, say of English, are of several kinds, which are used in sequence. The first are the deep-structure rules. Here are the ones that we'll use in this book:

$$\text{Sequence} \longrightarrow \text{(Conj) NP + PP}$$

$$\text{PP} \longrightarrow \text{Aux + VP (Place) (Time)}$$

$$\text{VP} \longrightarrow \left\{ \begin{array}{l} \textit{be} + \text{Predicate} \\ \text{V} + \text{S}^1 \\ \text{V} + \text{Predicate} \\ \text{V} + \text{(NP) (Prep-Phrase) (Prep-Phrase) (Manner)} \end{array} \right\}$$

$$\text{Aux} \longrightarrow \text{Tense (Modal)}$$

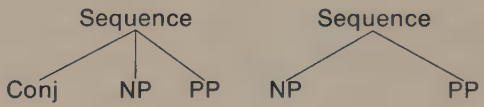
$$\text{Predicate} \longrightarrow \left\{ \begin{array}{l} \text{A} \\ \text{Adv} \end{array} \right\}$$

$$\text{NP} \longrightarrow \text{(D) N}$$

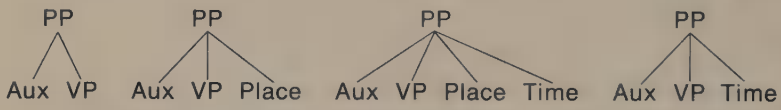
$$\text{D} \longrightarrow \text{Art}$$

Though these rules of Old English look forbidding,[5] they're actually simple and account for much of Present-Day English too.

Aux stands for auxiliary elements like modals (words like *can* and *will*) and tense. **S**[1] is an embedded sentence, one which is inserted within a main sentence. **VP** is a verb phrase, **A** is an adjective, **D** is a determiner, and **Art** is an article. The braces indicate a single obligatory choice of the elements enclosed; the parentheses, an optional choice. Thus the first rule can provide two structures, as illustrated in the following deep-structure **tree,** or graphic presentation:



For either **PP,** we then have one of four choices:



Since any of the four possible predicate phrases can have an initial conjunction, we have a total of eight possible structures. As **Aux** can also be rewritten with one of the five modals, the number is doubled to sixteen. If we count all the modal possibilities for each of the sixteen, we enlarge the total number of possible deep structures to forty-eight. If *be* is selected, a sentence like "He is happy" or "He is here" (from **NP** + *be* + **Predicate**) can ultimately be generated. If *be* isn't selected, any one of three kinds of structures containing a verb can be generated. The first structure embeds a whole new sentence in the basic structure, as in "She said, 'I'm hungry.' " A second possible structure has a predicate which, in turn, can be an adjective or an adverb, as in "The work looks easy" or "Jets go fast." A third possibility has an optional **NP,** which can provide for a direct object, as in "He ate candy." This third sequence can also include up to three kinds of prepositional units. None of the units or the **NP** is obligatory. For example, the structure might be just **V,** or an intransi-

[5] Modeled on Noam Chomsky, *Aspects of the Theory of Syntax* (Cambridge, Mass.: M.I.T. Press, 1965), pp. 106–07.

tive verb, as in "The miners migrated." If all three optional prepositional units are used, the structure can be something like "The miners migrated (to the west) (for years) (with determination)." "With determination," of course, is an adverb of manner, as specified in the rule.

Insertion of Words

These examples presume completion of the second step in the generative-transformational process. That is, we must insert the words that replace the various deep-structure elements. Suppose we've generated a sequence of elements like **Art + N + Tense + V + N.** The individual human mind specifies features for each of these. Let's say that the article has been arbitrarily selected as the definite *the,* which we'll write as [+ definite]. We'll also presume the noun subject to be [+ animate], the verb to have a matching [+ animate] feature, and the noun object to be [+ human]. Words possessing these features are plugged in, giving a sequence like "The shark **Tense** eat people." Why can't we have "The people **Tense** eat the shark"? The deep structure above doesn't contain an article before the direct object (as *shark* does in our sentence), and *shark* isn't [+ human]. Why can't we have "The stone **Tense** love liberty"? Neither the subject nor the object possesses the right features for our structure.

It's at this point in the process that dialectal variation begins. One speaker might employ one word, and another a different choice. For instance, a Southwestern cowboy today might prefer the term *lasso* in the sequence "The rider _____ the steer," as opposed to an Easterner's choice of *rope.* Presumably all speakers of a language have the same rules and procedures of universal grammar, as well as the same deep-structure rules of their particular grammar. But each person's lexicon is a little different from everybody else's, as is his pronunciation.

Transformational Rules

To complete the syntactic component, we next apply transformational rules, or **T** rules, to change the deep structure to a surface one. We'll illustrate nine of the common **T** rules. **T-infl** arbitrarily changes the generalized tense to a present or past inflection, or gives certain nouns a plural form, as when *foot*

becomes *feet*. If **past** is selected for the verb *rope*, the form eventually becomes *roped*. If someone's idiolect has **ripe* as the past, he will use that nonstandard form. If **present** has been selected, **T-agr** must be applied so that the verb will agree with the subject. Then /s/ is added to the base *rope* because *rider* is third person singular. **T-contr** contracts word combinations like *can* + *not* to *cannot* or *can't*.

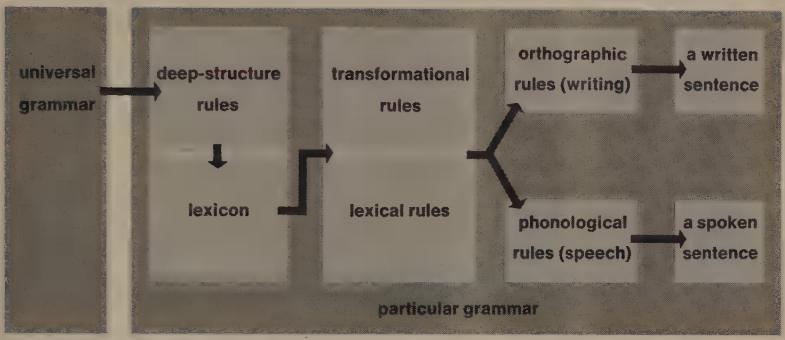
Technically, transformational rules operate within surface structure. They alter the sequence of inserted words of the deep structure in certain ways. By adding or deleting words, or by changing word order, they can significantly alter the syntax. For example, **T-ques** inverts the order, changing a declarative sentence to a question. When it converts "The team has lost" to "Has the team lost?" there is also an obvious semantic difference in addition to the syntactic one. Or consider **T-pass,** which can change an active-voice structure like "The Colts defeated the Cowboys" to the passive "The Cowboys were defeated by the Colts." **T-del** can delete "by the Colts" if so desired. **T-adj** can transpose an adjective, changing "The man is stern" to "The stern man." **T-adv** provides stylistic permutation of adverbs by changing, for example, "He left quickly" to "Quickly he left" or "He quickly left." **T-perm** stylistically permutes a verb-preposition sequence like "The sower went out" to "Out the sower went." These nine transformational rules are context-sensitive; that is, they can be applied only to the structure specified.

About the same time as the transformational operation, lexical rules can be applied to the context to determine the meaning of each of the words in a given sequence. Lexical rules give directions for choosing the right meaning from among all the possible meanings of a given word, once it has been inserted into a particular sequence. Finally, phonological rules provide an actual phonetic representation of the sentence, specifying the pronunciation. If a written version is preferred, then orthographic rules operate to provide the proper spelling and punctuation of the sentence.

Summary

There are six steps, then, in the generative-transformational process: application of the rules and procedures of universal grammar, generation of the elements of a given deep structure,

Figure IV:
Flow-Chart of the Generative-Transformational Process



replacement of the elements with the actual words, provision of the semantic meaning, transformational change of the sequence to a surface structure, and provision of a phonetic or orthographic form. The flow-chart in Figure IV will illustrate.

The deep-structure rules, unlike those of particular grammar, change quite slowly. The transformational rules (which, among other functions, supply the suffixes that may be in the language), the lexical rules (which select the meaning), and the phonological rules (which provide the pronunciation) change more rapidly. These rules, perhaps a little abstract at the moment, will be richly illustrated as the various chapters unfold.

Collectively, these rules compose the grammar of competence of any language, or the speaker-hearer's total linguistic knowledge of his language. As has been said, they can't be directly verified. They do account for all the language data in actual performances, which *can* be taped, directly investigated, and described. Analysis of a particular conversation, however, applies only to that conversation. Moreover, the analysis of a particular conversation usually includes a description of noise, false starts, unplanned changes of structure in mid-sentence, and other variables stemming from spontaneous speech. Often the conversational situation is unanticipated, including an unexpected listener. A grammar of a particular performance can never be representative of one's linguistic capacities, which are infinite in potential. Rather, the capacities derive from one's competence. They are the source of rules that can generate an infinite number

of novel sentences, which can be actually produced as a given person's performances of speech and writing. The sentences can be understood by other native speakers of his language. Of course, each person's linguistic competence differs minutely from everybody else's. Each person knows and uses a few words not used by others; he employs special pronunciations and structures that help make his every performance partly idiolectal and thus unique to him.

We'll conclude this general description of the generative-transformational process by summarizing the nine **T** rules discussed, since we'll be using them throughout the rest of the book:

T-adj:	moves the adjective in front of its noun headword
T-adv:	moves the adverb forward to certain positions
T-agr:	makes the verb agree in number with its noun subject
T-contr:	contracts certain word combinations
T-del:	makes certain deletions
T-infl:	gives a past or present inflection to the tense, and number to a noun
T-pass:	changes active voice to passive
T-perm:	moves certain kinds of units forward to certain positions
T-ques:	changes a declarative sentence to a question

Activities

1. Every language contains the rules and procedures of universal grammar. Why does this fact help make languages very much alike? What then causes a speaker of A to be unintelligible to a speaker of B?
2. What are the two principal sources or means of creativity in language? Which is infinite? Which is finite? Why? Take the example "The cobra bit Ali" and substitute other words, one by one, until no word of the original sentence remains. What substitutions particularly show the influence of your idiolect upon your choice of words?
3. Why can't the rule **Sequence → PP + NP** be part of universal rather than particular grammar?
4. Tape a spontaneous discussion in the classroom. If you were to make a linguistic analysis of the structures therein,

how valuable would it be as an analysis of the language itself?

5. Match the following deep structures with the sentences given:

   a. Conj + N + Tense + Modal + V
   b. Art + N + Tense + *be* + Adv
   c. N + Tense + V + Adv
   d. N + Tense + V + N + Manner

   (1) Birds eat worms greedily.
   (2) And John could go.
   (3) The interview is on time.
   (4) Everybody left in a hurry.

6. Apply the stipulated **T** rule to the particular sentence:

   a. We can eat rice **(T-pass, T-del)**
   b. We can not eat rice **(T-ques, T-contr)**
   c. The boy is exhausted **(T-adj)**
   d. Roy got up **(T-perm)**
   e. The saucer flew swiftly **(T-adv)**

Chapter Five

Old English (449–1100)

The growing Germanic power caused the Romans to desert their Celtic colony of Britain about A.D. 410. They were pulled back to Rome from the land that they had called Britannia since Julius Caesar's arrival in 55 B.C. The unsuccessful defense of their capital against the barbarians meant the end of cultured, literary Rome in 410, when the Visigoth Alaric sacked the city. Other Germanic tribesmen, barely literate, invaded Britain, but left few records of their incursions. Our little direct information about them and their dialects consists of brief runic inscriptions cut into stone.

Beginnings of English

English might be said to begin with the landing of the Jutish brothers Hengest and Horsa in A.D. 449. They landed in the southeast of Britain, which the Celts called Kent. Arriving in small groups during the fifth and sixth centuries, the Angles, Saxons, and Jutes brought their families, won small wars, and occupied the conquered areas. As soon as they started crossing the North Sea to England, at least four dialects began to develop. The boundaries roughly conformed to the areas in which the tribes settled. The Jutes—who settled in Kent, the Isle of Wight, and just north of Wight—spoke what came to be called Kentish. The Saxons, west of Kent and primarily south of the Thames, spoke West Saxon. The Angles, who landed north of the Thames, spoke Anglian, usually divided into the similar dialects of

Mercian and Northumbrian. Because the four dialects were closely related, communication was no problem.

The Anglo-Saxon Heptarchy

Geographical and military conditions helped determine where the migrators settled in Britannia. The weaker, outnumbered Jutes remained in the southeast. Expansion to the north was blocked by the Thames outlet to the North Sea and by the Saxon settlements just above, in what became the kingdom of Essex (i.e., *East* + *Saxon*). The Saxons, solidly blocking the west, had spread as far west as they could until they met the hostile Scots in Wales and Cornwall. Their boundary roughly coincides with the eastern limits of modern Wales, although they eventually drove to the tip of Cornwall. The powerful Saxon kingdoms of Sussex (*South* + *Saxon*) and Wessex developed west of Kent. What's the derivation of *Wessex?* There was no **Norsex*. Lands to the north were occupied by the Angles, who also advanced as far westward as the Scots and fleeing Britons would permit. The Angles controlled the land up to the Firth of Forth, where the Picts blocked them. Their territorial limits roughly coincide with the modern Welsh and Scottish boundaries. Wresting more territory from the Celtic inhabitants than either the Jutes or the Saxons, the Angles established the kingdoms of East Anglia, Mercia, and Northumbria. Thus by the early seventh century there were seven Anglo-Saxon kingdoms, usually called a heptarchy, as shown in Figure V (page 50).

Until Alfred the Great became king in 878, power moved from one kingdom to another over the centuries. All seven were never prominent simultaneously. Early in the seventh century, Northumbria was politically and intellectually dominant. Mercia was the leader in the next century, until Egbert, the king of Wessex from 802–39, gained the allegiance of the other kings. Two of the tribes (not the Jutes) were dominant, as witnessed by the name *Anglo-Saxon.* As a result, fewer people spoke the Kentish dialect. West Saxon extended from the Kentish boundary, roughly along the Thames and over to the Severn, then up to Bristol Channel, including Cornwall after a time. Ultimately Cornish died, as well as Breton in Britain. Some Breton speakers retreated across the English Channel to Britanny, where their descendants still speak the language. Mercian was spoken in the

Figure V:
Anglo-Saxon Heptarchy of the Seventh Century

Midlands area, between the Thames and the River Humber, which included the Saxon kingdom of Essex. North of the Humber, Northumbrian was spoken. How was the place-name *Northumbria* derived?

The Germanic invaders apparently exterminated many of the Celtic inhabitants of what is modern England except in Cornwall and the Cumbrian mountains. They didn't really intermingle, as the Roman conquerors had. The Saxons were so dominant that their name was applied indiscriminately to the Angles and Jutes too, although the word *English* ultimately came from *Angle*. Usually the Celts were killed or had to retreat, as their enemies' footholds were extended. Celtic was one of three language groups with which the four Old English dialects came into contact.

Celtic Loanwords

In view of the Anglo-Saxons' actions, we shouldn't be surprised that Celtic had little influence on English. Honorably, the Germanic men of the long boats often retained the Celtic names of places that they conquered. They particularly borrowed the names of the rivers used as their invasion routes to the interior. Thus we have *Avon, Dēvon, Exe, Kennet, Lea, Ouse, Usk,* and *Wye*. Scores of other Celtic place-names can be listed: *Abergavenny, Bryn Mawr, Carlisle, Cornwall, Cumberland, Dover, Leeds, London, Malvern, Wight*. Few words that aren't place-names have been verified as Celtic borrowings. Among these are *bannuc* (cake), *bratt* (cloak), *broc* (badger), *carne (cairn), crag, cumb* (valley), *glen,* and *loch*. (In this book we will italicize the word given in parentheses when it is a derivative, like *cairn* from *carne*. There will be no italics when the word constitutes the meaning, as in "valley," which is the meaning of *cumb*.) *Cumb* is common in place-names, as in *Greenscombe* and *Watercombe*. Addition of the place-names and a few other Celtic words didn't affect English syntax. The phonology was influenced little, since similar English sounds were usually substituted for the foreign ones. Thereby the items became completely naturalized.

Latin Loanwords

Latin was the second language of the three to come into contact with the Old English dialects. Unlike the Celtic lan-

guages, Latin contributed many words to English by 1100. Some replaced native words, accelerating the process of differentiation of the English vocabulary, or lexicon, from that of other Germanic tongues. Hundreds of Latin loanwords were already scattered among the Germanic dialects because of constant intercommunication with the Romans on the Continent. The Anglo-Saxons had about 175 Latin borrowings in their lexicon when they began landing in Britannia. These words primarily referred to household life, foods, trade, and military activities—i.e., ideas and objects for which there sometimes hadn't been native items. Most of these early borrowings died. Words like *ancor (anchor), flasce (flask),* and *mīl (mile)* have survived. By comparing such written examples with later spellings, linguists have reconstructed the pronunciation. Other borrowings were

biscop	disc	pīpe	sacc
butere	mūl	pund	strǣt
copor	pic	pytt	wīn

What are the modern versions?

Determination of the approximate time of borrowing is always difficult. For one thing, scholars must make several arbitrary decisions. For example, they must wrestle with the problem of how many different speakers must use a given word, and how many times it must be used by the various speakers before it can be considered a legitimate part of the language. Neither the Germanic nor the Danish invaders somewhat later could read the manuscripts and handwritten books found in the monasteries and other places of learning. They preferred to plunder the gold and silver. Not needing the manuscripts, they often burned these precious historical and linguistic records along with the monastery. Until the adoption of the Christians' Roman alphabet and the practice of writing on parchment, the Anglo-Saxons couldn't leave posterity any manuscripts. By that time, additional Latin words were entering Old English. A few, originally borrowed by the Celts, were reborrowed by the Anglo-Saxons. Many more were absorbed after St. Augustine introduced Christianity into Kent in 597.

The existence of these three separate periods of Latin borrowing complicates dating the entry of a particular item. We can't be absolutely positive that all the words just listed from the

Continental period were actually in Old English before 449. Even if we discover the first known use of a Latin "religious" word in English to be 871, we can't prove that it was borrowed in that year. Perhaps it was in the lexicon before 449 and mischievously didn't occur in any pre-871 manuscripts that are extant today. Maybe it was borrowed from some Christianized Celts after 597 and didn't occur in any manuscripts until 871. Logically, we might expect a word related to Christianity to have been borrowed directly from Christian missionaries after 597, but we can't really be sure.

The cumulative number of Latin loanwords in English by 1066 is of greater interest. Of the approximately 528 items known to have been borrowed by then, the scholar Mary Serjeantson estimates 175 to have been enlisted before A.D. 449, another 111 by A.D. 650, and a tardy 242 by the time of William the Conqueror's arrival in 1066 with his linguistic bag of Norman French words.[6] Latin and French loanwords were chiefly responsible for driving out about eighty-five percent of the Germanic vocabulary composing Old English.[7] However, such dropping and adding in the lexicon hardly disturbed the syntax. The deep-structure rules remained Germanic. Occasionally a transformational rule like **T-infl** was affected, as when the Latin inflected forms were borrowed along with the bases. Thus the later loanwords, *cacti* and *radii,* are only now being challenged by the "native" plural forms —*cactuses* and *radiuses.* The two nouns have kept their original Latin plurals over all these centuries, resisting the English pattern of an *-es* plural for such items. *Crises,* the plural of *crisis,* is still unchallenged.

English borrowed few Latin words from 449 to 650. These came primarily through the Celtic languages, for the native Latin speakers had left Britannia by 410 to try to defend Rome against the Anglo-Saxons' unprincipled cousins. Some of these second-hand loanwords were *ceaster* (camp), *munt* (*mount*), *port* (harbor), *torr (tor,* a high, rocky hill*)*, and *wīc* (village).

As Kent and the other kingdoms gradually became Christian, additional Latin names for Christian concepts and objects were needed by the end of the seventh century. The concepts and

[6] Serjeantson, *A History of Foreign Words in English* (London: Kegan Paul, Trench, Trubner, 1935), pp. 271–88.

[7] Harold Whitehall, "The English Language," in *Webster's New World Dictionary of the American Language,* College Edition (Cleveland: World, 1960), p. xxviii.

objects were previously unknown to Old English speakers, whose religion was part of the Valhallic tradition. They believed in an unpleasant chief god named Wyrd, who fatefully intervened in men's lives. He rewarded those who died bravely in battle by receiving their souls in a heavenly place for eternal wining, gambling, and other pleasures. Among the dozens of Latin loanwords primarily concerned with Christianity, we can list the following:

abbod	canonic	martir	papa
altare	deacon	mæsse	preost
apostol	discipul	nunne	scrifan
candel	Læden	organa	tempel

Most of these are close to their modern spellings. *Abbot* and *Latin* are perhaps furthest removed. The *-e* or *-a* on several words indicates that the Old English forms of *altar, pope* or *papa,* and three other words were pronounced with an additional final syllable. What are the other three? *Shrive* developed from *scrifan* through a consonant cluster to be mentioned shortly.

During this third period, loanwords not related to Christianity were also brought into the language. Here are some that the ordinary Englishman (instead of primarily the kings and aristocracy) adopted:

botany	**clothing**	**education**	**foods**
balsam	cæppe	dihtan	bēte
lilie	sioloc	fers	lopustre
plante	socc	scōl	rædic

Modern versions of the twelve can be guessed except possibly for *dictate,* which later formed combinations like *dictaphone* and *dictator. Dihtan* is the second Old English verb we've encountered; the first was *scrifan.* By now we should recognize the *-an* suffix that most Old English infinitives had. The /æ/ is the vowel in modern words like *bat.*

Besides giving about 242 loanwords to our language during the Christian period, Latin had a major linguistic and then literary effect. In a way the contribution compares with the vast spiritual and philosophical consequences of converting the Germanic "pagans." What happened was that the *Englisc* scribes began to

use the Roman alphabet. They added æ, *þ* (which became *th*), and *ƿ* (which became *w*), as well as *ȝ* (pronounced as /x/) later. Abandoning the laborious carving of runes onto stones, the Anglo-Saxons began to use manuscripts and adopted the practice of writing at least by the late eighth century. (Just think how many thousands of boulders would have been needed for the inscribing of the 3,182 lines of *Beowulf!*) This epic undoubtedly existed in oral form long before somebody, probably a monk, committed it to writing about A.D. 1000. Such written literature in English presented the first lengthy, direct view of the Anglo-Saxons. The view was very different from the sometimes confused, early reports by Caesar and Tacitus.

No doubt most of the early Anglo-Saxons were nonliterate. Certainly few Old English manuscripts predating 871 have survived. But from these few, especially from manuscripts of Alfred's reign (871–99) and of later West Saxon writers like the abbot Ælfric (died about 1020), the language has been reconstructed. Actually, the four or more Old English dialects varied less than our speech does today. There were no native dictionaries; so the scribes spelled words phonetically. Present-Day English developed from the Mercian line; however, Wessex was politically and culturally dominant after 821.

By the time Wessex conquered Mercia, the three once-Germanic tribes had become essentially English, unified into one people and one culture. Under Alfred the Great, the new Germanic invaders were stopped. They're usually called Danes, or Vikings, although there were some Norwegians. After getting out of the swamp and supposedly enduring some burned cakes, Alfred defeated the Danish leader Guthrum in 878. The poor Dane even publicly gave up the ultimate pleasures of Valhalla for the supposedly greater gift of Christian grace. After the liar Guthrum was beaten again, he agreed to the Danelaw. This agreement gave Wessex, which was essentially England now, control of everything south of a line from London to Bedford and then along the old Roman road to Chester. Also, Northumbria continued under English control. Alfred's successors finally conquered the last of the Danelaw in 954, after a decisive victory over a Danish-Scottish force at Brunanburh a little earlier. Northumbria was again contiguous. England was free of invasion until King Swegen's Danish army began systematic conquest of England in 1013. After his son Canute was acknowledged as the

ruler of all England in 1017, the Danes held the kingship almost until William's arrival at Hastings.

Scandinavian Loanwords

The language of the Danes, Old Norse, was the last of the three languages to influence Old English. This North Germanic tongue of Iceland, Norway, and Denmark wasn't really differentiated into separate dialects until about the time of the Norman Conquest. During the first millennium A.D., Old Norse probably still shared some deep-structure rules with English. Many of its word bases were virtually the same as Old English ones with somewhat different inflections. The differing suffixes and pronunciation probably caused confusion when Old English and Old Norse speakers tried to communicate.

The sharing of many words among languages descended from Germanic often has made it impossible to determine the source of a given Modern English word. It might have been in Old English all along and had an identical Old Norse cognate, or it might have been a Scandinavian item borrowed in the ninth to eleventh centuries. A further difficulty is that most of the Old Norse loanwords first appear in our extant manuscripts of post-1100, the approximate beginning of Middle English. A given word may have been borrowed into either Late Old English or into Early Middle English. It is difficult to determine the exact time, perhaps a fairly minor matter. Overall, there were almost two thousand Scandinavian borrowings during these centuries, not including many British dialectal forms that still exist.

The two thousand also exclude some fourteen hundred placenames. They principally contain *-by* (village), *-thorp* (village), *-thwaite* (clearing), and *-toft* (piece of ground). They're concentrated in Lincolnshire and Yorkshire, where Scandinavian centers flourished during the Danelaw period and later. *By* has been even more productive than *thorp,* including *Derby* (by of the deer), *Rugby,* and *bylaw.* This Old Norse loanword, together with the homonymous Old English preposition *be* or *bī* (by or near), has been extremely productive as a prefix. Besides *bypass* and *byword,* what other derivations might be listed? Scandinavian suffixes are found in some English proper nouns, like the *son* in *Johnson.*

Borrowed proper nouns are seldom as influential as ordinary loanwords in the history of a language. Most Old Norse items were fitted into the English inflectional and phonological system. Two general kinds of Scandinavian borrowings changed the pattern of Old English. First, the parent language of Old English had had an /sk/ cluster that became /š/. The pronunciation of *fisc, gesceap, scip,* and *scort* demonstrates their non-Scandinavian origin. The *sc* was later respelled as *sh*. What are the modern spellings? Many Scandinavian borrowings retained /sk/, which thus had to be reinstated in our language. The /sk/ in many modern words reveals their Scandinavian origin. Because our adjective is *short* rather than **skort,* we conclude that it came through West Germanic. It's from Indo-European **sqer-* (to cut), apparently not a borrowing from Old Norse *skort.* For words beginning with *sh,* we can be fairly sure of their native, non-Scandinavian origin. Some examples in modern English are *shake* (OE *sceacan*), *sheep (scēap),* and *shove (scufan),* as well as *fish, shape, ship,* and *short* listed above.

The English Dialect Dictionary records more than one thousand words containing /sk/. Once we delete Latin borrowings like *school,* even an ordinary dictionary can help attest to the Scandinavian origin of such words. Many of them relate to actions and things still useful today in everyman's lexicon. Consider the list below, the last two columns of which are Old Norse. What are their modern spellings?

scab	scowl	rannsaka	skrapa
scare	skulk	skālpr	skil
scathe	skull	skattr	skinn
scoff	skunk	skirra	skȳ

Since the skunk didn't live in Scandinavia (nor was he missed) and since the word isn't from Latin, we may wonder how the word got into English. It was borrowed from the Indians by American colonists, who soon learned that the pretty, striped animal disliked being petted. *Skunk* just accidentally begins with /sk/. In fact, the reintroduction of /sk/ from the Scandinavian may explain why the word isn't **shunk* today. Another observation might be made. The "negative" meaning of many of the above borrowings may imply that the Danes were ruthless characters. However, we mustn't blissfully add all other negative

/sk/ words to our list. The use of meaning as linguistic evidence is risky. *Scald* and *scourge* describe the sort of thing that the Danes did to their former Continental colleagues, but the words apparently came from Old French.

The second phonological influence on English by Scandinavian borrowings is reflected in many words with /k/ and /g/. Again using our dictionary, we can find words where *c* and *g* aren't pronounced as /k/ and /g/. Among others, we can list *cell, center, cheap, cheese, gem, giant,* and *year* (from *gear*). These can't be Scandinavian. By contrast, note /k/ and /g/ in certain other Old and Middle English words:

ceallian (call)	gelden (geld)	kilte
dic (dike)	gere (gear)	kindlen
dragan (drag)	gessen (guess)	kirrke
fracel (freckle)	kake (cake)	leg

What are the modern spellings of the last four? Interestingly enough, Old English already had *cirice,* which gives *church,* whereas the Scots use *kirk*. They apply the name *The Kirk* to the Church of Scotland. *Celt* fails the Scandinavian /k/ test, since it's sometimes pronounced with /s/ instead of /k/.

Many of the Old Norse loanwords eventually become monosyllabic and are high-frequency items in English today. Of course, a much-used one like *call* is intrinsically no more valuable than a long one like the Latin-derived *antidisestablishmentarianism.* Borrowings like *call* simply joined the familiar, ordinary vocabulary of English and have retained their utility. Old English plural personal pronouns were later replaced by Old Norse ones: *thei, their,* and *theim,* including *theim selfe* (the latter replaced the earlier *hemselve*). *Boðe (both)* and *same* were also borrowed from Old Norse, as well as the preposition *til* and the adverb *ðēah (though),* which remain important in syntax. Naturally the greatest Scandinavian influence was on Anglian, particularly Northumbrian, since Danes settled heavily in Northumbria, where their descendants still live.

A glance at *Beowulf* or other Old English texts will confirm the high mortality of Germanic words. Many were replaced by Latin or French items. Although Old Norse ultimately contributed a smaller share than either of these languages, Old Norse

borrowings squeezed out high-frequency Germanic words. Here are a few of the modern versions, followed by the obsolete, sometimes related, Old English word: *boon (bēn), cast (weorpan), cut (ceorfan, sniðan), get (gegietan), give (giefan), sister (sweoster), sky (wolcen), take (niman), weak (wāc).*

As Scandinavian adjective, noun, and verb bases slowly merged into Late Old English and Early Middle English, occasionally the entire paradigm was borrowed. Then, of course, the Scandinavian base couldn't be inflected by English rules. As a partial result, since some of the Danish, Swedish, and Norwegian inflections were also starting to be leveled, or die out, the process of English inflectional leveling was undoubtedly hastened. Whenever borrowings contributed toward a leveled or otherwise altered suffix, **T-infl** was naturally affected.

Two examples of the overall effect can be cited. The structure "They are" is entirely Scandinavian. First borrowed into Northumbrian, *aron* displaced the West Saxon *sindon,* the plural of the base *bēon.* Otherwise, today we might well be saying "We, you, they be," like "We, you, they go." Second, the verb suffixes *-s* and *-ing* are usually said to have been Scandinavian importations. The fact that third-person singular is the only inflected member of the present tense (except for the *am* and *are* of the mixed-up *be*) is extremely important. Except for the *-s* borrowing, today we might say *"He go."

In summary, there were three foreign influences on Old English. The Celtic languages furnished few loanwords beyond placenames, without syntactic effect. Latin gave hundreds of loanwords and perhaps had some effect on English syntax. Somewhat more than a hundred of these Latin borrowings survive today. By A.D. 1100 or so, Scandinavian loanwords in English were in the thousands, with still greater influence. This time transformational, lexical, and phonological rules were involved. Together, the three language groups considerably affected Old English; yet the lexicon remained primarily Germanic for a few more centuries. Some of the borrowings competed side by side with native words, until the native items finally bit the dust. The Old English lexicon was relatively small; one comprehensive dictionary lists fewer than fifty thousand words. By contrast, there are roughly four hundred fifty thousand entries in *Webster's Third New International Dictionary.*

Old English Literature

We have no idea of the number of lost manuscripts, which may have contained many additional words. Few Kentish sources remain. Apparently Northumbrian writing was extensive. The chief Old English preservations are runic inscriptions, charters, and biblical and short-verse materials. The earliest literature only dates back to about 750. It includes *Beowulf* (war-wolf) and the Northumbrian "Hymn," by Cædmon. The vision-poem "Dream of the Rood" was composed as a runic inscription about that time. We have a 156-line manuscript version of it from the ninth century. Only eight manuscripts with any quantity of original poetry are extant; more than half the poetry is Christian in theme and subject. These include the Cotton manuscripts (containing *Beowulf*), the Cædmon manuscripts, and the Peterborough *Anglo-Saxon Chronicle* (1070–1154). The extant prose of Alfred, Ælfric, and the archbishop Wulfstan (died 1023) is an additional literary contribution. Most of the extant manuscripts are West Saxon, the dialect used by linguists in describing Old English.

Old English Phonology

Before our study of a sample passage, a sketch of Old English phonology is needed. Otherwise, we're likely to give the letters the value of modern phonemes. If so, we'll mispronounce most vowels, and important later developments like the Great Vowel Shift won't make sense. For example, understanding the modern pronunciation of *ride* requires recognition of /i/ in Old English *rīdan* and its later shift to /aɪ/. The same process is true for the /u/ in *soon,* derived from the /o/ in *sōna*. Moreover, *sōna* has a second syllable; the vowel of this syllable is /ɑ/, pronounced as in *God*. For the /ə/ in *some,* we have to go back to the /ʊ/ in Old English *sum,* pronounced as in modern English *full.*

First we need to recall a rule from universal grammar: **S** → **Sequence** + **Intonation.** Rules of Old English particular grammar then permit us to rewrite **Intonation** into its more specific constituents. These rules do such things as place the stress on the first syllable of a word except in prefixed verbs and some prefixed nouns (e.g., *éare, forscríncan, begáng*). They add the appropriate pitch and juncture for the individual syllables in a

generated sequence. Thus we have something like *ge·hý·re* (*hear*), with the highest pitch on the stressed middle syllable.

The Old English consonants are fairly easy for us, in a simple approximation, since most of them possess roughly the same values as their modern forms. Consider the following chart:[8]

	labial	dental	alveolar	alveopalatal	velar	glottal
stops	p-b		t-d	c-j	k-g	
fricatives	f	θ	s	š	x	h
lateral			l			
apical			r			
nasals	m		n			
semivowels	w			y		

We've already seen /c/, spelled with *c*, in non-Scandinavian English examples like *ceap* (*cheap*) and *ciese* (*cheese*). The letter *c* can also be pronounced with /s/, as in Latin borrowings like *ceder* (*cedar*). Otherwise, *c* is usually /k/, as in *cuman*. When *f* is between voiced sounds, it becomes /v/, as in *ofer*. Elsewhere, it is /f/. What about the *f* in *eft* and *fela?* Like *f*, both *þ* and *ð* are pronounced /ð/ when they are between voiced sounds. How are *cweðan* and *þorn* pronounced? Distribution explains why our modern word *thorn* has /θ/ rather than /ð/. Distribution of *s* between voiced sounds is often /z/, as in *nosu* (*nose*) and *rōse*, as opposed to /s/ in *sunne*. As has been said, the earlier /sk/ became /š/ in Old English; so /š/ is listed on the chart.

There are some fourteen vowels in West Saxon. They can be

[8] All scholars use approximations either in the first stages of their analysis or in summary statements. As they investigate the details, problems often arise, resulting in qualifications or even alteration of that analysis. The above chart is such a summary statement, somewhat generalized for purposes of convenience and simplicity. Some scholars would suggest that /c, j/, listed as stops in the above chart, should be described by the term *affricate*, a sound beginning as a stop and continuing through a fricative release, so that the sound is neither a stop nor a fricative. The chart is also simplified in that the glottal /h/ may well have been the phoneme governing /x/, with /x/ then a special velar variant in distributions like *rūh* (*rough*). See Sherman M. Kuhn's "On the Consonantal Phonemes of Old English," in *Philological Essays: Studies in Old and Middle English Language and Literature* (The Hague: Mouton, 1970), pp. 16–49.

charted according to the height of the tongue and its advancement in the mouth as the vowel is produced:

	front		back round
	unround	round	
high	i	ü:	u
	ɪ	ü	ʊ
mid	e		o:
	ɛ		
low	æ:		ɔ
	æ		ɑ:
			ɑ

The front, unrounded pairs are differentiated by length, so that the symbols /i:, i/, /e:, e/, and /æ:, æ/ are sometimes used, with the colon representing the longer vowel of the pair. The transcription that we'll use in our book is adapted from the International Phonetic Alphabet, a system devised to represent the sounds of a wide variety of languages. The phonemic values of the above vowel sounds will be clearer if we make a list of illustrative words from the passage that we'll be studying shortly:

OE letter	OE word	pronunciation	OE letter	OE word	pronunciation
ī	stīgan	/i/ (beat)	y	syxtig	/ü/
i	hit	/ɪ/ (it)	ū	ūt	/u/ (boot)
ē	hē	/e/ (ate)	u	sum	/ʊ/ (book)
e	eft	/ɛ/ (egg)	ō	gōd	/o:/ (go)
ǣ	sǣ	/æ:/ (air)	o	on	/ɔ/ (jaw)
æ	wæstm	/æ/ (at)	ā	stān	/ɑ:/ (father)
ȳ	fȳr	/ü:/	a	and	/ɑ/ (hot)

An Old English Text

Now let's turn to a passage from a Late West Saxon translation of a Latin Vulgate manuscript from about A.D. 1000. Undoubtedly the unknown scribe's context determined his meaning, as is true today for all languages. After all, most words have several meanings. The meaning intended by the speaker or writer is considerably determined by the deep structure into which he inserts the word, except for cases of ambiguity or just bad diction. In our passage from St. Mark, four contexts give *on* its modern meaning, while three contexts call for "in." On two occasions, *tō* means "to"; the other two times, it means "for." Lexical rules perform much of the determination. The passage itself, from MS. Corpus Christi College Cambridge 140, will show us what Old English writing looks like. It will let us work directly with data in analyzing the key characteristics of the syntactic, semantic, and phonological components. Text A contains the passage, a translation, and a dictionary of the words used:

Text A:
Late West Saxon Translation of St. Mark, iv. 1–10

(1) And eft hē ongan hī æt þǣre sǣ lǣran. (2) And him wæs mycel menigu tō gegaderod, swā þæt hē on scip ēode, and on þǣre sǣ wǣs; (3) and eall sēo menigu ymbe þā sǣ wæs on lande. (4) And hē hī fela on bigspellum lǣrde, and him tō cwæð on his lāre, (5) Gehȳrað: Ūt ēode sē sǣdere his sǣd tō sāwenne. (6) And þā hē sēow, sum fēoll wið þone weg, (7) and fugelas cōmon and hit frǣton. (8) Sum fēoll ofer stānscyligean, þār hit næfde mycele eorðan, and sōna ūp ēode; (9) and for þām hit næfde eorðan þiccnesse, þā hit ūp ēode, sēo sunne hit forswælde, (10) and hit forscranc, for þām hit wyrtruman næfde. (11) And sum fēoll on þornas; (12) þā stigon ðā þornas and forðrysmodon þæt, (13) and hit wæstm ne bær. (14) And sum fēoll on gōd land, (15) and hit sealde ūppstigende and wexende wæstm; (16) and ān brōhte þrītigfealdne, sum syxtigfealdne, sum hundfealdne. (17) And hē cwæð, (18) Gehȳre, sē ðe ēaran hæbbe tō gehȳranne. (19) And þā hē āna wæs, hine āxodon þæt bigspell þā twelfe þe mid him wǣron.

Literal Translation

(1) And afterward he began them by the sea teach. (2) And him was great multitude to gathered, so that he on ship went, and on the sea was; (3) and all the multitude around the sea was on land. (4) And he them many things in parables taught, and them tò said in his teaching, (5) Hear: Out went the seeder his seed for sowing. (6) And when he sowed, some fell by the way, (7) and birds came and it ate up. (8) Some fell on stony ground, where it not had much earth, and soon up came; (9) and because it not had earth's thickness, when it up came, the sun it burned, (10) and it withered, because it root not had. (11) And some fell in thorns; (12) then sprang up the thorns and choked it, (13) and it fruit not bore. (14) And some fell on good land, (15) and it yielded up-springing and waxing fruit; (16) and one brought forth thirtyfold, some sixtyfold, some hundredfold. (17) And he said, (18) Hear, he that ears has for hearing. (19) And when he alone was, him asked the parable the twelve that with him were.

Dictionary

ācsian, verb, "ask"
ān, determiner (cardinal), "one"
ān, adverb, "only"
and, conjunction, "and"
æt, preposition, "by"
bēon, "be"
beran, *v.*, "bear"
bigspell, *n.*, "parable"
bringan, *v.*, "bring forth"
cuman, *v.*, "come"
cweðan, *v.*, "say"
eal, *d.*, "all"
ēare, *n.*, "ear"
eft, *adv.*, "afterward"
ēode, *v.*, past tense of *gān*
eorðe, *n.*, "earth"
feallan, *v.*, "fall"
fela, *n.*, "many things"
forscrincan, *v.*, "wither"
forswǣlan, *v.*, "burn"
for þam, *conj.*, "because"
forðrysmian, *v.*, "choke"
fretan, *v.*, "eat up"
fugol, *n.*, "bird"
gān, *v.*, "go"
gegaderian, *v.*, "gather"
gehȳran, *v.*, "hear"
gōd, adjective, "good"
habban, *v.*, "have"
hē, *n.* (personal pronoun), "he"
hit, *n.* (pro.), "it"
hundfeald, *a.*, "hundred-fold"
lār, *n.*, "teaching"
lǣran, *v.*, "teach"
land, *n.*, "land"

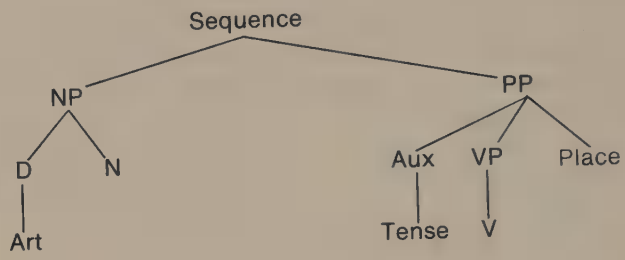
menigu, *n.*, "multitude"
micel, *a.*, "great"
mid, *prep.*, "with"
ne, *adv.*, "not"
ne + habban, *v.*, "not have"
ofer, *prep.*, "upon"
on, *prep.*, "on, in"
onginnan, *v.*, "begin"
sæ, *n.*, "sea"
sǣd, *n.*, "seed"
sǣdere, *n.*, "seeder"
sāwan, *v.*, "sow"
scip, *n.*, "ship"
sē, *d.*, "the"
sellan, *v.*, "yield"
sōna, *adv.*, "soon"
stānscyligean, *n.*, "stony ground"
stīgan, *v.*, "spring up"
sum, *d.*, "some"
sunne, *n.*, "sun"
swā, *conj.*, "so"
syxtigfeald, *a.*, "sixtyfold"
tō, *prep.*, "to, for"
twelf, *d.* (cardinal), "twelve"
þā, *conj.*, "when, then"
ðǣr, *adv.*, "where"
þiccnesse, *n.*, "thickness"
þorn, *n.*, "thorn"
þritigfeald, *a.*, "thirtyfold"
ūp, *adv.*, "up"
upstīgan, *v.*, "spring up"
ūt, *adv.*, "out"
wæstm, *n.*, "fruit"
weaxan, *v.*, "wax"
weg, *n.*, "way"
wesan, "be"
wið, *prep.*, "with"
wyrttruma, *n.*, "strong root"
ymbe, *prep.*, "around"

All this apparatus will be needed again and again. The text is our evidence, arbitrarily divided into nineteen numbered structures. The translation lets us follow the meaning word by word. The dictionary permits checking of the bases of items like *stānscyligean,* composed of *stān (stone)* + *stycce* (a bit of) + *en* (made of).

Our task for the rest of the chapter is to sketch key portions of the particular grammar of Old English. We'll need to recall the deep-structure rules, which generate elements in a grammatical sequence. Then we'll insert words in place of the elements. Lexical rules will supply additional information about meanings of the words. Transformational rules will change the sequence to a surface structure. Phonological rules will finally provide the proper phonetic form. The whole process was described in the fourth chapter, which we may want to review, particularly the flow-chart of Figure IV. The analysis of structures must be brief, with only simple ones described; otherwise, the rest of the book could be devoted to Old English and still be an incomplete analysis.

Analysis of an Old English Deep Structure

We can begin our work with the first part of the eighth structure, which we'll write as (8). Here is the deep-structure tree for "Sum fēoll ofer stānscyligean":



Sum is inserted for **Art;** *sǣd,* for **N.** *Feallan* is plugged in for **V.** What's **Place?** The noun *sǣd* is then deleted by the optional **T-del. Tense** is arbitrarily made past by **T-infl.**

Consider (11), which is much like (8). What element does *and* replace in the deep structure? This is the conjunctional item initiating many sentences. Note how frequently *and* begins a structure: (1–4), (6–7). We've already excluded (8). Now (9–11), (13–17), and (19). What about the twelfth structure, beginning with *þā* (then)? Sentences beginning with *þā* abound in extant manuscripts. They sometimes outnumber those initiated by *and.* Since *þā* is a conjunction, why don't we just change the rule to ***Sequence** → **Conj** + **NP** + **PP?** Then a conjunction would be obligatory for every sentence; but if we do so, (5), (8), and (18) couldn't be generated.

Let's start listing the structures where the **NP** subject precedes the verb, according to our rule:

(1) hē ongan lǣran
(2) menigu gegaderod, hē ēode and wæs
(3) menigu wæs
(4) hē lǣrde and cwæð

What's ailing (2) and (4)? There's no subject-word to tell us who was on the sea and who was cwæðing to the multitude. Of course, we know that the subject is *he* in each case, already given

for the verbs *ēode* and *lǣrde*. These are just sentences with coordinate verbs for the subject pronoun.

In the passage, *cwæð* is the only form of the infinitive *cweðan* that occurs—actually twice. Because Old English is an inflective language like the Latin seen in Chapter Two, *cwæð* automatically tells us things. It's third-person singular, preterit. So its subject can only be *hē, hēo, hit,* or a regular noun. A *hit* can't teach or even speak; it isn't human. It can't take a verb like *cweðan,* which must co-occur with a subject possessing [+ human] features. The only ambiguity left is between *hē* and *hēo.* However, the masculine and feminine genders weren't really differentiated in third-person singular pronouns for a few more centuries. To exclude the *hēo* possibility, the reader simply notes the masculine subject of the preceding verb, which is the subject of *cwæð* anyway. Of the numerous Old English sentences where the subject pronoun is omitted, the vast majority are like this one. Two or even three verbs may be coordinate to the single pronoun, as in (2, 4, 8). What happens is that the surplus pronouns are deleted by **T-del.** We've already seen the rule delete surplus nouns, as in the *sǣd* in (6, 8, 11, 14). In (7) and (12) the second *fugelas* and *þornas* are similarly deleted.

We might complete our subject-verb list:

(5) gehȳrað, ēode sǣdere
(6) hē sēow, sum fēoll
(7) fugelas cōmon and frǣton
(8) sum fēoll, hit næfde and ēode
(9) hit næfde, hit ēode, sunne forswǣlde
(10) hit forscranc, hit næfde
(11) sum fēoll
(12) stigon þornas and forðrysmodon
(13) hit bær
(14) sum fēoll
(15) hit sealde
(16) ān brōhte
(17) hē cwæð
(18) gehȳre, sē hæbbe
(19) hē wæs, āxodon twelfe, þe wǣron

In (5) the absence of a subject for *gehȳrað* shouldn't disturb us. The verb form is the imperative, second-person plural of *gehȳran,* which requires a second-person subject. As all nouns

are third person, only *gē* can be the subject. It is deleted by **T-del.** This "understood" *you* is the only pronoun that can be omitted today, barring structures with coordinate verbs. So (5) still observes the Old English pattern of **Subject + Verb.** Roughly the same explanation accounts for the subjectless *gehȳre* of (18). The form is the subjunctive, third-person singular, expressing a wish or desire. The *hē* subject is later deleted. The complex inflectional system will be discussed when we reach the transformational rules.

In (5)–(19) the verb precedes the subject in three structures. Which? Note the close resemblance of Old English deep structure to ours today. The one-word question *which,* without any noun, has just been asked, like the nounless *sum* in (8). The words "three structures precede the verb," automatically understood from the preceding sentence, have been transformationally deleted. Indeed, most of our nineteen Old English sequences are much like modern ones because deep-structure rules change ever so slowly.

The twelfth sentence is the easiest of the three verb-subject sequences, because structures initiated by *þā* usually have the subject trailing after its verb. The coordinate second verb, *forðrysmodon,* does follow the subject *þornas,* since it has a direct object, *þæt.* What does this **V+ N + V** order do to the deep-structure rule? Nothing. **T-perm** has simply permuted, or inverted, *stigon* in front of "þā þornas," producing the desired surface structure. The rule can be triggered by the initial *þā.* A little earlier we said that sentences beginning with *þā* usually have the verb first. In the passage *þā* occurs four times, but only in (12) is our generalization valid. Why not in (6, 9, 19)? The meaning in these is "when"; in (12), "then." Clearly, we must narrow the generalization: the inversion is usually done for sentences initiated by *þā* when it means "then." The actual meaning is selected by a lexical rule once *þā* is inserted as a conjunction. "Then" is selected as the meaning in (12) because the deep structure is a sentence, not a structure embedded in a sentence. The large syntactic and semantic difference between the two can be illustrated by "when it came up" and "then it came up."

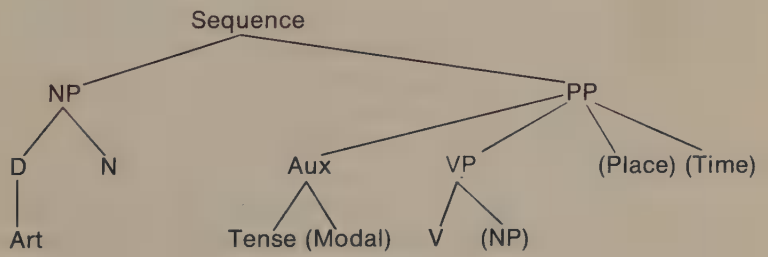
The verb-subject sequences in (5) and (19) pose too many problems for real resolution here. Old English, like our language today, often indicates emphasis by inversion. The complicated

structure of (5) originates as the direct object of the *cwæð* of (4). The inverted "Ūt ēode sē sǣdere" is itself the direct object of *gehȳrað*. Emphasis is on *ūt*, which is part of the verb sequence "ūt ēode" (like *óutgoing* or *íncoming*). The structure wouldn't have been *"Ūt sē sǣdere ēode." Let's just guess that the intonational rule giving *ūt* the heavy stress can optionally trigger **T-perm.** The verb sequence is then moved in front of "sē sǣdere."

In (19) we find the most complicated structure of the entire passage. Emphasis seems to be on Christ, or *hine*, who is being asked the parable. It's not upon the disciples who are doing the asking. If so, we can somewhat glibly explain the subject-verb inversion of (19) and (5) as a stylistic one done by **T-perm.** Thereby we have generated the last of the nineteen deep structures by **Sequence → (Conj) NP + PP.** This rule permits the three major types of Old English sentences:

Subject + V + X
Conj (*þā* meaning "then") + V + Subject + X
Conj (not *þā* meaning "then") + Subject + V + X

The first type might be shown in more detail. One of the most common Old English structures is **Subject + V (Object) (Adverb).** In this structure **X** can't be "empty"; it is "filled" by the two optional elements. Here is its tree:



The optional **NP** on the bottom line, which provides the direct object if there is to be one, can be followed by up to three prepositional phrases. Or **V** can be followed by S^1, a whole new sentence serving as a direct object, like our (18). The deep-

structure rules are then applied to the embedded S^1. The result is a sequence of grammatical elements. The derivation and character of each element give it intrinsic features.

Next we must insert words for the elements and use the applicable **T** rules. Suppose that the subject **N** on a given tree has these syntactic features: [+ common, + count, + concrete, + animate, + human, ± singular]. That is, only nouns from the lexicon that possess these features are eligible for insertion. So all verbs, prepositions, and other non-nouns are excluded. The noun must be common (not a proper noun or a personal pronoun, both of which are considered nouns in our analysis). It must be countable (not a mass noun that can't be made plural, like the modern noun *physics*). It must be concrete (not abstract, like *truth*) and animate (not inanimate, like *table*). It must be human and capable of having both a singular and a plural form.

Which words in the passage have these features? The dictionary can show us. Of course, the scribe had many thousands of nouns in his total lexicon from which to choose. His passage restricts us to twenty-one:

bigspell	hit	sǣd	þiccenes
ēare	lār	sǣdere	þorn
eorðe	land	scip	wæstm
fela	menigu	stānscyligean	weg
fugol	sǣ	sunne	wyrttruma
hē			

Which two do we reject because they are [− common]? The pronouns *hē* and *hit*. They're not [+ common] because they can't take adjectives the way nouns do. Thus we find "mycel menigu" and "gōd land," but not *"gōd hē."

The next step is to subtract all those that aren't countable. Note the difference between [+ count] and [+ singular]. Modern words like *scissors* and *trousers* are always plural; others like *athletics* and *mathematics,* always singular. Yet all four are [− count]. By contrast, *pencil* is [+ count] and can be [± singular]. To determine which of our nineteen Old English nouns have these features, we need to study many more passages. We must learn whether plural forms of *eorðe* and *þiccnesse* are used somewhere, or whether the words are always singular. The procedure is easy enough for [+ concrete, + animate, + human]. Which four

nouns are eliminated because they're [− concrete]? Which additional twelve are [− animate]? Which one is [− human]?

The survivors are *menigu* and *sǣdere*. Either can go with [+ human] verbs like *ācsian*. A fugol can fretan and gān, but can't lǣran. The scribe plugged in *menigu* for the subject **N** in (2) and (3). Remotely, *fugol* might fit, except that **N** has already been arbitrarily selected as [+ human]. The scribe used *sǣdere* in (5). *Fugol* isn't eligible, because it can't co-occur with "tō sāwenne." Birds may eat the grain, but they can't sow it.

Presumably the scribe was aware of such grammatical constraints. Extant manuscripts don't contain, for example, sentences in which a noun like *stān* co-occurs with *swimman*. That verb can be inserted only when the subject is animate. What if the subject is *oxa, fugol,* or *fisc?* Then the verb can be used, even though most oxen and birds aren't famous for that skill. Constraints, in short, were a part of the scribe's competence, whether or not he was conscious of them. Why do we have to hesitate before composing sentences like *"The dead baby lived to middle age"? If we don't hesitate, they may seem natural. If they do, alas, we should probably let our fingers do the walking to find a psychiatrist's telephone number.

Lexical Rules

In Figure IV the lexical rules were pictured as applying after insertion of the words. The process, not to mention the actual rules, still hasn't been wholly analyzed. There seems to be some overlapping with the syntactic features of the individual words. For example, the adjective *grēne* may have the feature [+ color] in its syntactic rather than dictionary listing. Similarly, *bigspell* has [− color]. Yet meaning apparently has some role in preventing the equivalent of *"green parable." Semantically, color words would seem to go only with a [+ concrete] noun like *house*. Today *green* also co-occurs with [− concrete] nouns like *color* and *tint,* but not with *picnic* and *truth*. In general, lexical rules provide the final directions needed for choosing the particular meaning of a word in a sequence, out of all its possible meanings. Such a rule finally selects "alone" as the meaning of *āna* in (19), but "one" in (16). Incidentally, *āna* wouldn't become the indefinite article for several more centuries.

When a lexical rule will directly affect a **T** rule, it is always

applied first. Thus "when" is chosen for *þā* in (6, 9, 19). The "then" determined in (12) requires **T-perm** to invert subject and verb, as we've observed. Again, **T-adj** must change the deep structure of (14), "And sum **past** feall on ðæt land – ðæt land **past** bēon good." The process deletes *ðæt* (the) in the main sentence and "ðæt land **past** bēon" in the embedded one. It moves the adjective into the vacated slot, producing the final surface structure of "And some fell on good land." Clearly, a lexical rule must first select a meaning for the adjective compatible with the noun *land.* If the base form *gōd* lacked a compatible meaning, the word couldn't be inserted into the embedded **N** + *bēon* + **A** structure ("the land was good") in the first place. The meaning selected for the adjective, for instance, can't be "honest" or "virtuous," since we don't have *"honest land." It must be something like "fertile" or "tillable."

Old English Lexicon

Two generalizations about the Old English lexicon deserve mention. First, some words are clearly competing. One word eventually wins out when its competitors are no longer used and become obsolete, or else shift their meanings. Our dictionary can again provide examples. *Bigspell* is viciously defeated by Latin *parabola.* How do we know? In our passage *cwæð* occurs twice, meaning "say." Actually, it's being vitally contested by Old English *secgan* (/sɛjɑn/), which becomes Middle English *sayen,* and our modern *say.* Poor *cweðan* gives us only the archaic, monotonous *quoth* for Poe's raven. Both verbs are Germanic descendants, neither being a suave Latin or a malicious Norse type.

Eft and *æfter* compete. Which wins? *Fretan* and *etan* joust, the modern *fret* taking a different meaning. What is the victor's modern spelling? The same is true for *fugol* and *brid,* giving us *fowl* and *bird.* Our scribe chose the loser, *fugol. Hundfeald,* after losing the meaning "hundred" to Old Norse *hundrath,* retains "hundredfold." *Lār* becomes *lore. Læran* is replaced by *leornian,* which loses out to *tæcan* in the modern context of what the teacher does to the student, but which describes what the student does. Can we resolve the foregoing riddle? *Micel* gives the chiefly Scottish *mickle,* as well as *much.* Part of its meaning is taken over by *grēat. Onginnan* loses out to *beginnan.* Of

stīgan only the British dialectal *styan* (rising) remains. What about confusing *þā?* One of its meanings is taken by *hwanne;* the other, by *þanne. Ðǣr* retains its "there" meaning. *Hwǣr* runs off with its other meaning. *Wæstm, ymbe,* and seven other words in our passage bite the dust. What are they? Their obsolescence explains why the now-familiar passage looked so foreign at our first shuddering glance. Those who don't know any German usually experience initial difficulty with Old English. The German speaker, however, knows cognates for a considerable number of the words in our passage.

A word may also seem strange because of the considerable shift of meaning over the centuries. An Old English word that looks familiar may denote something different today. Such changes have traditionally been divided into five kinds. A lexical rule may need widening in order to operate today, permitting the expansion or **generalization** of some words. There may be narrowing of a rule, for the modern **specialization.** We've mentioned the contest between *fugol* and *brid. Fugol* also meant "any feathered vertebrate animal," whereas *brid* was the general name of "the young of the feathered tribes." Today *fowl* is specialized, as in "wild fowl" or "fish and fowl," with *bird* generalized. *Dēor* (animal) similarly specialized to mean just "deer," after *animal* came in as the general word.

An Old English lexical rule may require alteration to convey a negative or moral judgment on a word today that did not have such a judgment in Old English. This process is called **degeneration.** Or a West Saxon judgment may be erased by the process called **elevation.** An example of degeneration is *cnapa,* which meant "serving boy or young man," whereas its descendant *knave* suggests a rogue or dishonest person. Elevation is illustrated by the changes of meaning in *cniht,* which could denote "knight" as well as a boy or servant. Our modern *knight* is someone honored by his sovereign. The fifth kind of lexical change is **splitting,** in which part of the meaning is shifted to one word, and another part to a second word. We've observed a similar process before, in the assumption by *hwanne* and *þanne* of two meanings of *þā.* Actually, in splitting, the words that are being split must derive from the same source. Thus Old English *strēawian,* which could mean "straw" or "strew," has become our noun *straw* and the verb *strew.* The modern verb doesn't denote "to straw." We must strew straw, if that's our desire.

The second lexical generalization is that prefixes, suffixes, and free forms (capable of standing alone, like *and*, as opposed to *mis-* and *-nes*) have combined with other elements to form thousands of new words. This capacity for combination was an important characteristic of Old English. Despite the obsolescence of the three verbs prefixed by *for-* in the passage, this prefix survives in verbs like *forbear, forbid, forget, forgo, forsake*. We do have *shrink* and *preshrunk* today, but not **forshrink*. Incidentally, because *shrink* begins with /š/, should we expect it to be of Scandinavian origin? Hundreds of words are formed with the prefixes *ge-* and *on-*. Which three words in our passage are formed this way? Unlike the obsolete *ymbe*, the modern descendants of *eal* and *ūp* continue to do well. For *all* we can list older words like *almighty, almost, already, also, altogether*. It's the father of recent combinations like *all-American, all-clear*, and *all-star*. For *up* there are *upbeat, upbraid, upcountry, upgrade, uphill*. We'll tiptoe around the graves of *ealceald* (all-cold), *ealcyn* (all-kind), *ūphūs* (upper chamber), *ūplang* (tall), *ymbgangan* (go round), and *ymblǣdan* (lead about). They rest in the large cemetery containing most Old English words.

Several highly productive prefixes don't occur in our passage, like *ā-, be-, fore-, mis-, of-, ofer-, tō-, un-, under-*, and *wiþ-*. A glance at a modern dictionary will prove the continuing utility of these native elements. Just as in modern times, addition of an Old English affix altered the meaning of the base. For instance, the free form *gān* is marvelously productive, combining with twenty prefixes to form other verbs:

agān (happen)	fulgān (fulfill)	togān (separate)
æfter- (follow after)	ge- (go)	under- (undermine)
be- (go over)	in- (enter)	ūp- (rise)
bi- (commit)	of- (require)	ūt- (exit)
for- (forgo)	ofer- (overrun)	wið- (go against)
fōre- (precede)	ōþ- (escape)	ymb- (go round)
forþ- (proceed)	þurh- (go through)	

Its combinations with suffixes provide modern words like *gang, gangster, goer* (from *gangere*), and *gangway*. Burns's "The best-laid schemes o' mice an' men / Gang aft agley" (go oft awry) might be "moused" in here.

We've mentioned the *-an* suffix, which provides the infinitive form of verbs like *lǣran*. Many are still productive. Consider the *-nes* attached to *þicce,* which usually forms feminine abstract nouns. Today it converts an adjective like *quiet* into what noun? We've observed *-feald* with numerals, giving us *twofold* all the way up to *billionfold* or beyond. The *-ere* (agent) of *sǣdere* is useful. Its descendant names those who *drink, eat, paint, read, swim, teach, work*. The *-en* (made of) of *stānscyligean* was once widely used with *gold, wood,* and other nouns denoting material. Now we're more likely to say "lead cup" than "leaden cup," or "silk pajamas" than "silken pajamas." Nine suffixes not in our passage are highly useful in their modern version. There are numerous combinations for *-dōm* (*kingdom*), *-full* (*eventful*), *-hād* (*childhood*), *-ing* (*scaffolding*), *-lēas* (*childless*), *-līce* (*eagerly*), *-scipe* (*friendship*), *-sum* (*handsome*), and *-wīs* (*lengthwise*). *-Wise* is particularly creative, to some people's disgust.

Old English free forms are just as productive as prefixes and suffixes. The resulting **compound** often provides us a view into the peoples' minds. As *gebēorscipe* means "feast," how did the lords often enjoy themselves? What may *dægcandel* and *eorþhūs* mean? Since *lǣcecræft* is "medicine," what did the doctor use instead of penicillin? This is a tricky question: since *lǣce* means "leech" as well as "doctor" (as we sometimes conclude upon seeing his bill), what is a *lǣcehūs?* Why is a *lārhūs* a school? *Mōdsēoc* is, literally, "mood-sick," just as *stīþmōd* is "stiff-mood" or "stern." *Mildheort* is equally obvious. *Medu* (*mead*) occurs frequently in extant manuscripts. From its several compounds we can list *medubenc, meduburh, medudrēam* (festivity), *medudrinc, medufull* (mead-cup), and *meduheall* (banquet hall).

Transformational Rules

In illustrating the deep-structure and lexical rules of Old English to this point, we've necessarily mentioned four **T** rules already familiar to us. We've seen **T-del,** which deletes the "understood" *gē* and repetitious nouns or verbs in certain coordinate structures. **T-infl** adds number to a noun and tense to a verb. **T-perm** is a general rule that permutes specified units in given structures, like a verb to a pre-subject position when *þā* (then) introduces the sentence. **T-adj** takes an adjective from its posi-

tion after *bēon* and permutes it before a [+ common] noun, discarding the repeated subject and *bēon*.

Now that we understand how the words are plugged into a deep structure and how they acquire their meanings there, a specific transformational example will be useful. The deep structure of (1) in our passage is "And hē **Tense** onginn—hē **Tense** lǣr hī—æt þǣre sǣ eft." A translation is "And he **Tense** begin **X** by the sea afterward," in which **X** is the embedded "He **Tense** teach them." We can't say *"began teach" today, just as our scribe couldn't. Both he and we could apply **T-ing** or **T-to.** For us, the transformation drops the duplicated "he **Tense**" in the embedded sentence, and adds either *-ing* to the embedded verb to make "began teaching," or else *to* just before it to make "began to teach." The latter is what happens, for (1) is like the modern structure minus *to*. *-An* does the work of *to* in converting the base *lǣr* to an infinitive. In fact, *to* later replaced *-an* in that function. **T-ing** and **T-to** have continued to gain importance and frequency of application up to the present because of the greatly more sophisticated auxiliary structures in the language. They prevent things like *"began teach" or *"begin taught." Some verbs, of course, won't permit both rules, as in "wanted to go" but not *"wanted going," although we have both "like to go" and "like going."

In altering the deep structure of (1) to a surface one, our scribe could have kept embedding an **S** indefinitely by means of $\mathbf{VP} \rightarrow \mathbf{V} + \mathbf{S}^1$. A modern example is "He said that I said that he said that I said . . . ." He did employ **T-adv** stylistically, by moving **eft,** the adverb of time, in front of the subject. As we've seen, **T-infl** arbitrarily makes **Tense** into **past.** The sequence thereby becomes "And eft hē **past** onginn lǣr + -an hī æt þǣre sǣ." Suppose we compare this structure, which is like the modern "And afterward he began to teach them by the sea," with (1). Clearly, **T-adv** must be applied again in order to invert the Old English adverb of place.

T-pro is also needed, so that the two pronouns will be close together. At this point we might well complain, "But why didn't the West Saxon stop with the 'modern' structure, instead of going on to create a foreign one?" Actually, **T-adv** and **T-pro** are optional in Old English. Personal pronouns are usually positioned just before or after the verb. If two or more occur, **T-pro** ordinarily moves one close to the other. We've seen "hē ongan hī"

in (1). There's also "hē hī . . . and him" in (4). For single occurrences, the pronoun subject precedes the verb, as was earlier observed. Today, we can't apply **T-pro** to a structure like (1). We still have the rule, and it still moves the pronoun forward in certain structures. We can't say *"He gave to me an apple" or *"He gave an apple me." Yet **T-pro** must change our deep structure *"The cop locked up him" to "The cop locked him up."

In how many structures in the Old English passage is an adverb of place moved forward? The answer is five: in (1), thrice in (2), and in "mid him wǣron" in (19). However, we find (in translation) "was on land" (3), "fell by the way" (6), "fell on stony ground" (8), "fell in thorns" (11), and so on. If many other passages are added to our little corpus, we'll discover the usual order to be **V + Place.** That is, **T-adv** is applied infrequently for adverbs of place, unlike other adverbs such as those of manner. The same is true today. "Gladly I went" and "I gladly went" are stylistic variations of "I went gladly," all of which have approximately the same meaning.

Three other **T** rules are frequently used. Though naturally altered over the centuries, two of these significantly differentiate Old English surface structure from ours today. **T-ne,** later replaced by **T-not** and sometimes followed by **T-do,** is one. The other is **T-agr,** which supplies complex agreement within the extensive inflectional system. For us it has become a simple rule supplying agreement within the few remaining inflections.

T-ne optionally adds *ne* to the deep structure. Thereby the negative sentence retains the same deep structure as the positive. For example, "hit wæstm ne bær" (13) results from application of **T-ne** to the syntactically almost identical "hit wæstm bær." Our "hē cwæð" (17) is almost identical with "hē ne cwæð." When the subject **N** is a personal pronoun, as is true in all four uses of *ne* in (8–10, 13), no inversion is necessary. If the subject is [+ common], however, **T-perm** usually permutes the *ne* and verb to pre-subject position.

Why are three of the four usages *næfde* rather than "ne hæfde"? **T-contr** makes the sequence into a contraction. *Habban* is then inflected by **T-infl.** So we find *næbbe, næfst, næfð* (I don't have, you don't have, he doesn't have), and so on. **T-contr** also combines *ne* with *willan* (wish), *bēon,* and *wesan.* Thus *nolde* (wouldn't), *neom* (amn't), *nis* (isn't), *næs* (wasn't), and *næron* (weren't). The rule is still applied to these words today. It ex-

plains our modal contractions like the preterits *couldn't, mightn't, shouldn't,* and *wouldn't,* as well as *mustn't.* It explains forms like *hasn't, haven't, isn't,* and *aren't. Amn't,* from *ne* + *eom,* probably helped develop our *ain't.* The speaker of Old English apparently used *neom* without any loss of dignity or prestige.

He also applied **T-contr** to *āgan* (*ought,* possess), giving *nāh* and *nāgon* (I don't possess, I didn't possess). From *witan* (know) we find *nāt, nāst, nāt,* and *nyton* (I don't know, you don't know, he doesn't know, we don't know). Otherwise, the rule doesn't apply. Structures are like "ne bær" and "ne brōhte," with no **nabær* and **nabrōhte.* As a result, we don't use **boren't* and **broughtn't* today. Historically, *ne* changed to *not* but kept its pre-verb position. An obligatory **T-do** developed to provide our "didn't bear" and "didn't bring."

T-contr isn't limited to modals and a few verbs. There are pronominal contractions like *nic* (*ne* + *ic,* not I) and *nān* (not one). *Nǣfre, nāhwǣr, nahwæðer, nāwiht,* and *nāþing* are the predecessors of certain modern adverbial negatives. Which ones? Some died, like *nāhwǣrn, nāthwæt* (something unknown), and *nāthwilc* (someone I know not who). If a person uses *nohow, nowhat,* or *nowheres* today, he might argue that he's intuitively employing a rich historical compound. After all, he's not to blame for the severe degeneration of these three, in contrast to the continuing respectability of the others!

Old English Nouns

Unlike **T-ne, T-agr** is rather general. It's probably the most complicated of all Old English rules. It supplies the agreement required by the noun for certain other kinds of words co-occurring in the sequence, requirements that have mainly vanished from English. The noun seems to be central to the deep structure. When a particular noun is plugged in, it carries features of gender and case. **T-infl** provides its number. **T-agr** then gives matching, relevant features to the co-occurring form classes. Before we can really understand this rule, we need certain information about Old English nouns. Consider *stān.* It must take one case feature from [+ nom, + gen, + dat, + acc], depending upon the deep structure. If the **N** is in subject position, **nominative** must be chosen. If it's possessive or a few other things, **genitive** is the

case. The noun is **accusative** if it's the direct object, and **dative** if it is the indirect object or follows most prepositions. As *stān* is [± sing], **T-infl** arbitrarily selects one of them, since, in the abstract, the base *stān* can be either singular or plural, the latter by later addition of the plural suffix.

In addition, each noun has its own grammatical gender. Regardless of the biological accuracy, *stone* is always masculine, *year* is neuter, and *lore* is feminine. Today, of course, "natural" gender operates. A personal pronoun referring to any of these three nouns must be *it*. The word *king* takes *he*. *Student* may be *he* or *she*, hopefully not *it*, although we hear stories of precocious chimpanzees. At this point it might be helpful to see the whole declension of *stān*, together with those of four more subclasses. A few other subclasses are too small to bother with.

	stān (*m.*)	**nama** (*m.*)	**fōt** (*m.*)	**gēar** (*n.*)	**lār** (*f.*)
sing. nom.	stān	nama	fōt	gēar	lār
gen.	stānes	naman	fōtes	gēares	lāre
dat.	stāne	naman	fēt	gēare	lāre
acc.	stān	naman	fōt	gēar	lāre
plur. nom.	stānas	naman	fēt	gēar	lāra
gen.	stāna	namena	fōta	gēara	lāra
dat.	stānum	namum	fōtum	gēarum	lārum
acc.	stānas	naman	fēt	gēar	lāra

The *stān* type accounts for about a third of the nouns; the *nama* type, a tenth. Little by little, *naman, tungan, ēagan (eye),* and many others have adopted the *-s* plural of the *stān* type. Only *oxen* and *children* retain *-en* today, with *brethren* fading. The other three subclasses in our list include considerably fewer nouns. The *fōt* type explains why we use *feet* instead of **foots*. Indeed, the Old English inflection explains irregular plurals like *lice, men, children,* and *knives*. Similar preservation of Latin plural forms gives us *crises, dicta,* and *indices*. However, the great majority of modern nouns have a simplified *stān* inflection, with grammatical gender lost. That is, the noun form is the same for the former nominative, dative, and accusative functions. Thus the word *boy* is unchanged in these positions: "The boy is polite," "Jack gave the boy money," and "He saw the boy."

Of course, genitive case remains, with the singular and plural sounding alike. In writing, *boy's* versus *boys'* differentiates the number. As for the plural form, **T-infl** adds one of the following: /s, z, əz/, as in *ships, boys,* and *judges.*

Let's use *sǣ* in (1) of our passage to illustrate the application of **T-agr** to the determiner, adjective, personal pronoun, and verb, in that order. *Sǣ* is [+ fem, + dat, + sing]. The base determiner *sē,* which will precede its noun, must gain the same features from **T-agr,** as selected from over three dozen forms. Some of these have several variations within the same inflection. A **P** rule later converts the feminine form *sēo* to *þǣre,* the form actually chosen by the scribe. Any other form is as unacceptable as *"she himself" is today. In general, most Old English determiners have telescoped into *ðē,* which gives *the,* with all trace of number lost. The indefinite *a(n)* develops from the numeral *ān.* Another determiner form, *ðis,* gives *this,* with the plural *ðēs* developing in Middle English. *Ðæt,* already contesting with *ðē,* develops into *that,* with *ðōs* becoming the plural form. Today there is a clear contrast among *the-this-that. This-these* and *that-those* require agreement in number. We don't say *"this girls" or *"those boy."

We can use the "gōd land" of (14) to illustrate **T-agr** for the adjective. As has been said, *gōd* originates in the deep structure **N** + *bēon* + **Adj.** It must gain the matching inflectional features of its subject noun if it remains there. However, the adjective usually precedes its noun. So **T-adj** moves it in front of *land,* which is [+ neut, + acc, + sing]. **T-agr** and a later **P** rule give *gōd,* the final form. The process of adjective agreement looks simple. But once the degrees are added, there are almost two hundred possible forms, of which only *-er* and *-est* survive. Why don't the modern *good, little, many,* and *bad* fit the *-er* and *-est* pattern, giving us **gooder?* Their Old English forms can answer:

gōd-betera-betst	lytel-læssa-læst
yfel-wyrsa-wyrresta	micel-māra-mæst

In Middle English the word *badde* appropriates the "base" position from *yfel,* by then changed to *ēvel.* Today, of course, *evil* requires *more* and *most* to make the comparative and superlative, as opposed to **eviler.*

Unlike the determiner and adjective, the personal pronoun

may receive its case feature because of its own deep structure. If inserted for a subject **N** in deep structure, it is [+ nom]. Elsewhere, it depends upon the noun for inflectional information, like its two colleagues. The forms are somewhat more complex because person must be indicated, too. There is a dual number, which means "two only."

	first person	**second person**	**third person**		
			masc.	*neut.*	*fem.*
sing. nom.	ic	ðū	hē	hit	hēo
gen.	mīn	ðīn	his	his	hire
dat.	mē	ðē	him	him	hire
acc.	mē	ðē	hine	hit	hīe
dual nom.	wit (we two)	git (you two)			
gen.	uncer	incer			
dat.	unc	inc			
acc.	unc	inc			
plur.	wē	gē	hīe	[no gender	
	ūre	ēower	hira	designated	
	ūs	ēow	him	for the	
	ūs	ēow	hīe	plurals]	

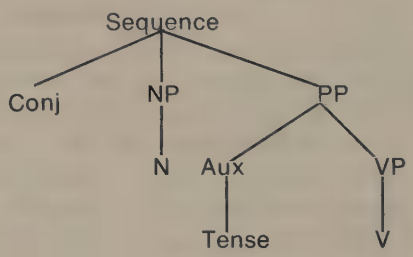
Consider "Ūt ēode sē sǣdere his sǣd tō sāwenne" (5). As *his* clearly refers to *sǣdere,* we must know the inflectional features of that noun. It's [+ masc], and its subject position in deep structure requires [+ nom]. As it's [+ third], we can rule out dual number, which has only first-person and second-person forms. Anyway, if the number were dual, the feature must be [+ dual]. **T-infl** makes *sǣdere* [+ sing]. When we look in our list for a pronoun with [+ masc, + nom, + third, + sing], the answer is *hē.* However, the scribe used the modern *his,* just as we say "his seed," not *"he seed." The pronoun needed actually modifies the neuter *sǣd,* which requires it to have [+ gen] rather than [+ nom]. Once this case feature from *sǣd* is added to the [+ *masc,* + third, + sing] determined by *sǣdere,* **T-agr** must select *his.*

Except for the vanished dual number, the pronoun has retained most of its inflections. Dative and accusative have merged into one case, usually called **objective.** The first-person singular

pronouns haven't changed much. *My* and *mine* are differentiated, as in "my hat" and "hat of mine." The merging of the second-person singulars into *you* is easy to see, with the biblical *thine* and *thee* still around. *Thou* is used by some Quakers when addressing one person. For third person, the dative replaces the accusative. The *it* forms lose initial /h/ except in dialectal pronunciation. The *hēo* forms become purely feminine, with *hers* developing. The plural pronouns are actually more complex today. The added /z/ gives *ours, yours, theirs.* Loss of second-person plural is sometimes compensated for by the dialectal *you-uns, you-all,* and *youse.* As has been said, the four third-person plural "natives" have been replaced by Scandinavian loanwords.

As for verbs, **T-agr** can't be as neatly illustrated as it can for determiners, adjectives, and pronouns. After all, verbs are parallel with nouns in the deep-structure rule **(Conj) NP + PP.** In some ways they may be as central to deep structure as nouns are. By contrast, determiners come subordinately from **NP.** Though deriving from **PP,** adjectives are equally dependent on the noun for inflectional information. Except for case in certain structures, pronouns are just as dependent. The noun subject does determine the person and number of verbs, together with certain other features.

"And fugelas cōmon" (7) can illustrate. The subject is [+ animate]. Thus the word inserted for **V** must be [+ animate], as are *cuman* and many other verbs:



As this deep structure has no direct object, the verb must be [− transitive]. The noun subject is [+ third, − sing]. **T-agr** gives these two matching features to *cuman* after its insertion. Here the direct influence of the noun stops.

The scribe made an arbitrary decision before plugging in *cuman*. Remember the rule **Aux → Tense (Modal)?** He decided against a modal for (7). Deep structure thereby dictates the affixing of **Tense** to the verb rather than to a possible modal. In modern terms, the difference can be seen in *should come* as opposed to *came*. **T-infl** makes *cuman* [− pres], but the deep structure determines the **mood** of the verb. The mood clearly isn't subjunctive. As the subject isn't the second-person *ðū* or *gē*, the mood can't be imperative. It's indicative. Thus four features collectively determine the final form of *cuman:* [+ third, − sing, − pres, + ind]. A **P** rule later supplies *cōmon,* or /kómɑn/, the scribe's choice.

Vocalic Change

Cōmon is one form of the base *cum-*, once we deduct the *-an* suffix marking the infinitive and then add the *-on* suffix. Note some other verbs in the passage that have undergone vowel, or vocalic change: *bær, cwæð, fēoll, forscranc, frǣton, ongan, sēow, stigon.* Only *cōmon, frǣton,* and *stigon* have anything added. The *-on* marks the third-person plural preterit.

Now consider *āxodon, brōhte,* and *sealde*. Again there is vocalic change, with addition. Part of the addition is a dental consonant /t,d/, unlike the *-on*. The modern forms of these verbs retain it, as in *ask-asked, bring-brought,* and *sell-sold,* as opposed to *come-came*. Let's list the other four with /t,d/: *forswǣlde, forðrysmodon, gegaderod,* and *lǣrde*. Our dictionary shows that their bases don't undergo vocalic change but do add the dental. Which two have *-ian?* Actually, *-i-* belongs more to the base, so that *-an* may be considered as their suffix too. All seven of these verb forms are preterit; all have a dental suffix. *Cōmon* and its eight colleagues are preterit, but not one has the dental suffix. In short, the capacity to take /t,d/ characterizes one of the two major subclasses of Old English verbs. We've seen that the other operates by a vocalic change. The same difference is found in the participle as well.

The vocalic subclass, containing more than three hundred verbs, has seven subsets. Here are two representative examples of each:

	base	pret. sing.	pret. plur.	participle
1.	glīd-	glād	glidon	gliden
	rīd-	rād	ridon	riden
2.	crēop-	crēap	crupon	cropen
	flēog-	flēag	flugon	flogen
3.	melt-	mealt	multon	molten
	scrinc-	scranc	scruncon	scruncen
4.	cum-	cōm	cōmon	cumen
	scer-	scær	scǣron	scoren
5.	et-	ǣt	ǣton	eten
	fret-	frǣt	frǣton	freten
6.	bac-	bōc	bōcon	bacen
	weax-	wōx	wōxon	weaxen
7.	feall-	fēoll	fēollon	feallen
	slǣp-	slēp	slēpon	slǣpen

From the list we can make two observations. First, the preterit plural, marked by *-on,* has vanished from the language. What historical argument does a person have for saying *"We ridden yesterday," instead of "We rode yesterday"? Second, eight of the above verbs are dentals today—*glided, crept, melted, sheared, fretted, baked, waxed, slept.* Such is the history of two-thirds of the vocalics, greatly outnumbered by the dentals from the first. This alteration in the inflected forms of many verbs explains why the verb has changed more than any other form class.

As for the other six vocalics in the above list, the base vowel changes in the past singular and participle:

rīd-	/i ⟶ ɑ, ɪ/	cum-	/ʊ ⟶ o:, ʊ/
flēog-	/ɛ: ⟶ æ:, o/	et-	/ɛ ⟶ æ:, ɛ/
scrinc-	/ɪ ⟶ ɑ, ʊ/	feall-	/æ: ⟶ ɛ:, æ/

The change explains their irregularity today. We might note that the ex-diphthongs *ea* and *eo,* represented by *flēogan* and *feallan,* were monophthongs by the eleventh century.

Dental Addition

Three dental subsets are usually posed:

1. gehȳr-	gehȳrde	gehȳrdon	gehȳred

2. gegader-	gegaderode	gegaderodon	gegaderod
3. habb-	hæfde	hæfdon	gehæfd

The *gehȳran* type contains by far the most verbs. Many of the second type have become obsolete, with many remaining. Of the four members of the last type, the high-frequency *habban, libban* (*live*), and *secgan* (*say*) have survived.

From this representative list, several observations about dentals can be made. The preterit plural, marked by *-on,* has vanished, as has the *ge-* prefix. Most don't undergo vowel change, unlike all of the vocalics. The Old English dental paradigms help explain the modern forms—*hear-heard-heard, gather-gathered-gathered,* and *have-had-had.* Moreover, the preterit and past participle have become homonyms, unlike the vocalic pattern, as in *drank-drunk.* Resisting the trend, a few dentals have become vocalic, like *dig* and *fling. Dove* is often used today as the preterit of *dive,* originally a dental. It retains the *-ed* participle.

How do we classify modern verbs like *burst, cut,* and *rid,* which have no vowel change or suffix? For consistency, we'll say that they have a "zero" suffix, **ø.** Their base is unchanged but might have been, and **ø** can serve as a useful "space-holder." These "zero" verbs have two characteristics: each is monosyllabic and ends in /t/ or /d/. Note the paradigm of the dental verb *settan: sette, setton, setted.* After the *-ed* was eventually lost, *set* was left with no trace of its dental bloodline. Let's not confuse it with the vocalic *sittan,* which has a paradigm of *sæt, sǣton, seten.* Has this non-upstanding verb remained vocalic, enlisted in the dental ranks, or rejected both for **ø**? A related current problem is *lie-lay,* again probably resulting because the two are near-homonyms. Actually, the Old English speaker had to differentiate the paradigms of *licgan* (*læg, lǣgon, legen*) and *lecgan* (*legde, leidon, leid*). Both verbs have several variations for each form. We should sympathize with him, not blame him, if today we carelessly say *"He had just laid down." Overall, our "zero" verbs once belonged either to the vocalic subclass, like *berstan,* or to the dental one, like *settan;* but they've lost all trace of this ancestry.

A few verbs never belonged to either subclass. *Witan* is one; only *to wit* and *dimwitted* remain of the now-archaic verb. *Dōn, gān,* and *bēon* form another undisciplined band. The irregularity of *dō, dēð, dyde,* and *dōn* is to blame for our /dəz, dɪd, dən/, as

opposed to the comedian's **doed*. The preterit of otherwise nice *wendan* gets a piece of the action, with poor *ēode* dying after disinheritance from *gān*. So today we have *go-went-gone*. With true poetic justice, *wend* as an intransitive verb becomes archaic, losing out to *go*. *Wesan* stands convicted of the same crime. It supplies the preterits *wæs* and *wǣron* to the paradigm of *bēon*.

Bēon is especially interesting. Though never a modal, it can occur in auxiliary structures, as we saw in our old friend "wæs gegaderod." (Several problems in this passive transformation will be resolved in the next chapter.) Today *be* retains separate preterit forms for the singular and the plural (*was-were*), unlike all verbs and modals. Note that *be* is a copula, a linguistic form linking a subject to its predicate. *Be* is not a verb. That's why, in the Old English deep-structure rules, **VP** is rewritten as *bēon* or **V**. We may recall the replacement of the native *sindon* (from *bēon*) by the Scandinavian *aron*. Thus we say "they are" instead of *"they sin." *Be* is really our most mixed-up word. We say "I am, you are, he is," as opposed to *"I be, you be, he bes." The latter is the pattern of all verbs except a handful including *do, have,* and *say*. Their vowels change in the third-person singular present, giving /dəz, hæz, sɛz/.

We may remember that Scandinavian is usually credited with furnishing our verbs with the *-ing* participle and *-s*. Both the vocalic and the dental subclasses have *-ende,* the present participle and origin of the modern *-ing*. The *-ende* is frequently used, as in "ūppstīgende and wexende wæstm" in our passage. The *-s* is of special interest because it's the only inflection preserved in the whole present tense. The Old English speaker had to separate "ic hæbbe, ðū hæfst, hē hæfþ, wē habbaþ." A closer look at this sample paradigm reveals why the modern third-person singular is *has,* instead of **haves,* on the model of other verbs. The replacement of *-eth* by *-s* took centuries. Indeed, *-eth* still occurs in archaic forms and in poetic or biblical texts. As late as the King James Bible (1611), there are "unleveled" forms like the second-person *art* and *wast*. Contractions like *doth* and *hath* are occasionally used.

Six modals form a neatly irregular group: *cunnan, durran, magan* (*may*), *mōtan* (*must*), *sceolan, willan*. Not verbs, they provide the basis for the modern modals. The structures in which they occur are not as elaborate or perhaps even as systematic as ours today. Some of them have changed substantially, as in

Cædmon's reply to his king when asked to sing. He says, "Ne con ic nōht singan" (not can I not sing). Here we see the frequent double (and sometimes triple) negative. *Con* is the modal for the infinitive *singan*. **T-ing** and **T-to** can't apply to Old English modals, nor can they be applied today. *"I can't swimming" and *"I can't to swim" are unacceptable.

The number of pages devoted to **T-agr** may have indicated its complexity. It gives matching inflectional features of nouns to determiners, adjectives, personal pronouns, and verbs. Thereby it helps make Old English a relatively inflective language. Our passage has demonstrated another fact: word order is comparatively free. Also, the *-an* suffix and the dative case reduce the need for connecting words like prepositions. We've seen the considerable separation of *ongan* from *lǣran* in that verb sequence in (1). There, *-an* takes the place of *tō* as in the modern "began to teach." "And hē sǣde him" is the structure following the end of our passage. Again, *tō* isn't needed, because the dative *him* means "to him or them." Yet we find *tō* in (2, 4) and so on. There are other prepositions, such as *æf (of), æt, fram, mid, ofer, on,* the loanword *til, wið,* and *ymbe*. Naturally they're uninflected, as are the other connecting words, or conjunctions—*and, būtan* (*but*), *fōr, swā, ðēah*. They're all simply less important syntactically than they are today because we don't have dative case to provide needed syntactic information.

A word should be said about adverbs, most of which are formed from adjectives by the addition of *-e* to the base. Among many, here are some examples: *hlūde* (*loud*), *lange* (*long*), and *yfele*. As the purists may not know, loss of this *-e* in Middle English leaves some adjectives and adverbs as homonyms. Consider *fæst, riht,* and *strang*. Today we say "fast driver—drive fast," "right guess—guess right," but not *"strong ruler—rule strong." The last usage probably isn't accepted because *strange* and *stranglīc* competed as adverbs in Old English, and the latter won. Had there been no *stranglīc,* we would probably be using *strong* adverbially. Addition of *-līc* (body) to adjectives creates increasing numbers of adverbs by Late West Saxon, as in *glæd-glædlīce* and *slāw-slāwlīce*. Actually, there was already a practice of derivation of new members of form classes from others. A modern example is the noun *act* and the verb *act,* as well as *active, actively, activeness, action,* and *activity*. The process is quite productive today. An Old English illustration is

dēop (a.), *dēop* (n.), *dēope* (adv.), *dēoplīc* (a.), *dēoplīce* (adv.), *dēopnes* (n.). Some adverbs have no such "cousins"—*eft, hēr, sōna, ðǣr, ūp.*

Over the past several pages we've summarized key transformational rules, in the process of which four form classes have been discussed. Prepositions, conjunctions, and adverbs have been sketched. What remains is the phonological component, already partly described. Applied last in the generative-transformational process, **P** rules supply the proper pronunciation or spelling of the sequence. They can also give us mispronunciation or misspelling with no trouble at all. Since it's the sounds of the language that someone hears, the rules governing them are important to his interpretation of what he hears. We'll only summarize, not detail, the complicated formulas actually composing them.

Old English Phonological Rules

Naturally the rules provide all the phonemes and all their possible distributions. Constraints prevent phonemic combinations not in the language. For example, as Old English clusters are usually composed of two consonants, constraints prevent /skl-/, as in our *sclerosis*. On the other hand, we find numerous predecessors of the modern /r/ clusters. Examples are *scrincan* (/š/), *prāfost* (*provost*), *brōðor,* and *grēat. Fnēsan,* with the rare /fn/, gives us *sneeze,* apparently because people misread the *f* as *s*.

Clusters with initial /k/ are common, as in *cræft* and *crop. Cwudu* has been respelled as *quid,* one of the few /kw-/ survivors from Old English. What's the modern version of *cwellan* (kill)? Sometimes the /k/ has been lost. Yet *cniht* and *cnēow* have been respelled as *kn,* perturbing those who "rite about nights on their nees before their ladies."

Several two-consonant clusters have lost their initial consonant in like manner. Among the victims are /wr, wl, gn/ as well as the /h/ in four others. *Wrist, wrītan, wlisp,* several obsolete compounds formed with *wlite-* (looks), *gnæt,* and *gnagan* (*gnaw*) will illustrate. Which one of the six words has been respelled to indicate the leveling? Of these former clusters, *wr-* items are the most numerous today. Initial /h/ clusters are especially productive, such as /hw, hl, hn, hr/. We've noted *hwanne,*

hwǣr, and *hwǣt,* unphonetically respelled *wh-* and often pronounced with /w-/ when the word isn't under stress. When stressed, it retains the /hw/. The stressing constitutes a kind of test, if we know the Old English spelling. *Whole* (from *hal*) fails, since it never had /hw/. What about *whisper?* In any case, /hw/ remains with us today. *Hlēapan* and *hlehhan* provide *leap* and *laugh.* Of /hn/, *hnecca* and *hnutu* give us *neck* and *nut. Ring* and *roof* come from *hring* and *hrōf.*

The loss need not be from a cluster. Many Old English spellings indicate that the words have lost one or more phonemes over the centuries. Three kinds of omission are usually described —initial, medial, and terminal. *Hit-it* illustrates **aphaeresis** /əférəsɪs/. *Betst-best,* which we've seen earlier, shows us **syncope.** *Prættig-pretty,* where the terminal /g/ is gone, exemplifies **apocope.** It also illustrates elevation, since the Old English meaning is "wily or crafty." Terminal omission can be extreme, as in clippings like *gym* and *zoo,* from *gymnasium* and *zoological garden.*

Some phonemic change doesn't involve a loss, but only an inversion, or **metathesis.** For instance, *brid* becomes *bird,* with a change of vowel. *Hros, wlips,* and *ācsian* give us *horse, lisp,* and *ask,* respectively. If we experience a slip of the tongue, a jocular explanation might be that our brain was intuitively reconstructing an earlier stage. Many of our slips are of this nature. The unstressed /pərti/, as in "pretty good," may have had such an origin.

A word may also gain a phoneme. Curiously, a once-omitted sound can later be replaced. The word *often,* its spelling containing its autobiography, is typical. The Old English *oft* is added to, for *often.* Syncope takes away the /t/. Today, people sometimes self-consciously reinsert the /t/ in speech. A phoneme may be added to make a sequence easier to pronounce. Thus *ganra* and *brǣmel* are now *gander* and *bramble.* There may be a careless addition, as in *guma* and *vermin,* which give *groom* and the dialectal *varmint.* The process of phonemic loss and gain continues through the centuries in every language. The central kind of addition in Old English is compounding, observed in words like *wiðstandan.* Overall, **P** rules create these original phonological sequences and permit historical alteration of the sequences.

Our description of English up to A.D. 1100 has been quite long. However, the aspects of Late West Saxon that have been

summarized will appear in Middle English as old friends, sometimes senile and on the way toward death. The number of Latin and Scandinavian borrowings will grow, combining with a large French influx to squeeze out additional native words. Some transformational, lexical, and phonological rules will undergo simplification, particularly **T-agr** as many inflections are leveled. Few rules will be added. The necessary length and complexity of this chapter will greatly simplify what follows in the rest of the book, on the way to the 1970s. Middle English will look extremely modern by comparison with Old English.

Activities

1. Why do we expect the modern dialects of English speakers north of the Humber to be quite a bit different from those south of the Thames? Why does the vocabulary of this more northern group contain many words not used by people south of the Thames?
2. Collect some of the stories about Alfred.
3. List three Latin loanwords that don't have the English plural form, such as *cacti*.
4. Write the modern spelling of each of the following Latin loanwords in Old English. Remember the pronunciation of *c*.

cealc	pipor
cēse	plūme
cytel	Sæternesdæg
mangere	sicul
mynyt	weall

5. Take a map of England and find two place-names formed from each of the following: *-ceastor* (e.g., *Lancaster*), *port* (*Portland*), *tor* (*Torre Abbey*), and *-wich* (*Greenwich*). Do the same with a map of Lincolnshire or Yorkshire, this time locating Scandinavian borrowings with *by-, thorpe, -thwaite,* and *toft*. Note that some have variant spellings. Form four words like *bypass* from the prefix *by-*.
6. In a good modern dictionary find nine Scandinavian loanwords containing /sk/, /k/, and /g/. Repeat the

exercise for non-Scandinavian English words with /s/, /š/, /c/, and /j/. Remember that the *sc-*, *sh-*, *sk-*, *c-*, *ch-*, *g-*, and *k-* spellings often indicate the origin of the word.

7. Find Modern German cognates for the following words in the dictionary accompanying Text A:

gegaderian	gōd	lār	sǣd	weg
gehȳran	land	ofer	ūp	

List the nine words in the dictionary that have become obsolete.

8. Determine the modern spelling of the following Old English words. How has the meaning of each word generalized, specialized, degenerated, or elevated? Here's some help: there are three examples of each.

bereærn (barley place)
corn (grain)
dōm (legal judgment)
eorl (man)
girle (youngster)
godsibb (sponsor, godparent)
lǣwede (not of the laity)
nice (foolish)
salarie (salt money)
steorfan (die)
stigweard (one who manages the table)
tægl (horse's caudal appendage)

9. Find three modern words containing each of these prefixes (e.g., *ago*) and suffixes (*earldom*):

a-	be-	fore-	mis-	off-
over-	to-	un-	under-	with-
-dom	-ful	-hood	-ing	-less
-ly	-ness	-ship	-some	-wise

10. Before application of **T-adj,** the adjective to be inserted must match certain features of the noun it is to modify. Consider the modern noun *boy,* which is [+ concrete, + animate, + human]. Which features do the adjectives *misty, wooden,* and *girlish* lack that prevent their insertion?
11. Present-Day English isn't inflected for gender. When gender is important, pronoun forms can suggest it, as in

"The student called the girl, and he broke their date." Though *student* has no inherent gender, the masculine pronoun indicates *student* to be [+ masc] here. The pronoun modifying *date* is [+ plur], combining him and her. A few nouns have natural gender, like the [+ masc] of *priest,* or the [+ fem] of *priestess*. List three pairs like *uncle-aunt*. List three nouns that are made feminine through addition of *-ess*.

12. Find three modern nouns preserving the plural form of the *foot* type, three of the ø type, and three of each of the regular /-s, -z, -əz/ types.
13. List the thirteen different determiner forms in Text A (p. 63). Don't forget that cardinal numbers are determiners.
14. Attempt to justify use of the dialectal "you-all" or "youse guys." Why is a structure like "you three people" useful, considering the loss of the plural form for *you?*
15. Find three verbs that still preserve forms of each of the seven vocalic subsets. List three "zero" verbs.
16. From a linguistic point of view, why do we sometimes use structures like "drive slow," upsetting purists who insist on "drive slowly"?
17. A modern spelling may indicate a possible earlier cluster. Collegiate dictionaries with a good etymological section can supply the Old English spelling and confirm the supposed cluster. Otherwise, there's no easy way to discover the loss. Use a dictionary to try to find one or more examples of each of these: /l/ from /hl/, /n/ from /hn/, and /r/ from /hr/. Locating two examples of other "clusters" is easy because of the spelling: /n/ from /kn/, /n/ from /gn/, /r/ from /wr/, and retained /kw, hw/.
18. Which of the following words have undergone metathesis, aphaeresis, syncope, apocope, and/or compounding?

crudan-curd	ēare-ear	ne wolde-nolde
crul-curl	esquyer-squire	noumpire-umpire
ēage-eye-/aɪ/	napron-/épərn/	stānas-/stonz/

Chapter Six

Middle English (1100–1500)

Any broad span of time is open to a variety of chronological divisions. Is the Middle English span best divided by political events? If so, we should probably begin with 1066, the Norman Conquest, and end the first period at 1204, when Normandy was lost to the English king. We might stop the second period at 1337, the beginning of the Hundred Years' War. The last period might end in 1485, when Henry VII slew Richard III at Bosworth Field and initiated the Tudor line.

However, a linguistic history is best revealed by linguistic developments, which can only be roughly approximated. Thus 1100–1250 will be designated as Early Middle English, when only the nonruling classes used English. The Norman influence hadn't really been felt upon the language. The ruling classes spoke and wrote Anglo-Norman. The period 1250–1350 covers the time when English was reasserting itself throughout the country. The dominance of the language characterizes 1350–1500, or Late Middle English. As we'll see, political events influenced each of these arbitrary divisions, so that the political dates given above will be useful. By 1500 we'll be noticing the rather modern look of our language, for it changed more in the preceding four centuries than during any other stage. In time and magnitude of change, it was *middle,* between the old and the modern.

Norman Invasion and History to 1250

Any discussion of the period 1100–1250 requires information about the Normans who did the notorious conquesting.

Though their language and English were mutually unintelligible, the Normans were related to the Vikings who began ravaging England in 787. Unlike the Vikings, who took the dangerous boat ride across the North Sea, the Normans' ancestors moved handily down the European coast to the Frankish kingdom, across the Channel from Britain. There the "Norþmann" established himself in the tenth century. The word is in quotation marks to emphasize the syncope, which helps give us *Norman.* The invaders based their kingdom upon their great military power. Intensely enjoying the French culture, they adopted it. Generally forgetting their own language, they adopted the local French language too, to which they added many Scandinavian words. Actually, the dialect that William and his nobles spoke in England, once they conquered it, should be termed **Anglo-Norman.** For the sake of economy, we'll call it French anyway.

The Normans seldom missed an opportunity to take someone else's land. They were fully aware of the prize seized in England by their former colleagues, the Vikings, who had been occasionally rebuffed by men like Alfred and his inhospitable Danelaw. After the Danish Harthacanute died in 1042 and the English kingship returned to Alfred's line, William, the Duke of Normandy, watched developments across the Channel with rising ambition. Edward the Confessor died childless after twenty-four years on the English throne. The witan, or king's advisory council (remember the verb, "to know"), chose Harold as the successor. The Earl of Wessex and the King's chief adviser, Harold was nonetheless not of royal blood. William was furious; however, he couldn't be too legalistic about ancestry because of his own illegitimacy. Actually a second cousin of Edward, William loudly claimed the throne. The Pope sent him a consecrated banner, showing church support. The Norwegian King ineptly tried to assist William, but was killed and his army routed by Harold at Stamford Bridge.

William landed near Hastings three days later, probably with less than fifteen thousand men. For a time the English used a wall of shields to defend themselves against William's cavalry and infantry. Their defeat partly came about through the Norman strategy of shooting arrows into the air, which fell vertically upon them. Harold, pierced in the eye by an arrow, died instantly. William was crowned on Christmas Day. Over the next four years he ruthlessly crushed periodic resistance. One of the

punishments he devised was to lay waste the area between York and Durham, where, interestingly enough, many Danish descendants lived. Almost all lands were confiscated by William from the English who had resisted and were given to his Norman lords as military rewards. The English nobles, in fact, were almost wiped out. A few survivors retained their estates by vowing allegiance. Needless to say, they soon learned French.

William changed the government to conform to the much more centralized Norman practice. Church backing was insured by gradual replacement of the highest English prelates with Normans. He imposed feudalism on the country. Thereby landowners promised military and other services in return for feudal tenure. Except for the change of boss, most Englishmen weren't greatly affected. Living in hovels as tenants on the large estates, they found the work no easier than it had been under English overlords. Undoubtedly the serfs on the more isolated estates had a linguistic effect on their Norman lord. His prosperity derived from their toil, and they didn't know his language. He must have been forced to learn some of their language to give orders. In essence, the country had become Norman.

William and his lords couldn't understand English and apparently were uninterested in learning it. This instituting of French as the language of the ruling class for two centuries was by far the most important linguistic result of the Conquest. French was used in the court, in governmental activities, in military camps, and in the extensive commerce with France. The Norman nobles frequently visited their home estates. Soon the survivors of the English ruling class and the educated began to use the conquerors' tongue. They also wrote in Latin, but French was the literary medium. Romances, stories of knights and their deeds, were a Norman favorite, usually composed in French. Most Englishmen, however, kept speaking English, the only language they knew. They had no voice in governmental, religious, and military affairs. But the tongue didn't die.

Had England not lost Normandy, the future of our language might have been bleak. By 1205 only Aquitaine and part of Poitou remained of the Continental holdings of John, who, after signing the Magna Carta, ruled until 1216. In 1244 the French King and Henry III of England decreed that it was illegal for anyone to own land in both countries. Thereby the English ruling class lost their double allegiance. The separation presumably

speeded what might be called the return to English. Moreover, the French monarchs continued efforts to regain the remaining "French" lands, particularly the rich duchy of Aquitaine.

Loanwords

With French still the official language of the ruling classes in 1250, this seems the appropriate place for discussion of loanwords. The borrowing is dramatically accelerated in these centuries because of the artificial imposition of a foreign tongue upon native English speakers. Thousands of Latin and French words are added. The Scandinavian borrowings into the northern English dialect continue to influence the other English dialects. As in Late West Saxon, the lexicon changes. French is a kind of massive substitute for the tiny Celtic borrowing in Old English. A major difference is the loss of much of the native *wordhord* (*word-hoard*). No longer needed, many Germanic words simply aren't used or are replaced. The borrowing again speeds the continued leveling of the numerous inflections.

Latin. The Latin influence upon English extends the story begun during the Continental period. We also observed the borrowings from 449 to the introduction of Christianity in 597, and from then to the Conquest. Not just "religious" terms, but also words relating to botany, clothing, education, foods, and other things were borrowed by 1066. Somewhat more than a hundred of these approximately 520 loanwords survive today in Standard English, a term used to describe the language of educated, cultivated members of the English speech community. Most of the words lost were replaced by the French items descended from them. The naturally close resemblance between the two sets of words makes it hard to decide whether a given Middle English term is directly from Latin or from French. Examples are *miserable* and *register*, which may have come from either. The Latin borrowings in Middle English are fewer than those of either the pre-Conquest or Modern English eras, especially the Renaissance. As in the latter, the words must come from writings rather than from contact with native speakers, who were hardly speaking Latin by then because of the substantial changes in Latin dialects into Italian and so on. Nevertheless, the items number in the hundreds. They relate to law, learning, religion, science, and

miscellaneous matters, as illustrated by the following modern derivatives:

client	allegory	bull	digit	accede
conviction	desk	collect	elixir	commit
corpus	index	lector	equator	conflict
gratis	item	limbo	ether	immortal
pauper	library	requiem	orbit	infect
subpoena	minor	sanctuary	recipe	infirm

Several dozen others enjoy comparable frequency today. Perhaps the most important individual additions are key word-forming elements. For example, some items in the above list come from combinations like *com* + *mittere*, *re* + *quies*, and *sub* + *poena*. English has naturalized the Latin prefixes *ex-*, *infra-*, *per-*, *re-*, and *sub-*, which compete in modern utility with native ones like *un-* and *under-* discussed in the fifth chapter. *Com-* has three forms, depending upon the context. It's *com-* before /p, b, m/ as in *-press*, *-bine*, and *-mute*. It's *col-* before /l/, as in *-lision*. It's *con-* before all other sounds, as in the /d, j, k, s/ of *-dense*, *-join*, *-cussion*, *-sent*. What are some other illustrations? The same thing has happened to Latin suffixes. From *-ate* we derive adjectives like *delicate* and *fortunate*. From another *-ate* suffix we get verbs like *alienate* and *radiate*. *Different* and *fluent* belong to a group of adjectives formed from verbs and what suffix? *Morose* and *verbose* utilize what suffix?

French. Some naturalized prefixes and suffixes derive partly from Latin, partly from French, a second source of borrowing in Middle English. The prefixes include *de-*, *dis-*, *in-*, *inter-*, *pre-*, and *pro-*. *In-* has several variations, as in *illustrate*, *immaterial*, and *irrational*. The suffixes include the adjective-forming *-ious*, as in *dubious* and *various*. The suffix *-tion* converts verbs to nouns, as in *conversation* and *definition*. French apparently gives us *ad-*, *-able*, and *-ment*. *Ad-* has numerous alternations, as in *accept*, *adjourn*, *affect*, *agree*, *allot*, *appall*, *assuage*, *attain*. Illustrations of the other two suffixes are *durable*, *miserable*, *argument*, *segment*.

It's impossible to decide whether Latin or French deserves credit for hybridization, a related development helping characterize English today. This is word-formation from elements of

originally different languages. Thus Old English verb derivatives have joined with the formerly French *-able* to create *answerable, bearable, eatable*. We've observed some English bases with the French feminine-forming *-ess*, such as *goddess* and *shepherdess*. There are fairly large numbers of the reverse, where the suffix is native and the base is borrowed. Let's illustrate with French bases and twelve familiar Old English suffixes:

dukedom	noblest	powerful	preacher
falsehood	preaching	colorless	princely
faintness	courtship	quarrelsome	costwise

The bulk of the many hundreds of French loanwords entered between 1251 and 1400.[9] Few if any can be positively dated before the Conquest. Among the very earliest we find *bacon, castle, market, prison,* and *service*. Not unpredictably, the words often relate to those areas where the Normans had the greatest influence upon England, if we overlook the killing of the nobility and the razing of some of the countryside. They made important governmental and church reforms, while stimulating intellectual and artistic activity. Here are typical modern derivatives:

arts: art, beauty, cloister, image, palace, pillar
foods: beef, dinner, mutton, pork, sauce, veal
government: crown, government, minister, parliament, state
law: court, felony, judge, justice, sue, traitor
militarism: army, battle, lance, peace, siege
rank: baron, duke, herald, peer, servant, vassal
religion: angel, baptize, preach, saint, sermon, virgin

Scandinavian Languages. A third source of borrowing is the Scandinavian languages. As has been indicated, most Danish and Norwegian loanwords first appear in Middle English manuscripts. Some of these were likely in use before 1100. Few of the Old Norse borrowings in Old English remain; some died out during the Norman occupation. Some survivors are *call, fellow, husband, law,* and *outlaw*. Because the dating of Scandinavian loanwords is so mixed up, the previous chapter may be referred

[9] Otto Jespersen, *Growth and Structure of the English Language* (New York: The Free Press, 1968; reprint of the ninth edition), p. 87.

to for typical examples in Old and Middle English when needed. We should recall their /sk, k, g/ characteristics, among others. We may also recall that our biblical translation in Text A uses the native Old English forms of the plural personal pronouns. Indeed, they tenaciously resist the attacking Scandinavian *they, their,* and *them* long after 1100. Even today people say "Get 'em," probably thinking the usage to be a contraction of "Get them." Actually, it's probably the last gasp of the native plurals. That is, *'em* is the contracted, dative plural *him.* In Middle English the Scandinavian *same* meets little resistance.

The Ormulum

The *Ormulum* (Orm's little book), a northeast midland manuscript of about 1200 that is now in the Oxford Bodleian Library, contains the first sizable number of Scandinavian words—about 120 of them, as well as twenty French borrowings. The *Ormulum* is of particular interest because Orm, a pious Augustan canon, composed this incomplete, twenty-thousand-word paraphrase of the gospels partly to reform and standardize spelling. As he doubles a consonant that follows a short vowel, his poem admirably indicates English pronunciation of the day. We'll later find that in the early fifteenth century the long vowels begin a long process called the Great Vowel Shift, whereby they are raised and moved toward the front of the mouth when produced. Orm's consonant-doubling confirms the vowels that are to shift and those that are draft-exempt. Consider his sentence "Onn Ennglissh wrītenn rihht þe word." The doubled consonants indicate a short vowel in *onn* (in), two short vowels in *Ennglissh,* and a short second one in *wrītenn.* The vowels in these words won't shift later. Not so for *write,* which he spells with a single consonant. Its /i/ noted in the fifth chapter will eventually become the diphthong /aɪ/. We must beware of his doubling in *rihht,* for the /i/ in *right* does finally go to /aɪ/, one of the few inconsistencies in his spelling system.

Before concluding our discussion of 1100–1250, we might look at a brief sample of the *Ormulum.* The poem exemplifies the tedious religious writing in Early Middle English. Religion was the usual subject of the few who chose to write in their native tongue instead of the prestigious French. The sample won't show us what most writing of the day looks like because Orm's opti-

mistic hopes for standardized spelling didn't materialize. Apparently his huge, repetitious book was never widely read or even copied. Hand copying, of course, was the only means of book publication in England until William Caxton introduced the printing press in 1476. Other writers didn't follow Orm's practice of doubling a consonant before a short vowel. Now for the sample from his Preface, which we'll call Text B, followed by a literal translation:

Text B: From the Preface to the Ormulum With a Translation

Ðiss iss nemmnedd Orrmulum
Forr þi þatt Orrm itt wrohhte,
Itt iss wrohht off quaþþrigan,
Off Goddspellbokess fowwre.

This book is named Ormulum
Because Orm it wrought,
It is wrought of the *quadriga*
Of Gospel-books four.

Because the *Ormulum* is only two centuries older than the Old English passage in Text A, there's no reason to make a detailed analysis. We'll reserve our more comprehensive observations for later samples from the *Canterbury Tales* and a Wycliffe translation. However, a few observations should be made about the syntactic and semantic components.

The deep structure of each of Orm's first three lines derives from a familiar rule, **Sequence → (Conj) NP + PP.** Which one contains the conjunction? The third structure, which includes the fourth line, has how many prepositional phrases? It's generated by the unchanged **VP → V (NP) (Prep-Phrase) (Prep-Phrase) (Manner).** The **V** becomes *wrohht,* with *iss* added by **T-pass.** Actually, "is named" and "is wrought" are quite modern sequences, as we'll see. We find a subject-verb order in all three structures. In the first one, **T-infl** may operate to help provide the form *boc*. However, the leveling of some of the noun inflec-

tions means that [+ fem, + nom, + sing] is now the base, rather than the Old English *bōc*. Once the singular form becomes **ø** for nouns in general, this part of **T-infl** will be lost to the language. **T-agr** imparts the matching features of *boc* to its determiner, giving *þiss*. Again, *þiss* in Old English was [+ neut, + nom, + sing], another indication that gender is being leveled in order for *þiss* to co-occur with a once-feminine noun. *Boc,* of course, is deleted by **T-del.** In the second structure, **T-pro** permutes the direct object *itt* in front of *wrohhte*.

Now for Orm's semantic component. The verb in the first line has remained dental, but with what notable difference from the Old English infinitive *genamian?* In Old English, "for þām" means "because"; whereas Orm uses "forr þi þatt." Obviously much has happened to the Old English *weorc* (*work*). A derivative form has been shifted to a verb base, *worken*. How has metathesis operated on this dental subclass? Today, of course, the verb follows the regular *work-worked* pattern. *Wrought* is no longer its inflection, existing only as an isolated adjective. It's unrelated to either *wreak* or *wreck,* both of which derive from *wrecan* (drive). In the case of Orm's *off,* we should be glad that his orthographic scheme wasn't adopted. Our preposition *off* is discriminated from *of* by the doubled *f*. In reality, his whole scheme was naive. To differentiate long from short vowels, all he had to do was employ different letters, like the *e* and *ɛ* used in our phonemic transcription.

A treatise could be written about the genitive *Goddspell-bokess*. We saw the *spell* base earlier, in *bigspell* (*by-story*). The modern *gospel* comes from *gōdspell*. By Middle English the shortening of the *o* has made *gŏd* (*good*) homonymous with *God*. The pious Orm presumably interprets the base as *God,* intending "four gospels" by his mouthful. The compound illustrates the diminishing Germanic tendency to combine bases into new words. It literally means "God-story-books." So when someone mentions a "spellbinding story" today, we can smilingly accuse him of what historical repetition?

History: 1250–1350, 1350–1500

Let's advance the calendar to 1250–1350. This is the second of our arbitrary periods, when English begins to reassert itself. Several political and military events play a role in the slow move-

ment away from French and back to English as the language of the country. In fact, the following remarks can be anticipated by opening a history to the right date and noting certain events. Most of them involve patriotism and nationalism. When we read about Henry III's love for things French, including his marriage to Eleanor of Provence in 1236, we can imagine the public reaction to the streams of fortune-seeking Frenchmen descending upon the island. They were new oppressors, taking high positions in church and state. In 1258–65 the English descendants of the Norman conquerors fought the Barons' War, capturing the king but then losing the battle of Evesham. This effort to drive out the foreigners, supported by the middle class at a representative assembly, contributed toward the nobles' beginning use of English by 1300. Manuals for the learning of French began to appear, and these later treated it as a foreign language. Meanwhile, the greatly accelerated flow of French borrowings helps document the wider use of English, even as more native words became obsolete.

Edward I, succeeding Henry III in 1272, was the first truly English king in centuries. Relations with the French, who were determined to regain Aquitaine, were poor. Patriotism rose sharply on both sides with the Hundred Years' War beginning in 1337. By the end of the bloody struggle in 1453, after the burning of the "witch" Joan of Arc, England was again English. The only Continental possession left was the city of Calais. French was a language spoken across the Channel, not in England. The native tongue had triumphed. The war overlaps deeply into the last period of Middle English, 1350–1500, which can be characterized by the phrase "dominance of English." Once again the language was "official," used in the court and politics, and, of course, by the people, who had kept it alive for three centuries. By this time, anyone who spoke French in England was apparently bilingual. By 1385 English was in general use in the few existing schools. After the introduction of printing, there was a large expansion of the reading public. After 1489 all statutes and parliamentary records were written in English only.

The concentration of power and people in London, plus Caxton's contributions that we'll discuss shortly, helped make the London dialect the standard one by 1500. However, there was still great diversity in the language as spoken and written in the different areas of the kingdom, sometimes within the same

county. Orm's attempt at a standardized spelling had failed because people seemed not to have known his big book. The tradition of grammars and dictionaries that instruct people how to spell and, alas, how to pronounce and construct sentences along prescribed lines wasn't developed yet. In fact, what was probably the first purely English dictionary didn't appear until 1604. Even then, Robert Cawdrey designed *A Table Alphabeticall* to help the reader only with learned borrowings from Hebrew, Greek, Latin, and French.

Four Major Dialects

Four major dialects of Middle English are usually singled out: Northern, West Midland, East Midland, and Southern. Predictably, they don't coincide with the four Old English dialects. The flat East Midland area was suitable to agriculture. The social-political-intellectual prestige of Cambridge and London heavily contributed to the development of a prestige dialect in that area. Crudely, Old English Northumbrian became Northern. Mercian, spoken between the Humber and the Thames, was divided into West Midland and East Midland. In the latter dialectal area, London was a kind of average between the extremes of the north and the south. West Saxon and Kentish became Southern, though the Jutes' descendants have maintained a somewhat different English even until today. The decline of Wessex, coupled with the emergence of London, helps explain why the rather dominant West Saxon dialect was superseded by London Standard. Technically, this was a special part of East Midland. In many respects the history of our language since 1500 is that of London Standard.

Undoubtedly some linguistic influence was exerted by the Great Individual Writers, as literary scholars often term the span of 1350–1500. It's impossible to document their influence directly, despite earlier critics' claims of Geoffrey Chaucer's supposedly key role in developing London writing as the standard. If we inspect the plentiful documents of the day,[10] echoes of Chaucer and William Langland simply don't appear. After all, Chaucer, the court poet, wasn't transcribing samples of speech

[10] In *A Book of London English 1384–1425*, ed. R. W. Chambers and Marjorie Daunt (Oxford: Clarendon Press, 1967).

as a model for writers of wills and petitions to copy, much less as samples for analysis. Literary men have preconceived artistic purposes.

Except for singling out a few earlier works like "Dream of the Rood" and *Beowulf,* critics usually designate 1350–1500 as the beginning of English literature. They hail Langland's alliterative *Piers Plowman,* John Wycliffe's prose sermons, the unknown poet's *Sir Gawain and the Green Knight* and *The Pearl,* and especially Chaucer's tragic romance *Troilus and Criseyde.* Caxton translated and published some twenty-two folio volumes, including Malory's powerful *Morte Darthur* in 1485. Besides printing Chaucer's *Canterbury Tales,* he attempted to "fix" English orthography, in which each word would have an unvarying spelling. His introduction of a printing press to London had revolutionary effects upon standardizing the written word.

Chaucer's Prologue

Chaucer and Langland used about ten percent French and Latin items in their works. Chaucer's conservatism and popular qualities suggest, however, that his writings may not have introduced these numerous foreign words into English, although they are the first record we have of the words. The borrowings may have already been in English. There were probably pre-Chaucerian writings that have been lost, works in which some of the words might have been used earlier.

We'll conclude our chapter with a discussion of two other passages. The first is from Chaucer's Prologue to the *Canterbury Tales,* primarily to illustrate the pronunciation, leveling of inflections, and use of French loanwords. A passage from the Wycliffe Bible will then permit some larger generalizations about Late Middle English. Now for Chaucer's opening lines, together with a tentative phonemic transcription:

Text C:
From the Prologue to the Canterbury Tales

Whan that Aprill with his shoures soote
hwɑn θɑt ɑ:prɪl wɪθ hɪs šu:rəs so:tə

The droghte of March hath perced to the roote,
θə dru:xt ɔf marc haθ pɛ:rsəd to: θə ro:tə

And bathed every veyne in swich licour
and ba:ðəd ɛvrɪ vɛən ɪn swɪc lɪku:r

Of which vertu engendred is the flour;
ɔf hwɪc vɛrtyʊ ɛnjɛndrəd ɪs θə flu:r

Whan Zephirus eek with his sweete breeth
hwan zɛfɪrəs e:k wɪθ hɪs swe:tə brɛ:θ

Inspired hath in every holt and heeth
ɪnspi:rəd haθ ɪn ɛvrɪ hɔlt and hɛ:θ

The tendre croppes, and the yonge sonne
θə tɛndrə krɔpəs and θə yʊŋgə sʊnə

Hath in the Ram his halve cours yronne,
haθ ɪn θə ram hɪs halvə ku:rs ɪrʊnə

And smale foweles maken melodye,
and smalə fu:ləs ma:kən mɛlɔdi:ə

That slepen al the nyght with open ye
θat sle:pən al θə nɪ:xt wɪθ ɔ:pən i:ə

(So priketh hem nature in hir corages);
sɔ: prɪkəθ hɛm natyu:r ɪn hɪr kʊrajəs

Thanne longen folk to goon on pilgrimages
θan lɔ:ŋgən fɔlk to: gɔ:n ɔn pɪlgrɪmajəs

As has been indicated, we can't ever be certain about the accuracy of reconstructed phonology. Chaucer might not recognize his own lines or even understand them if he could hear them read according to our transcription. Yet his fairly regular meter and other evidence suggest that this is a reasonable guess by someone living six centuries later. We're more certain of his spelling, because of scholars' careful collation of all known

Chaucerian manuscripts, including the Ellesmere ones, the best source.

Changes from Old English spelling immediately stand out, indicating the replacement of certain letters by their modern counterparts, not always for the better. Consider the vowels. The æ has dropped out. Thus *hæþ* has become *heeth; smæl, smale;* and *slæpan, slepen.* Scholars believe that Chaucer's speech had no /æ/. There's a new diphthong in *ye,* from the Old English *ēage,* later to become /aɪ/. The vanished diphthong *ea* has been lost from the spelling, as in *halve.* Among the stops, *c* is now *ch* when pronounced /c/, as in *swich.* Note the /sw/, which isn't to be simplified as *such* for a while longer. The recent French borrowing *March* follows the *ch* pattern. When pronounced /k/, the *c* is now *k* in *folk, maken, priketh.* The /lk/ in *folk* isn't yet simplified to /k/. *Priketh,* of course, will later be respelled with *ck.*

Of the fricatives, *ð* has vanished. Likewise *þ* in Chaucer's *hath, breeth, heeth,* and *priketh,* though it appears in Wycliffe and some later writing. The *sc* of Old English *scūr* has become the modern *sh* in what word? The /x/ is spelled *gh,* as in *droghte* and *nyght.* This palatal is later lost. The /y/ pronunciation of Old English *g,* along with the lost diphthong, has been respelled in *yonge,* from *geong.* Overall, several words containing doubled consonants haven't lost one of them yet. Thus Chaucer writes *croppes, sonne, thanne, yronne.* The latter word is interesting, for the *ge-* in its source *gerinnan* has been changed in what way? The letters of the *hw* cluster have been metathesized, despite retention of the /hw/ pronunciation. What are the two examples in Text C?

Our transcription exemplifies a major change in Middle English, the neutralizing of unstressed syllables. Consider only the examples from the first quatrain:

shoures	soote	the	perced
roote	bathed	every	engendred

Every unstressed *e* is now the schwa /ə/, which is also the pronunciation of *a, u,* and *o* in unstressed syllables. Having appeared in the language at least by 1000, /ə/ is replacing unstressed vowels right and left by Late Middle English. Today it's our highest-frequency vowel.

What is cosmic about the statement that /ə/ is being used

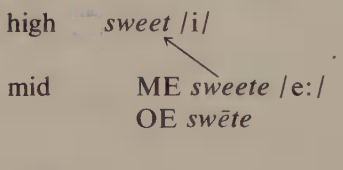
more often? For the answer, we must rely on our native-speaker competence. How many syllables do we have in each of the modern words *sweet, pierced, root,* and *bathed?* Yet Chaucer has two. Clearly, in Middle English the unstressed vowel has become /ə/. Later it's lost, or apocopated, in such words. What happens to the inflectional last syllable of Chaucer's two verbs above? Right: it vanishes, with the /t, d/ suffix attaching itself to the verb base. This process generally explains why the Old English preterit plural form drops out of the language.

Now note the [− sing, + pres] forms, *maken* and *slepen* in (9–10). After the later loss of /n/ in these and some other verbs, /ə/ vanishes too. Thus we have the uninflected modern "when fowls make and sleep." The same thing happens to the Old English infinitive suffix *-an,* as in *slǣpan* (/slæ:pɑn/). In Chaucer it's *slepen* (/sle:pən/). After the later loss of /n/ and /ə/, together with the change in /e/ during the Great Vowel Shift, the infinitive becomes our monosyllable /slip/. A similar process affects nouns, as in *roote* and *shoures.* Likewise, the singular and plural forms of the weak adjectives collapse into *-e,* like *soote.* All adjectival gender and number distinction is lost. In consequence, that part of the function of **T-agr** drops out of the language, since the adjective no longer accepts the matching inflectional features of its noun. In general, then, vowel neutralization in unstressed syllables causes the loss of many inflections. Word order thereupon tightens.

We might take a final look at Chaucer's eight words above. Consider *perced.* Leveling later occasions a phonological change from /pɛ:rsəd/. Why is the result /pɪrst/? The terminal dental can't be lost, and yet English doesn't have a cluster */rsd/. Either /s/ becomes the voiced counterpart /z/, as in /rzd/; or /d/ is made into voiceless /t/ to harmonize with voiceless /s/, as in /rst/. The latter happens. Along the way, /ɛ/ shifts to /ɪ/. We retain the *-ed* spelling today. The change is less complicated for *showers,* although the Middle English last syllable is still apocopated and the /z/ attached to the base. For the monosyllable *the,* there can't be apocope, or the whole determiner would be lost. *Engendred* undergoes metathesis in its apocopating bath, for English lacks */drd/. So /ə/ wedges between /d/ and /r/, for our /ɛnjéndərd/. Comparable mayhem is done to *every,* explaining Chaucer's possibly syncopated /ɛvrɪ/. In Present-Day English it's still trying to squeeze out /ɛvərɪ/.

His language is strategic in time. These are the last moments before the beginning of the Great Vowel Shift, which isn't complete until the eighteenth century. In the process, every long vowel becomes a different one, higher and farther to the front. The shift strikingly differentiates these vowels and their modern counterparts. Our spelling faithfully mirrors Late Middle English, since it's essentially Caxton's and precedes the shift. Of course, the unrepresentativeness of our orthography causes many of us to feel a little helpless when we have to write something but don't have a dictionary handy. Yet a reformed spelling would erase most of the direct evidence of English phonological history.

Returning to Chaucer's *maken,* we are informed by the *e* that the word is disyllabic, having two syllables. The *a* tells us that the vowel of Old English *macian* has become /ɑ:/ in Middle English, and later goes through the Great Vowel Shift to get to the modern mid-front /e/ in *make*. Any spelling reform to **mek* would destroy this visual history as effectively as the Danes destroyed many manuscripts. Since the shift only begins about 1400, we'll wait until the appropriate later moment to describe it in detail. Two more of Chaucer's front-vowel shifts, *sweete* and *inspired,* can suffice for now. The former is best shown graphically:



Inspired is a French or Latin borrowing that convivially goes along for the Great Vowel ride. However, its medial vowel is already high front, /ɪnspi:rəd/. It can't go any higher without becoming a consonant. This can't happen. Grimm's and Verner's laws permitted only consonantal participation. We know what the Great Vowel Shift permits. So the new low central vowel /a/ is invited to join in as a diphthong, ultimately providing our /ɪnspaɪrd/. Again, the apocopated /ə/ has gotten into the act. How?

The lexicon of Chaucer's three quatrains is also worth glancing at, if only to count the loanwords. These introductory, elegant lines aren't at all typical of the colloquial nature of much of the

Canterbury Tales. Otherwise, we might have to conclude a native French ancestry for the jolly Geoffrey. The lines almost seem to be the result of his having deliberately plugged into the deep structure about two French items for every native one. Naturally he didn't do so. He was simply a lot better in French than are most English speakers today. Nevertheless, after deleting the Latin borrowing *Zephirus*, we should be able to tabulate up to fifteen French loanwords without much difficulty. What are they?

We can't pronounce the words according to French stress rules, or the meter will be ruined. When a word is naturalized, the intonation and sounds are forced into the native pattern. For example, *engendrer* and *tendre* have a /-dr/ that English lacks. An inserted /ə/ gives the Modern English /dər/. Or consider *corages*, which is [+ count]. Why is the base form [− count] today? Part of the explanation may be the glittering new import's failure to compete with *hert*, boastful of an Old English grandfather (*heorte*) and an even earlier lineage. To survive, the immigrant surrenders its "heart" meaning to *heart* and moves across town to the "courage" neighborhood.

As the linguist necessarily counts items and their frequency in his corpus, we might do so too. Count the occurrence of each of the following words in the quatrains: *whan, that, with, the, of, to, and, in*. Which wins? Which two place? Which one shows? The rest have two appearances each. We shouldn't be surprised at the high frequency of these key structural items from Old English, both here and in other works. Conjunctions, for instance, are derived in the very first deep-structure rule. The auxiliary *hath* might easily have been added to the above list, because of its perfectly modern sequencing with what three past participles in Chaucer? In fact, its curtain calls illustrate the suddenly increasing sophistication and frequency of auxiliaries. We'll need to broaden the Old English rule for **Auxiliary** in a moment.

A few more native words deserve comment. A compounding of the earlier *æfre* and *ælc* (*each*) provides Middle English *everich*. How does the word become Chaucer's *every?* Remember that the unstressed suffix *-līc* changes to *-ly*. The vowel is retained. Otherwise, we'd say *"ever man," as is sometimes heard in nonstandard speech. *I* develops from *ich* in a similar way. Chaucer's *eek* doesn't have the good fortune of *every*. Derived from *ēc* and related to the verb *ēcan*, it valiantly jousts with

the compound *eallswā*. Of course, *also* wins, with *eke* becoming archaic. The verb, however, seizes a preposition to survive as "eke out." *Holt* (grove) becomes archaic too. The Great Vowel Shift and respelling explain *heath*. *Croppes* (shoots) persists, meaning "product or yield" today. Finally, there's a modern employment of *to* in "longen to goon." If we recall, the infinitive *gān* was obstreperous in Old English. In Chaucer's lines the /n/ isn't leveled yet. Once it is and the vowel obeys the Great Vowel mandate, *go* emerges. Soon all infinitives will lose the Middle English *-en* suffix. Then **T-to** can insert the *to* and provide a wholly modern structure like "long to go."

These generalizations about Chaucer's phonology, inflectional leveling, and lexicon equip us for another sample. Because Middle English arbitrarily covers four centuries, what we must do is extrapolate a general description. It should derive from many writings, including the *Ormulum,* all of Chaucer, and other texts. Prose is usually a better sample for study than is poetry. Verse may be especially unrepresentative of speech because of meter and rhyme. There is one single, almost continuous record of language changes over the centuries, the many translations of the Bible. Had there been prose translations of certain Old English prose writings every century or so until now, those records would undoubtedly be preferable.

Wycliffe Bible

Lacking these prose translations, we turn to a passage from the Wycliffe Bible, Genesis iii, 1–3. It will be the chief source for our general description of deep structure, the lexicon, and **T** and **P** rules. Certainly this first appearance of the entire Bible in English is closer to us linguistically than to Late West Saxon, despite the six centuries separating us from it. Even if few of his countrymen could read and write, Wycliffe claimed that they wanted a Bible in the vernacular. Literate Englishmen had necessarily read it in French or Latin until then, since few books had been rendered into English. Wycliffe's colleagues and helpers completed the monumental translation from the Latin Vulgate about 1395. The considerable difference between it and an Old English rendition of about 1000 will be sketched in the rest of the chapter. The verses have been renumbered to accord with the linguistic structures therein.

Text D:
Passage from the Bible in Old English

(1) Ēac swylce sēo nǣddre wæs gēapre ðonne ealle ðā ōðre nȳtenu ðe God geworhte ofer eorðan. (2) And sēo nǣddre cwæð tō ðām wīfe: (3) "Hwī forbēad God ēow ðæt gē ne ǣton of ælcon trēowe binnan Paradīsum?" (4) Ðæt wīf andwyrde: (5) "Of ðǣra trēowa wæstme ðe synd on Paradīsum wē etað, (6) and of ðæs trēowes wæstme þe is on middan neorxnawange God bebēad ūs ðæt wē ne ǣton, ne wē ðæt trēow ne hrepedon ðī læs ðe wē swelton."

Text E:
Passage from the Wycliffe Bible

(1) But & þe eddre was feller þan eny lyuers of þe erþ: þe which made the Lord God. (2) Ðe which seide to þe womman: (3) "Why comaundide God to ȝou þat ȝe schulde not ete of ech tre of paradise?" (4) To whome answerde þe womman: (5) "Of þe fruyt of treese þat ben in paradise we eten. (6) Of þe fruyt forsoþ of þe tree þat is in þe mydyll of paradise comaundide us God þat we schulden not eten & þat we schulden not touchen it lest parauentour we dyen."

Deep Structure

The deep-structure rules come first. They provide the underlying subject-verb order, as we discovered in Old English in Text A. Actually, Text E isn't quite typical of the usual surface-structure order of Middle English. **T-perm** makes inversions like (3, 6), as we'll see. The important thing to note for now is the subject-verb sequence of "þe eddre was feller" (1) and "þe which seide to þe womman" (2). The "þe which" means "who"; it's the subject of *seide*. Both it and the comparative "was shrewder than any creatures of the earth" are like ours. So are their predecessors in Text D, the *-re* comparative suffix having become *-er*. The relative-pronoun sequences in (5, 6) are equally modern in both texts. Both have embedding, with (D5) looser syntactically than (E), "trees that be in paradise." "The tree that is in the middle of paradise" (E6) even has today's verb form.

Modals and Auxiliaries

The modal-auxiliary structures, quite close to ours today, have an important and expanding role. Whether their increased variety and frequency are due to inflectional leveling, we don't know. We observed earlier that a relatively uninflected language often has a tighter word order. In Middle English, some of the tightening may result from development of a sophisticated modal system and the auxiliaries *have* and *be*. Here's a modern example: "should have been slain" is the only possible order for the four words. In a sense, all three words preceding *slain* are "helpers" and would seem interchangeable. Instead, they're rigidly ordered, as can easily be proved. In Old English, any one of six modals could be generated by **Aux → Tense (Modal).** The same rule can generate "ye should not eat" and "we should not eat and touch" (E3, 6). Note the second-person plural *gē* employed in (D3).

Now recall Text C, where *hath* can't be a modal in its three occurrences. Each requires the past participle of the verb, as in "hath inspired." Modals co-occur only with the verb base or infinitive form, as in Chaucer's "can hyde" and "shalt make." Nor is *hath* a verb in "hath inspired"; the latter word is. *Be* isn't a modal either, in Middle English or in today's speech. Even in the dialectal "I be gone," *gone* is the inflected verb. This kind of *be,* distinct from the copula, is an auxiliary like *have*. It follows a modal when there is one. So for one of the few times in our book, we must enlarge a rule. Otherwise, "have" and "be" auxiliary structures can't be generated. The history of English is more often a process of narrowing of the functions of rules. Note that even if we broaden the auxiliary rule to accommodate more of Chaucer, it still must generate a variety of sequences:

1. And by my drem it is now sene. (is seen)
2. Whilom ther was dwellynge in my countree. (was dwelling)
3. I have wept many a teere. (have wept)
4. We han ben waitynge al this fourtenyght. (have been waiting)
5. Ne hadde he ben holpen by the steede of bras. (had been helped)

6. I sholde have dyed, ye, longe tyme ago. (should have died)
7. When that he sholde han be slayn. (should have been slain)

All seven are quite modern. The last two include the modal *should,* plus the auxiliaries *have* and *be*. Therefore, the **Aux** rule must be broadened to the Present-Day one: **Aux** → **Tense (Modal)** (*have* + *-en*) (*be* + *-ing*). This is a fundamental change, as is always true when a deep-structure rule is altered.

Consider the economy of the new version. Chaucer employs different optional parts:

Tense + be + -ing + **V**	(2)
Tense + have + -en + **V**	(3)
Tense + have + -en + be + -ing + **V**	(4)
Tense + **Modal** + have + -en + **V**	(6)

How do the *-ing* and *-en* become attached to **V,** as they must be? Neither Middle English nor modern speakers use *"being dwell" (2). *"Shall had die" (6) is equally unacceptable. What's needed is **T-aux.** It inverts the auxiliary suffix and the following verb unit. Each suffix can be permuted only once. This is a highly precise movement, so that the generalized **T-perm** won't work.

Let's practice with (4), where **T-aux** is applied thrice. **T-infl** first arbitrarily makes **Tense** into **present. T-aux** inverts **present** and *have,* then *-en* and *be,* and finally *-ing* and *wait*. The result is "have + **present** be + -en wait + -ing," for "have been waiting." Actually, **T-aux** existed in Old English to invert the **Tense** and the modal or verb. For example, **Tense** + *be-* is generated in (D1). **T-infl** makes **Tense** into **past.** Then **T-aux** permutes the sequence to *be-* + **past.** A **P** rule later provides /wæs/ or *wæs,* as we know.

What about Chaucer's (1, 5, 7)? To account for them and for Orm's "iss nemmnedd and wrohht," we need **T-pass.** The passive transformation wasn't illustrated in Old English because the modal and verb structures were quite elementary and seldom required it. West Saxon never had anything like the expanded sequence "should have been slain." Remember "wæs gegaderod" in Text A? If really conscientious, we discovered that it could be generated only by **T-pass.** The same rule does the trick for Middle

English, changing active voice to passive. Thus "My dream sees it now" (1) becomes "It is seen by my dream." Specifically, **T-pass** inverts the subject and object, adds *be* + *-en,* and inserts *by* before the object. The process is the same for (5, 7):

the steede of bras **past** have + -en help- he →
he **past** have + -en be + -en help- by the steede of bras

when that X **past** shal have + -en slē- he →
when that he **past** shal have + -en be + -en slē by X

To derive Chaucer's finished (1), **T-perm** permutes "by my drem," and **T-adv** does the same for *now.* In (5) **T-ne** first adds the negative particle, and **T-ques** inverts *ne hadde.* In (7) we don't know who's to do the slaying, and so he's called "X." Our old friend **T-del** deletes "by X," which wasn't very revealing anyway. **T-aux** moves the verb suffixes to the right place in all three structures.

Modal-auxiliary structures, developed in Middle English, steadily gain in variety and frequency of use through the centuries. Though existing by 1000, progressive constructions like the equivalent of "was running" don't occur very often before 1400 and don't really enjoy high frequency until Modern English. Our six West Saxon modals continue in that function throughout Middle English, with the kind of growing complexity that has been illustrated. Their spelling has become *can, dar, mai, mōt* (*must*), *shal, wille.* Moreover, their number is expanded by four: *oghte, ginnen, dōn, nēden.*

Remember West Saxon *āgan,* meaning "possess"? In Middle English that function is served by its derivative *owen,* the predecessor of *owe* and *own.* Its new derivative, *oghte,* temporarily becomes a modal, as in Chaucer's "oghte moeve" and "oghte doon." It must be added briefly to our list, until its death about 1500–1600. Even with such a short tenure, it's mixed up. Note two of Chaucer's structures that our expanded **Aux** rule can't generate: "But alle power aughte ben desired" and "I oghte have take hede." The homonymous verb *oghte* takes *to* like other verbs, as in his "oghte to passen" and "oghte secreely to hyde." *Ought* can't be classified as a modal today, because a structure like "ought pass" (as opposed to "ought to pass") is archaic if not obsolete.

The other three modal additions are defective, perhaps a major reason why they've died too. *Ginnen* means that action is commencing. Its origin returns us to "ongan lǣran" of Text A. *Onginnan* was already initiating a modal career in Late West Saxon. As a verb, however, it was defeated by *beginnan*. Aphesis then gives us the modal *ginnen,* with *gan* its high frequency preterit. For example, Chaucer says of his Oxford clerk: "And bisily gan for the soules preye." *Ginnen* is defective, in that it doesn't combine with the auxiliaries. We never find *"gan have gon" or *"gan have being gon." Like *oghte,* its homonym is a verb, taking *to* like other verbs. Actually, Chaucer's "Gan to swere" is a wholly modern use of **T-to,** like our "began to swear." Neither *ginnen* survives, of course, except in the archaic *gin* and *gan.*

Dōn, the ninth modal, is equally temporary. It enjoys considerable frequency and operates like the original six, except for the inability to combine with the auxiliaries. Thus we find "dooth carie" and "dost retorne" in Chaucer. Like *oghte* and *ginnen,* a homonymous form operates as a verb. This **V** also co-occurs with a modal, as in his "that mighte don him harm." *Nēden,* the tenth, joins the modal ranks of Late Middle English more or less permanently. Both it and *dōn* gain even higher frequency by Early Modern English. At this point we should underline one characteristic of the **Aux** rule. It never permits more than one modal per sequence, whether West Saxon or Present-Day English. "I can go" is acceptable, but not *"I might can go." *Might* and *may* simply aren't adverbs, despite their nonstandard use by some speakers.

Lexicon

The semantic component of Middle English is our next consideration. We'll concentrate on the biblical passages. After generating the six structures therein, we plug in the appropriate words. The lexical rules help select the right meaning. Naturally the vocabulary from which the words are chosen for (E) is quite different from that used for (D). The vocabulary available for (E) is probably larger, and it becomes more so even later. As many native words were obsolete by Late Middle English, they don't exist as choices for (E), although available for (D). Moreover, six items in (E) are new loanwords and so weren't available

for (D). Note the five French borrowings: *commaunden, fruit, paradise, touchen, parauentour.* Which four Old English words in (D) die? Perhaps there's poetic justice in *peradventure*'s later becoming archaic, as contrasted with the continued utility of *lest.* In (D) *gēap* (cunning) develops into *yepe* before becoming obsolete. Its Middle English replacement, *fell,* may be losing ground except in phrases like "fierce and fell."

Nǣddre, the generic word meaning "serpent or the devil," has experienced aphaeresis. What was lost? The French loanword *serpent* later appropriates the meaning, particularly the "devil" part. What's the modern version of Old English *snaca* (reptile, snake), which eventually gains the generic quality? What's the modern spelling of *edder,* which often means "viper"? Our momentary concern with French borrowings in Wycliffe doesn't mean that the translators had a deep love of that language. In fact, they wanted a "native" Bible. Consciously or unconsciously, they would have been more likely to select a Germanic word when possible. The number of originally French items in the brief passage simply illustrates the great lexical influence of that tongue, especially the use of biblically important words like *fruit* and *paradise.* That is, the bulk of the Latin, French, and Scandinavian words borrowed up to about 1500 are so thoroughly naturalized that Englishmen were probably unaware of the nonnative origin. There's no such thing as a "native" word, a term employed here only for convenience. The situation is equally true today. Few people could write anything if required to use only originally Germanic items. The fifteenth-century Londoner or the modern American would have to look up the etymology of almost every word, while execrating the ridiculous instructions if not the instructor.

In (E) the probably Scandinavian *dyen* has replaced the once high-frequency *swelten.* The latter is in a losing battle with *sterven* and the probably Scandinavian *slēn,* both of which we've seen earlier. How is the meaning of *swelter,* the modern survivor, related to "die"? Do *die* and *slay* differ today? The passage doesn't require any plural personal pronouns, or we'd observe the Scandinavian replacements. They occur elsewhere in the Wycliffe Bible and earlier in Orm. The two passages reflect competition between native words too. *Nȳten* (D) loses to *nēat* (beast, cattle). *Cattle,* from a French word meaning "personal property," eventually squeezes out *neat,* except in "neat's-foot

oil." *Middel* (half, middle) is contesting with variant spellings. Which meaning does *mydyll* have in (E)? Which is lost to the descendant of Old English *healf?*

In (E) there are five West Saxon compounds now respelled into a form somewhat shrouding the original elements. *Answeren* conceals the *and* + *swarian*. *Forsooth,* little changed but seldom used today, is from *for* + *sōþ*. *Sōþ* has what modern descendants? *Lest* successfully hides its origin from *ðȳ* (inflection of *sē*) + *læs* + *ðē*. So does what later-obsolete combination of *lyue* (*life* or *live*) + *ouer* (*over*) in (1)? *Wīf,* earlier a female person, was often compounded as *wīfmann*. By Late Middle English it's *wimman* and then *woman*. In (E) the word indicates an adult female human being, having replaced the generic *wīf* of (D). Chaucer's "wif" of Bath may be a woman or a wife. How do the two words differ today? Related to these is *hussy,* coming from the modern *house* + *wife*. Has the word degenerated or elevated? Like *ginnen* and *lest,* it has lost several phonemes.

Transformations

Now let's consider the transformational rules. **T-aux** and **T-pass** have already been illustrated. In Middle English the latter permits complex structures like "should have been slain." Some other **T** rules are altered by the large inflectional leveling. We've mentioned what seems to be the sequential causes: the change of unaccented vowels to /ə/, the loss of inflectional /-n/, and finally the loss of /-ə/ in the vestigial inflection. For example, **T-agr** no longer provides a matching feature of gender. We've observed the lack of gender in Chaucer's adjectives. The determiner *sē* and its inflectional horde telescope into *þe* and then *the*. **T-agr** is no longer applicable to the determiner or the adjective in **case** either. However, *that* remains from the neuter determiner, along with *this,* contrasting with *the* and requiring agreement in number. Almost all nouns have adopted the *stānas* form for the plural, as changed to *-es* from the Old English [+ masc, + nom, − sing] form. **T-infl** continues for them, but inflection for gender has been lost from most nouns. The dual number of pronouns dies, but most other pronominal inflections persist. Note the clearly designated plural number in "to ȝou þat ȝe schulde not ete" (E3). We don't have this today; we're forced to use something

like "you two." The modern "To whome" (4) derives from the dative *hwām,* victor over the accusative *hwone.*

The infinitive suffix *-an,* made into *-en,* is lost from verbs. We've noted the transitional stage of **T-to** in Chaucer's "longen to goon," where *-en* is still retained. This can be contrasted with "The hooly blisful martir for to seke," only five lines later. Here /n/ is lost from *seken,* but not yet /ə/. About a hundred lines later Chaucer retains /ə/ and *to,* but drops *for:* "And peyned hire to countrefete cheere." He apparently has three choices, demonstrating the expanding variety of verbal structures.

Once /ə/ is apocopated, **T-to** becomes almost completely modern. No doubt the change results from loss of the rest of the infinitive suffix. Simultaneously **T-to** can no longer apply to the modals, which in Old English simply added *-an* to the verb base, as in "wilt behealdan." **T-to** operated similarly on some Old English verbs. For instance, "secgan hȳrde" (have heard say) and "licgan þence" (think to lie). These few verbs occasionally resist the change of **T-to** in Middle English, refusing the *to* in the transformation. As a result, we find "he herde speken" (heard spoken) and "I þenke telle" (think to tell). By 1500 *oghte* seldom appears without *to.* So it survives only as a special kind of verb. Modern modals never take *to,* as in *"can to go." The word order thereby becomes yet more rigid. Incidentally, the change of the suffix *-an* to *-en* is exemplified by the three loanword verbs in (E). All are analogically fitted into the new *-en* pattern as they are borrowed. The copula *bēn,* the son of *bēon,* doesn't occur as an infinitive in the passage. In (5) we do have "treese þat ben in paradise."

Verbs

In Middle English, like today, the verb continues to use **T-agr** to acquire the matching inflectional features of person and number from the noun. **T-infl** is simpler for verbs in Middle English than in Old English, thanks to the continuing apocopation of the *-on* preterit plural suffix. Like **T-to, T-ing** is much more frequently used. One function creates the verbal noun, or gerund. In Chaucer's "Of prikyng and of huntyng for the hare / Was al his lust," the gerunds convey a sense of action. After adding *-ing* to the verb base, **T-ing** shifts the word to temporary noun status. Modern gerunds thereby retain a verbal function, despite their

filling a noun slot. The situation is historically confused because of the rising eminence of the present participle *-ing*. This developed from Old English *-ende* (for example, "Wæs sē engel sprecende"). However, both the participle and the original two *-ing* suffixes (from *-ung*) deriving nouns from verbs enjoy much higher frequency in Middle English.

Verbs continue the shift to the dental group if not already a member. The vocalic ones, still rather numerous, are often of high frequency. Of the six verbs in (E), only *ēten* is vocalic. Has it surrendered yet? Many are dead by 1100, like *hnīgan* (bow). Others die soon after 1500, like *agrīse* (frighten). For many of the survivors, dental forms develop that compete with the vocalic forms for centuries until the entire verb shifts over. Of the vocalic group, the first subset is the most tenacious. It remains vocalic or dies except for the later shift of *glide*. Even *glid* as the preterit still doesn't cause many raised eyebrows today. *Bide, bite, ride, rise,* and *write* continue merrily unshifted. The second subset is none too resistant. *Bow* and *creep* shift in Middle English. *Flee* and *dive* shift later, with *dove* now trying a comeback. *Bid, fly,* and *freeze* continue unchanged. The third subset is least resistant, losing *delve, help, melt,* and *yield* to the dental side. The siren call isn't heard by *drink, find, sing, swim.* In the fourth subset, *shear* turns dental in Middle English. *Break, come, steal,* and *tear* are unperturbed. In the fifth, *mete* holds out, despite competing forms, until the late sixteenth century. *Eat, see, speak,* and *tread* continue. In the sixth, *bake* and *load* become turncoats. *Shake* and *stand* stand unshaken. In the seventh, *crow, hew,* and *sleep* surrender. But *beat* and *fall* fight rather than switch. Examples of the seven subsets of Middle English vocalic inflections are as follows:

1.	wrīten	wrǭt	writen	(y)writen
2.	chēsen	chę̄s	curen	(y)cǭren
3.	drinken	drank	drunken	(y)drunke(n)
4.	stę̄len	stal	stēlen	(y)stǭle(n)
5.	mę̄ten	mat	mēten	(y)mę̄ten
6.	fāren (go)	fōr	fōren	(y)fāren
7.	fallen	fēl	fēllen	(y)fallen[11]

[11] Examples are from Fernand Mossé, *A Handbook of Middle English*, translated by James Walker (Baltimore, Md.: Johns Hopkins Press, 1952), pp. 69–73. The ę and ǫ stand for /ɛ:/ and /ɔ:/, respectively.

As has been said, the dental verbs expand their ranks. The three West Saxon subsets are rearranged into the two modern divisions. One forms the preterit by adding /t/ or /d/, as in *hǭpen, lērnen, māken, stēren* (*stir*). Thus we say /hopt/ rather than */hopɪd/. The other, which includes most verbs, adds a vowel and the dental. Some of these undergo a base change in the process, as in *dǣlen, fēlen,* and the dying *lǣren* (teach—remember our favorite *lǣran*). Some dentals become obsolete by Early Modern English, like *hǣrien* (praise). Recruits, however, are flowing in from the French and Scandinavian, which almost invariably take the dental form. In the Wycliffe passage five of the six verbs are dental, including the Scandinavian *dyen*. The French *commaunden* and *touchen* readily follow that banner.

The dissident verbs from Old English are undaunted by all this conversion. After all, they never belonged to either subclass anyway. *Witen* (know), not yet archaic, has extremely high frequency. So do the Middle English trio *bēn, dōn,* and *gǭn,* already discussed. One of the common combinations is "at dōn," which telescopes into *ado* and the related *to-do*.

Other Transformations

Three other **T** rules undergo change. The first, **T-contr,** is gaining in frequency. Its expanded use has continued until today, despite purists' efforts to exorcise the contraction from the language by failing all themes containing one. In Chaucer the various forms of *bēn* combine with *ne* dozens of times. The results are *nam, nart, nis, nas, nere*. What are the modern versions? His modal *wille* is by far the most productive. There are *nil, nilt, nolde, noldest, nilling*. *Witen* also participates, as in *noot* (*ne woot*) and *nost*. So does *hāve,* in *nath* and *nadde*.

The second is **T-ne,** which can be renamed **T-not.** The reason is simple: the more modern *not* or *nat* is beginning to replace *ne* as the negative element inserted before the verb. Examples are Chaucer's "But not wot I whoo did hem wirche" (*wirche* means "carve") and "Axe hym thyself, if thou nat trowest me." Often one or more other negatives occur as reinforcement or emphasis. Note Chaucer again:

He nevere yet no vileynye ne sayde

There nas no man nowher so vertuous

At this time no one had yet thought to question whether two negatives make a positive, or five negatives a—quick, what's the answer?

T-perm permutes the verb before its subject in at least two kinds of sequences. It creates an interrogative adverbial sequence in Wycliffe (3): "Why comaundide God to ȝou?" Note that this is little changed from Old English (D3), except for the occurrence of *to*. **T-perm** also acts upon conditional structures, as in "comaundide us God þat we schulden not eten" (E6). This contrasts with the surprisingly modern subject-verb order of (D6), "God bebēad ūs."

There's a third kind of inversion in Middle English. Consider Chaucer's direct question: "O! hastow slayn me, false theef?" *Hastow,* of course, is a contraction of "hast thou." He constantly permutes the modal to make interrogative sequences, as in "Can he wel speke of love?" Even when there is no modal, the verb is inverted. For instance, in *The Second Shepherd's Play* we find "Gaf ye the chyld any thyng?" A modern translation requires insertion of *do,* to effect "Did you give the child anything?" Clearly our "do" structure has developed out of the syntactic changes occasioned by inflectional leveling. In Middle English, however, **T-ques** creates questions. It provides for the movement of **Tense** plus either **Modal** or **V.** The above examples can illustrate:

He **Tense** can wel speke of love → **Tense** can he wel speke of love

Ye **Tense** give the chyld any thyng → **Tense** give ye the chyld any thyng

T-ques can't be applied unless there is to be a question. This rule gains ever greater utility after 1500 and strikingly affects surface structure.

Phonology

Always the last but not the least in the generative-transformational process, the phonological component has been in great part necessarily discussed in terms of Orm and Chaucer. Perhaps the major change is the loss of the inflectional /-n/ and then /-ə/.

We've also emphasized the difference between long and short vowels. The former are just starting on the Great Vowel Shift that will convert each of them to a different phoneme, and /i/ and /u/ to diphthongs. All the others in stressed syllables will essentially remain unchanged.

Some of the clusters characterizing Old English have become single consonants. *Hrepian* was fleetingly seen about 1000 (D6), but its disappearance from the Bible was hardly due to its cluster. Besides reduction of /hr-/, we find /hl-/ simplified to /l-/. Some of the respelled /hw/ words were glimpsed in Wycliffe, from *hwā* and *hwȳ*. What are their modern versions? The loss of *-ch* in unstressed syllables gives us *-ly*. Our pronoun *I* is another result. The *ge-* prefix has become *i-* or *y-*, as in *yronne*. It's on the way out, as shown in Orm's *nemmnedd*. How is metathesis seen in such spelling as *edder* from *nǣddre?* We've observed Chaucer's *th,* which can be voiced or voiceless, for the dying *þ* and *ð*. The interchangeability of *u* and *v* explains why we're sometimes momentarily mystified by spellings like *vnto* and *vp*. The *y* persists for the most part, despite phonemic unrounding to /i:/, as in *fȳr* (eventually *fire*).

Overall, little happens to phonology in Middle English except for the late loss of the inflectional schwa, together with the replacement of unstressed vowels by /ə/. This replacement might hardly have caused a ripple, except for the consequent effects upon **T-infl, T-agr, T-to,** and even **T-pass.** Alterations in the last two of these significantly affect surface structure. So the center ring of Middle English holds the syntax, composed of a tightening word order and expanding modal-auxiliary structures. The lexicon is in the second ring, peopled by many new citizens and without a lot of the old standby Germanic performers. Phonology is relegated to the third ring, all the way across the linguistic tent. In Modern English, the Great Vowel Shift will demote the syntax to the outer ring, since the surface structures of 1500 are essentially the same as ours today. Old and new immigrants will fairly well replace the natives in the now-crowded second circle. The GVS will appropriate most of the spotlight for phonology.

Activities

1. Find at least two examples of words containing each of the following French and Latin borrowings:

ad-, ac-, af-, ag-, al-	ob-	-ate (*a.*)
com-, col-, con-	per-	-ate (*v.*)
counter-	pre-	-ent
de-	pro-	-ess
dis-	re-	-fy
ex-	sub-	-ious
in-, il-, im-, ir-	trans-	-let
infra-	-able	-ose
inter-	-age	-tion

2. What's the origin of each of the following words (French or Germanic)? Explain why you are or aren't surprised about each: *churl, earl, knave, lady, lord, man, marquis, prince, viscount.*
3. Make a list of the doubled consonants in Orm's verse. Try to find whether any of the words had such doublets in the Old English spelling. On the basis of modern pronunciation, attempt to justify each of Orm's doublings.
4. The reduction to /ə/ of vowels in unstressed syllables continues today. Sometimes the /ə/ is then lost. Find two words that have lost a medial syllable. Here's a clue: does the spelling have a "silent" vowel?
5. Present-Day English probably has only five modals: *can, may, must, shall, will.* List the seven basic structures generable by the rule **Aux** → **Tense (Modal)** (*have* + *-en*) (*be* + *-ing*). Compose two sentences in which *dare* and *need* are used as modals.
6. What are four modern words containing the base *mand,* a variant of *mend,* from French *commander?* Are they instances of hybridization or of a borrowing of the whole word?
7. An important early bilingual dictionary was Randle Cotgrave's *A Dictionarie of the French and English Tongves* (London, 1611). It lists numerous words derived from the French *command,* together with English cognates or meanings. Numerous other words employ the prefix *com-*. Cotgrave's French example is given first in the following list:

command (commaund, power, authoritie)
commande (commende)

commandé (commaunded, bidden, charged)
commandement (a charge, commaund, precept)
commander (to commaund, bid, charge)
commanderesse (a commaundresse, gouernesse)
commanderie (a commaunderie or commaundership)
commandeur (a commaunder, bidder)

Which of these eight have survived? Are there any other cognates today?

8. It's sometimes difficult to distinguish historically between the present participle *-ing* ("was gobbling") and its two homonyms. One homonym, meaning "kind of, descended from," can be illustrated by *farthing* (diminutive of *fourth*). The other, with multiple functions, creates verbal nouns out of verbs and some nouns, as in "Digging is hard work." In our analysis, both are economically created by **T-ing,** along with the earlier one like "needed helping." Find three examples of each of the homonyms. The "kind of" type is scarcer.
9. The tragic history of vocalic verbs, which usually die or surrender to the dental troops, is illustrated by fourteen of the following. Which eleven have remained vocalic?

ache	blow	fret	load	swallow
awake	choose	grind	sit	weep
beat	climb	grow	slay	well
begin	crow	hasten	sleep	wheeze
bind	draw	hew	smart	wield

Chapter Seven

Modern English (1500–1900)

Modern English is arbitrarily assigned the dates 1500–1900. As has been said, the traditional names of linguistic eras are generally unsatisfactory. Yet the tradition of assigning 1500 as the beginning of Modern English is too long established for us to abandon it. Anyway, even if scholars could agree on dates and names for significant previous eras, the language will presumably persist. If *Modern* is employed for the time up until 1900, what will English be called in the twenty-fifth or the thirtieth century? *Post-English* may be appropriate by then, just as the language in Italy today is no longer Latin but its direct descendant, Italian. Consistent with earlier chapters, we'll designate the first period (1500–1660) of the four centuries as Early Modern. Historians usually call this period the Renaissance. Because it is the most important modern period linguistically, as well as intellectually and literarily, it merits most of our attention. The second period will be called Authoritarian English, dating from the Restoration of the Stuart kings in 1660, beyond the French Revolution to 1800. Late Modern, a somewhat redundant name, is useful for the nineteenth century.

Early Modern (1500–1660): History

When did the Renaissance begin? Historians usually suggest 1485 as the date marking the end of the feudal system introduced by the Normans. That end came somewhat dramatically at the Battle of Bosworth Field, when Henry Tudor stood over Richard

III's body. Concluding the bloody Wars of the Roses, during which many Old English manuscripts were probably lost forever, the battle also signaled the beginning of the Tudor line of kings. Crowned as Henry VII, Henry and his Tudor descendants turned England away from medieval policies toward modern attitudes. Then began the Age of Exploration, with its vast commercial, political, and intellectual effects upon the country, together with the linguistic effects that we'll consider shortly. John Cabot traveled to the New World in 1497, after Columbus had proved its existence five years earlier. Not having a Suez Canal, Vasco da Gama had to go all the way around the Cape of Good Hope to reach India in 1498.

Other exterior events affected England. The Protestant Revolution started in 1517, when Martin Luther nailed to the door of the Wittenberg church his ninety-five theses attacking papal abuses. The Reformation in England was partly consummated by Henry VIII's divorce and his establishment of the Anglican Church. His daughter Elizabeth, one of the greatest of monarchs, ruled from 1558. After France finally took Calais, England defeated the Spanish Armada in 1588. Suddenly England was a world power. After Elizabeth's death in 1603, the Stuart kings ruled until Charles I's beheading in 1649 effectively concluded their absolute monarchy. An experimental Commonwealth ended in 1660, when Parliament invited Charles II to assume power. These exterior events affected the English language, as we'll see.

Education

Education became increasingly important during the Renaissance. It expanded greatly under Elizabeth, a patron of the arts. Many new grammar and elementary schools were started in London, which at first were attended by the sons of London merchants and of nearby farmers. Upon perceiving the superior quality of the system, the nobility generally abandoned the practice of hiring a tutor to live in their castle and instruct their children, and sent their sons to the new schools. Enormous advances were made in literacy, with perhaps up to half of London literate by 1600.

These advances partly derived from Johann Gutenberg's invention of movable type in the mid-fifteenth century, when he

printed the famous edition of the Bible bearing his name. Heretofore, the laboriously hand-copied manuscripts had been too expensive for most people. William Caxton's printing press permitted a cheap, mechanical way to reproduce texts that never varied in spelling or in any other way. One page of type could print an indefinite number of copies, each copy orthographically identical. The freezing of English orthography was thereby hastened. Today's spelling problems result from the continuing changes in the sounds represented by the letters of the frozen spellings.

Books

In the history of communication, the invention of the printing press and movable type may rank second only to the invention of writing itself. Now anyone could afford the one-penny broadside poem hawked by a peddler. The common man now had a reason (or the means or the occasion) to learn to read. Writers could draw upon a huge new audience, with a distinct possibility of making money if their books sold. Such developments further heightened the interest in literacy and learning. A related result was the publication of the great Shakespeare First Folio. Since the first modern copyright law wasn't to be passed until 1709, printers had been making an easy profit by producing cheap, incorrect versions of his plays for the Londoners who flocked to the live performances. The Folio of 1623, the supposedly correct edition, was published partly in self-defense.

The writing of books was a distinguishing feature of the Renaissance. Although our concern is naturally with the many language books, we might pause briefly over the literature. The quality of literature had been poor during the century following Chaucer's death. Drab imitators like John Lydgate and Thomas Hoccleve vainly tried to maintain Chaucer's standards. Except for Sir Thomas Malory's *Morte Darthur,* published by Caxton in 1485, great literature didn't re-emerge until the Middle Renaissance. Undoubtedly some of the dynamic conditions of the day contributed to the reigning of three geniuses during the height of the Renaissance: Spenser (1552?–1599), Shakespeare (1564–1616), and Milton (1608–1674). The epics *Faerie Queene* and *Paradise Lost,* together with Shakespeare's sonnets and dramas, are so famous that they need only be mentioned.

Grammars

As for books concerned with language, there was a virtual flood of grammars, translations of the classics, language-teaching manuals, spellers, and some dictionaries. We'll consider them in that order, which is roughly chronological, noting their influence on the English language as we proceed.

The grammars were usually modeled upon Latin grammars, for example, on the comprehensive study by Varro (116–27 B.C.). He and the earlier Greek grammarians were philosophers foremost. Zeno (fourth–third century B.C.), for instance, was more concerned about the nature of knowledge and reality than with actual Greek language data. The first grammars from London adopted the classical philosophical tradition, which was later distorted into a prescriptive attitude. In general, these were more intellectual than linguistic, but they did contain more direct information about language than can be found in the Greek works or in the Latin imitations of the Greek.

One of the earliest grammatical studies in England was John Stanbridge's *Accidence* (c. 1496). Only twenty-six pages, it went through many undated editions. The book is a study of Latin inflection and word order as grammatical devices, explained in English but necessarily illustrated by Latin. All but the last two pages analyze the eight "partis of reason": "noune, pronowne, verbe, partycyple, adverbe, coniunctyon, preposicion, interiec-cyon." We can redesignate *partycyple,* of course, as *adjective.* How many later books about English have copied this pattern, descriptive of pre-1496 Latin? Following Stanbridge, the first Latin grammars began to appear in English editions about 1513.

Translations

Early Modern writers also busied themselves with systematic introduction of classical humanistic works into English. Translations began to pour out after Sir Thomas Hoby rendered Castiglione's *Book of the Courtier* from the Italian in 1561. A major one was Sir Thomas North's masterful translation of Plutarch's *Lives* (1579). Another was John Florio's version of Montaigne's *Essais* (1603). Probably the most famous was the King James Version of the Bible in 1611, partly motivated by a

missionary desire to present the Englishman with the Bible in vernacular, *current* English.

The works that were being translated often contained Latin or Greek words for which there were no English equivalents. Because many of these words were important to the authors' humanistic philosophy, they were borrowed into English, both to capture the flavor of the original and especially to avoid semantic loss. Otherwise, English had no native word by which to specify the particular object or concept used in the original. Sometimes translators and original writers deliberately introduced polysyllabic "inkhorn terms" into their works. They naively thought that English needed enrichment, believing the lexicon inadequate to convey the intellectual and aesthetic conceptions of the classics.

For instance, Sir Thomas Elyot italicized *devulgate* on one page of *The Governour* (1531), with *describe* and *education* a little later. What has happened to the first of the trio? Yet because of the successful careers of the other two words, we must be cautious in condemning such effort to add supposedly learned items to the lexicon. Some of these borrowings have survived; others never gained even momentary membership in Early Modern.

Or consider *translatum, synonymum, contrarium, diversum, proprium,* and *phrases.* The six words were borrowed directly from the Latin by Roger Ascham in *The Scholemaster* (1570). Though the first four of the words already had a naturalized English form by then, he still preferred italicized Latin to the English spelling. *Phrases* enjoys such high frequency today that its appearance in the Latin sextet seems out of place to us.

Language-Teaching Manuals

The Scholemaster, reprinted dozens of times, was probably the first serious book about the methodology of language teaching. It was but one of many works pouring from London presses that dealt with the teaching of a foreign language. Ascham expressed his purpose in the subtitle: "Or plaine and perfite way of teachyng children, to vnderstand, write, and speake, the Latin tong, but specially purposed for the priuate brynging vp of youth in Jentlemen and Noble mens houses." He wanted to guide

gentlefolk toward fluency in Latin, while teaching them basic educational principles.

In *The Scholemaster* the ideal nobleman is urged to study the liberal arts, of which grammar is an indispensable part. The grammar is naturally Latin and perhaps Greek. The pedagogical method recommended is *imitatio*. Thus the child learns by imitating Cicero. Ascham advanced some unsound premises that we still sometimes hear today. For instance, "The providence of God hath left vnto vs in no other tong, saue onelie in the Greke and Latin tong, the trew preceptes, and perfite examples of eloquence." Actually, the precepts and eloquence come from the writer, not from the language.

Despite such extravagant praise of the classical tongues, Ascham was basically nationalistic. This attitude was partly a reaction to the overreverencing of classical learning and languages that we've seen in the grammars and translations. It helped lead to a driving interest in the vernacular, contrasting with the attitude that English was somewhat barbaric and needed enrichment. Curiously, the nationalistic inspiration may have been classical. The Roman Senate required everyone to address the senators in their vernacular, Latin. They thought it beneath their dignity to permit any other language to be spoken in the Senate. Even in Greek and Asian provinces, Roman magistrates observed the rule. What difficulties the trembling, monolingual Egyptian peasant must have experienced when brought before the imperial judge!

Another nationalistic spokesman was Richard Mulcaster, who defended the power of English speech and writing in *The First Part of the Elementarie* (1582). In the thirteenth chapter he argued that "the antiquitie of our tung, the peples wit, their learning, and their experience" assure artistic capacity and the pedagogical qualities needed for the language to serve as the vehicle of children's instruction. His enthusiasm for the vernacular was echoed by some other scholars of the day. This was a kind of extension of the "dominance of English" observed in Late Middle English.

There were two other important language-teaching manuals. The first was by David Rowland, who translated *A Comfortable Ayde for Scholers* in 1568. He announced his book to be "full of varietie of sentences, gathered out of an Italian Authour." His novel approach to Latin teaching was to give an English sequence

like "But to retorne to our purpose," and to follow it by a handful of unidentified Latin quotations conveying the same thought.

The other was *The English Schoole-maister* (1596), by Edmund Coote. It went through an almost unbelievable fifty-four editions by 1737. His title reveals the pedagogical purpose of this introductory manual. Specifically, he wanted to teach "the most easie, short, and perfect order of distinct reading, and true writing." He mistakenly believed the language teacher to be somehow the judge of his language. All other native speakers of the tongue must bow to the teacher's linguistic pronouncements. Coote's view is counter to the actual equality of dialects, each dialect being adequate for the speaker's purposes.

Spellers

Coote's book is partly a speller. Occasionally there is attention to orthographic standardization of borrowings. Preference is accorded *malicious* over the Latin-influenced *malitious*, and the Latin-influenced *German* over the French *Germain*. There is invaluable information about intonation, especially syllabification. In one chapter, all junctures are marked with hyphens. Here's an illustration: "If you di-li-gent-ly ob-serve these things, you can-not erre in a-ny word of one syl-lable." Coote's last twenty pages constitute the first attempt at an English dictionary.

Three other spelling books were important. Already mentioned for its nationalistic spirit toward English, Mulcaster's work contains a plea for modest orthographic reform, based upon rules developed from "ordinarie custom." He criticized those who "fly to innouation, as the onelie mean, to reform all errors, that be in our writing." A fifth of the book is devoted to a "generall table," a spelling list of seven thousand common words. It begins with *abaie* (suffer) and ends with the new creation *zealousnesse*. It includes inflected forms, as in *abhorrest, abhorst, abhorreddest, abhordst, abhorreth,* and *abhorring*. The second and fourth of these are designated as contractions.

The other two books were primarily spellers, especially significant because of attempts to phoneticize English orthography. The purpose was to make the spelling uniform and representative of actual speech. John Hart based *An Orthographie* (1569) upon his own pronunciation, making the book invaluable to modern phonologists. He recognized that *ch, th,* and *sh* mis-

leadingly represent single phonemes. However, the three new symbols that he created for the sounds in his twenty-five letter alphabet weren't adopted by writers.

Alexander Gil's *Logonomia Anglica* (1621, second edition) contained information on pronunciation and grammar as well. He utilized diacritical marks to indicate the pronunciation. There's another point of interest about his book, one of the chief Renaissance works on the English language. The examples are necessarily in Gil's English transcription, but his exposition is in Latin. Therefore, all the grammatical terms are Latin. It would be an easy step a century later for grammarians to describe English on the basis of Latin. Although the specific attempts at spelling reformation by Gil, Hart, and others failed, the general result was the further freezing of late fifteenth-century English orthography.

Dictionaries

In addition to the spellers, some of which enjoyed many editions, the first dictionaries were compiled during the Renaissance. Robert Cawdrey extensively drew on Coote's *The English Schoole-maister* for *The Table Alphabeticall of Hard Words* (1604). In this he defined three thousand borrowings. Cotgrave's bilingual *A Dictionarie of the French and English Tongves* (1611) was discussed in the previous chapter.

A few other dictionaries appeared before Edward Phillips's *The New World of English Words* (1658). Its subtitle indicates Phillips's narrow purpose: "Or, a general dictionary containing the interpretations of such hard words as are derived from other languages; whether Hebrew, Arabick, Syriack, Greek, Latin, Italian, French, Spanish, British, Dutch, Saxon, &." Entries often include an etymology. For instance, a technical Greek word borrowed through the Latin about 1387 is defined as follows: "*Embolism* (Greek), a casting in, it is commonly used for the casting in of the day, which is added to Leap year." For a Danish combination first used about 1430, the entry is "*Embolned,* (old word) swelled."

In Phillips's preface there is recognition that Latin has contributed "the main body of our Army of forraigne words, these are so numerous that they may well be thought to equal, if not exceed the number of our ancient words." Needless to say, the book is a rich source for anyone seeking to determine whether a

given item has yet been borrowed. Let's try *Renaissance*. Alas, the very name of the period isn't listed. The reason is simple. This French loanword, originally derived from Latin *re* and *nasci* (be born), wouldn't come into English until about 1840.

In general, Renaissance dictionaries filled the need for definitions of "hard words," many of which were loanwords. The spellers had a comparable purpose. That is, writers needed a supposedly standard English transliteration of borrowings from languages like Greek or Arabic, in which there was a different writing system. We have to wait until Nathan Bailey's *An Universal Etymological English Dictionary* (1721) for the first attempt to list the whole lexicon.

Overall, writers' language preoccupation stemmed in considerable part from the borrowings. More than ten thousand items joined English in just two centuries, with about half surviving today. Renaissance writers almost systematically expanded the vocabulary. They wished to enrich it with inkhorn terms and other words, while building an inventory of synonyms permitting discussion of a subject without the author's having to repeat the same words again and again. The large expansion proves their success, particularly when we remember that one comprehensive Old English dictionary lists fewer than fifty thousand words. This was one of the few moments when man was able to change his language directly.

Loanwords

In the overall history of borrowing into English, Latin has supplied the largest number of loanwords. Next in order are French, Scandinavian, Germanic, and Italian. Such additions have helped give English the largest lexicon of any Germanic language, including Modern High German.

Let's consider the Renaissance borrowings in detail. Naturally the principal suppliers were the chief languages of the great Renaissance writers—Latin, French, Italian, and Greek. That very fact sometimes obscures exact derivation. Most contemporary French and Italian words came from Latin, just as many Latin abstract or scientific items derived originally from Greek. Arbitrarily, we'll name the language that seems to have been the immediate source of a given borrowing. There was also borrowing from tongues that had contributed earlier words—Scandi-

navian and Celtic. Dutch and High German added many. Constant trade and other contacts contributed hundreds of items from India. Arabic, Spanish, Persian, and Malay provided others. We'll consider the classical languages first.

Latin. This is the fifth time that we've observed Latin loanwords coming into English. Actually, there had been a continuous inflow since 1500. We find additions like

acumen	census	decorum	medium	pallor	series
appendix	circus	focus	militia	pollen	toga
arena	corona	fungus	octavo	quarto	vacuum

By now, the "Latin quality" should be familiar, like the *ap-* prefix and the meanings associated with law and learning. After 1660 there was a reduction in the volume of Latin borrowings, although some words were to be borrowed as recently as Late Modern. In chronological order, examples are *serum* (1672), *mica, lens, status, nucleus, bonus, tandem, ego, bacillus* (1883).

French. We've seen the massive French input into Middle English. In the sixteenth century alone that language added whole categories of words:

arts: grotesque, hautbois, rondeau, vogue
military: colonel, pilot, pioneer, trophy, volley
society: bourgeois, partisan, portmanteau, viceroy
others: cache, gauze, machine, piqué, promenade

We may or may not have cultivated a "French" sense by now, the capacity to recognize certain *-e, -eau,* and *-que* endings. Several of the words in the above list retain a spelling and pronunciation close to the original, causing many English speakers to mispronounce the items until corrected. Consider the stress in *cache, piqué, grotesque. Potage, rouge,* and *garage,* all with /ž/, were to be borrowed later. Does *siege* have /ž/? The French flood continued for centuries, with the largest number of borrowings in Late Modern. In some restaurants today, the whole menu may be in French.

Italian. Only a handful of Italian words came into Middle English. *Alarm, brigand,* and *million* arrived via French. During

the Renaissance, Englishmen regularly took the "Grand Tour." Not adventurous like Cabot or Sir Walter Raleigh, they brought back Italian items interpolated into their speech because of the necessity of making themselves understood in Florence and Venice. Why might these have been borrowed: *ballot, carnival, gondola, lottery, macaroni, manage, mountebank, scope?* Before concluding that the grand tourers' only purpose was to have a good time, we might remember that words change in meaning. How has *carnival* degenerated, for instance? If we don't know about the Italian military power of the time, other items may imply it: *bandit, parapet, salvo, squadron.* Terms from the arts, commerce, and society can easily be added. Indeed, the flow was to continue down to our century. It can be illustrated by two more sample groups:

architecture: balcony, cupola, piazza, portico, stucco
arts: cameo, canto, fresco, miniature, stanza

Greek. Greek was the fourth largest source of loanwords in Early Modern. The majority of these arrived via Latin, or through French as borrowed from Latin. Here are some loanwords borrowed directly from Greek: *bathos, cosmos, pathos, pylon.* Great numbers are technical words related to medicine and science, and many have been formed into polysyllabic compounds. For instance, ten direct borrowings have provided numerous other words, as can be easily proved by checking an unabridged dictionary:

archae-	(ancient)	archaeology
auto-	(self)	autobiography
bio-	(mode of life)	biochemistry
chron-	(time)	chronology
idio-	(personal)	idiograph
logo-	(word)	logograph
morpho-	(form)	morphogenesis
-philic	(loving)	photophilic
-phobe	(fearing)	Francophobe
-phone	(sound)	homophone

As in Middle English but to a vastly greater degree, word-building elements were also absorbed. Through compounding

with elements already in the language, these Greek additions created still more items. In particular, besides the ten bases listed above and many others, about two dozen Greek prepositions came directly into English as naturalized prefixes. Some of the words formed from them are hybrids. These prefixes continue to be enormously productive today, similar to the Old English pattern of preposition-verb (*fōre-bodian*). For examples, we can simply check scientific and medical journals. Most of the prefixes are listed below, with one of many possible examples for each:

amphi-	(around)	-theater	endo-	(within)	-derm
ana-	(up)	-gram	epi-	(on)	-demic
anti-	(against)	-toxin	hyper-	(over)	-tension
apo-	(away)	-cope	meta-	(among)	-thesis
arch-	(rule)	-enemy	para-	(from)	-phrase
cata-	(down)	-logue	peri-	(around)	-scope
di-	(twice)	-lemma	proto-	(first)	-type
dia-	(through)	-gram	syn-	(with)	-cope[12]

Some have phonological variants. For instance, *syn-* becomes *syl-* (*syllogism*) or *sym-* (*sympathy*) in certain distributions.

Scandinavian. Now let's consider languages discussed in earlier chapters, Scandinavian and Celtic. What familiar phonological characteristic do the first two of the following borrowings share: *cozy, keg* (from *cag*), *oaf, simper, skald* (poet), *troll?* These exclude items descriptive of Scandinavia, observable in travel posters today: *auk, Edda, fjord, kraken, saga, ski, skoal.*

Actually, we could list numbers of other Scandinavian loanwords, except that they occur primarily in northern British speech today. In that area, a kind of national language spirit developed after 1500, as illustrated by John Jamieson's five-volume *Etymological Dictionary of the Scottish Language* (1808). His etymologies frequently go back to Scandinavian, when not to Celtic, a fact verified by the new edition of 1966. Jamieson declared Scottish to be a separate tongue, as different from English as Portuguese is from Spanish. He has a point, if

[12] The word lists in this chapter and that on Middle English come primarily from the compendious lists in Arthur G. Kennedy's *Current English* (Boston: Ginn, 1935) and Serjeantson's *A History of Foreign Words in English.*

exaggerated. If we particularly rely on Scottish English in searching for Scandinavian loanwords, thousands more must be added to the total. However, the lexicon that we're describing at any period in the history of our language is theoretically that of all English dialects of the time, even though an American may not be as interested in words used primarily in Scotland or Wales as he is in words used in the United States.

Celtic. British English contains numbers of borrowings from Munster Irish, Scots Gaelic, and Welsh since 1100. By contrast, only about a dozen of these are really familiar in America:

banshee	clan	loch	slogan
blarney	colleen	shamrock	Tory
bog	leprechaun	shillelagh	whiskey

All but which two of these are from Irish? *Tory* has undergone some elevation, once meaning a Papist or outlaw. What does it denote today? Another has done just the reverse, *uisge beatha,* or water of life. What's the word? A third lends itself to folk etymology, a false derivation based on chance resemblance. Is "sham + rock" its source? Note that most of these words relate to things characteristically Celtic.

Germanic. Germanic was the only other group of languages to give substantial numbers of words to English. Indeed, the total contribution of Dutch and some of the other non-Scandinavian ones approaches the Scandinavian total. In *A Dictionary of the Low Dutch Elements in the English Vocabulary* (The Hague, 1939), J. F. Bense conservatively tabulates 5,079 loanwords from Dutch, Flemish, and Low German. Even after deduction of 1,149 obsolete items and of 1,581 combinations and derivatives, the figure is more than 2,300. The number shouldn't surprise us, in view of the close linguistic, cultural, and geographical affinity of English with Low Dutch. After all, the English-Frisian dialects were among the last of the West Germanic group to separate. There was constant intercommunication with the Continental lowlands over the centuries. During the Renaissance, the English revered Dutch and Flemish art. In North America, the Dutch and English colonists lived side by side for a time. Such contact also strengthened some of the remaining Germanic words

in English, or even rejuvenated an otherwise obsolete item by introducing its Dutch cognate. The following typify the Dutch loanwords:

commercial	military	nautical	other
hawker	bulwark	cruise	boor
isinglass	furlough	skipper	booze
mart	knapsack	sloop	easel
muff	tattoo	splice	poll
spool	wagon	yacht	sled

High German has contributed fewer words, such as *Junker, kindergarten, plunder, poodle, sauerkraut, schnapps, waltz.* Perhaps the largest number relate to mineralogy, as in *cobalt, feldspar, quartz, zinc.*

At this point we may be wondering whether there are international or scientific words. That is, a few items of a given language have almost necessarily been borrowed widely. How can anyone talk about German government or money without using words like *Bundestag* and *Deutschmark?* When we go to Germany, we usually eat sauerkraut. Another kind of illustration is the vocabulary of chemists, who use the technical terms *cobalt* and *zinc* around the world, regardless of their native speech. Such words enlarge the total English lexicon, but they change nothing else about the language. If *sauerkraut* and *schnapps* suddenly disappear from menus everywhere, English can comfortably lose their names without syntactic or phonological effects.

Indian. The languages of the subcontinent contributed about nine hundred bases, exclusive of thousands of derivatives, after the East India Company was chartered in 1600. As usual, many came via other tongues, and some soon died. By contrast, English supplied far more items to Indian languages like Telegu than it took. Most Indian borrowings derived from Sanskrit or Hindi. They describe objects, life, and activities previously unknown to Europe:

bandanna	cashmere	dacoit	jute	pundit
brahman	cheetah	dungaree	loot	thug
bungalow	chintz	guru	rajah	yoga

We might add the Tibetan *yeti*. Did Sir Edmund Hillary find one?

Arabic. The Crusades, together with the Arabs' excellent early studies in science and medicine, help explain why Arabic has probably contributed more words than any language not yet discussed. Most entered via Latin, French, or other tongues, as in *alchemy, alcohol, algebra, almanac, cotton, mattress, zenith.* Incidentally, *al-* means "the," as in *Alcoran* (the reading), aphesized to *Koran*. What's the aphesized version of *alchemist?* Here are some direct borrowings:

afreet	ghoul	henna	mufti	sash
Allah	harem	jinn	roc	sheikh
fakir	hashish	Koran	salaam	wadi

Spanish and Others. Spanish was a strong competitor with Arabic, perhaps with as many direct borrowings. Here are some up to 1660:

anchovy	armada	booby	corral
iguana	mosquito	mulatto	renegade

Two other languages contributed a few handfuls:

Persian: bazaar, pajamas, serai, shah, shawl
Malay: gong, orangutan, paddy, pangolin, sarong

For the Malayo-Polynesian family, Hawaiian borrowings could just as easily have been listed. Some were given in the second chapter.

We've witnessed the enormous number of borrowings in Early Modern. The quantity is partly explained by scholars' strong interest in English, demonstrated by the outpouring of grammars, translations, language-teaching manuals, spellers, and early dictionaries. The major change in the language was the lexical expansion, since syntax changed comparatively little after 1500. A few small surface structures were added or dropped here and there. The Great Vowel Shift did alter the vowel sounds of many words, as we'll see shortly. To complete our discussion of the Renaissance, we turn to a sample from a famous translation, the King James Bible (1611). By studying this and other ex-

amples, we can arrive at a description of Early Modern. The last third of the chapter will sketch Authoritarian English and Late Modern, since little happened to English in those two and half centuries, by comparison with the lexical and phonological changes during the Renaissance.

Structure of Early Modern English

Any doubt about the modern qualities of the language of this "Authorized Version" is quickly dispelled. It was used until 1952, when the Revised Standard appeared. Even today, many people prefer the King James Version. The forty-seven scholars' product makes an excellent sample for analysis, when bolstered by other Renaissance works, particularly Shakespeare's. To permit direct comparisons with the Old and Middle English versions quoted in the previous chapter (Texts D-E), we'll use the same passage, Genesis iii, 1–3. The verses of Text F are renumbered according to the linguistic structures:

Text F: Passage from the King James Version

(1) Now the serpent was more subtil then any beast of the field, which the Lord God had made, (2) and he said vnto the woman, (3) Yea, hath God said, (4) Ye shal not eate of euery tree of the garden? (5) And the woman said vnto the serpent, (6) Wee may eate of the fruit of the trees of the garden: (7) But of the fruit of the tree, which is in the midst of the garden, God hath said, (8) Ye shal not eat of it, (9) neither shall ye touch it, lest ye die.

Deep Structure

As usual, a first concern about deep structure is the word order. Even the surface structure is **noun** + **predicate** in all but two inversions. **T-ques** easily provides these, "hath God said" (3) and "shall ye touch" (9). It inverts the auxiliary before *God*, and the modal before *ye*.

The large number of prepositions, all developed from Old English, is striking. If we bother to count them, we'll find thirteen

in (F), as opposed to eight in (D) and thirteen in (E). The explanation is the preposition's steady gaining in syntactic importance since West Saxon. As long as there was a dative case, the prepositional presence was implied and didn't have to be filled. Loss of the dative inflection in Middle English meant that the preposition had to be included. The especially frequent one is *of*. There are three in (D), seven in (E), and ten in (F). Note that the three "of" phrases in (F6) require most of the optional prepositional phrases of the deep-structure rules.

Modals

The modal-auxiliary trend begun in Middle English picked up steam in Early Modern. In (E) there are three instances of *schulde*, but no occurrence of the auxiliaries. In (F) we find which four modals and three auxiliaries? Shakespeare used ten different modals. *Will* enjoys tremendous frequency in his writings. *Shall* and *dare* maintain a distant second place. *Can* and *may*, equally reputable, occur still less often. *Must* is infrequent. *Ought* and *gin*, which joined the group meteorically, are in their final moments as modals. Shakespeare uses them only once, in "you ought not walk" and "Phoebus 'gin arise." The structures in which these eight modals occur are accommodated by our previously expanded rule, **Auxiliary** → **Tense (Modal)** (*have* + *-en*) (*be* + *-ing*).

Do and *need* are defective modals in Shakespeare. That is, they cannot combine with auxiliaries to produce sequences like *"did have eaten" and *"did have been eating." There are two problems in the analysis of *need*. The first is simple. A homonymous form is a verb, as in Shakespeare's "you shall never need to fear." This sentence also illustrates the modern use of **T-to,** which links two verbs. We can exemplify the other problem with more of his structures:

(1) You will needs buy
(2) I needs must lose
(3) Thou hadst need send

Note that (1, 2) seem to have two modals each, and that their order varies. If true, both generalizations contradict the modal principles. That is, there can be only one modal per sequence,

and the order in the sequence is inflexible. If there is an auxiliary, it must follow the modal. Alas, we seem to have a modal following the auxiliary *hadst* in (3)! An arbitrary decision can resolve the problem. *Needs* and *need* in structures with *will* and *must* will be considered as adverbs meaning "necessarily." Thus our modal principles are maintained. We thereby prevent things like *"might could eat," where *might* can't be analyzed as an adverb.

By 1660, with the demise of *ought* and *gin,* English had only six regular modals and the defective *do* and *need.* Within another century, the modal *do* was no longer to occur in ordinary prose. *Need* has never quite died, although its restraints have become still more severe. *Dare,* a fair-haired boy in Shakespeare's works, would unaccountably share the same fate. *Need* and *dare* occur as modals today in only a few specified structures—negatives ("I need not go, I dare not go") and questions ("Need I go?" "Dare I go?"). We don't even use their past form, as in *"Needed I go?" although "Durst I go?" is only archaic. *Daresay* remains a verb in a kind of frozen compound of **modal + verb.**

These restrictions upon *dare* and *need* usually prompt modern linguists to exclude the pair from the group of five modals developed out of Old English: *can, may, must, shall, will.* The five participate fully in the **Auxiliary** rule. *Must* has a small blemish, in that the past form has **ø.** Its four compatriots all take a dentalized, distinct past. Has the dental subclass ultimately triumphed over them too, as it generally has over the once large but outnumbered vocalic verbs? No, because the five modals aren't verbs and so can't have a past participle. Anyway, the modals always had /d/ in the preterit.

Progressive Passive

Let's complete the history of modals in Modern English. We've seen the **Auxiliary** rule expanded at least by Late Middle English. The optional unit "have + -en" was needed for Chaucer's sequence "should have been slain." **T-pass** transformed the structure into the desired passive. He additionally needed "be + -ing" for "should have been waiting." It was only a matter of time before **T-pass** would be applied to a sequence containing "be + -ing," to use more of the potential of the **Auxiliary** rule. This new expansion gave English the largest variety of modal-auxiliary structures in its history up to that time.

The first recorded appearance of the progressive passive was in 1769: "There is a good opera of Pugniani's now being acted." After the words are inserted, the basic deep structure is

X **Tense** be + -ing act an opera now

Then **T-infl** and **T-pass** operate. **T-del** deletes the unidentified, unnecessary "by X," giving

An opera **present** be + -ing be + -en act now

Three applications of **T-aux,** besides the **P** rules, result in "An opera is being acted now." Actually, two embedded structures must provide the needed modifiers *good* and *of Pugniani's,* and another transformational rule adds *there.* Thus English speakers began to apply **T-pass** to the *-ing* portion of the **Auxiliary** rule, available for several centuries. Almost certainly the structure came into the language before 1769, since writing is ordinarily more conservative than speech. Yet purists attempted to discredit it for a century before grudgingly acknowledging its utility.

Lexicon

Next in the generative-transformational model is the lexicon, in which there was a huge expansion. The addition was sometimes at the expense of Germanic words. The great majority of Germanic words obsolete today were no longer being used by 1660. For example, nine words in the Old English passage in Text A are obsolete; four, in (D). In (F), however, not one has died.

The survivors, both native and borrowed, continued to change substantially in meaning. Part of the explanation may be that the few early dictionaries were pitifully inadequate. A prolific, varied writer like Shakespeare, possessing one of the largest vocabularies of the day, had no convenient, unabridged dictionary to turn to. He presumably relied upon his understanding of the meanings of a given word as he heard it in speech; so the usual jousts between words went on. For example, he used the prepositions *among* and *amongst* interchangeably. There are three occurrences of *to* for every single use of *unto* in Shakespeare, as

opposed to the heavy dominance of *vnto* in the Bible. *To* precedes an infinitive, of course. In the Bible there are some "to" constructions with pronouns, as in *to thee-ward* and *to us-ward.* Dating from Middle English, these were to become archaic. Certain other Old English-derived suffixes persist today in adverbial combinations like *downward, inward,* and *toward.*

Prepositions

The meanings of prepositions in Modern English were more limited than in Middle English. There was still overlapping of meaning, though reduced. In Text A, *on* conveyed "on" and "in." Actually, Old English *on* could mean "on, at, during, in, into, among, against," in the sense of time, place, manner, circumstance, and condition. We find no particular ambiguity in the prepositions in (F). Today, *to* would be preferable to *vnto,* and perhaps *from it* rather than *of it* in (8). Consider these structures from Shakespeare, of which only (1) and (10) contain the prepositions that we would use today:

(1) let it lie on my head
(2) Be not jealous on me, gentle Brutus
(3) They do not point on me
(4) But thou wilt be aveng'd on my misdeeds
(5) lest he catch cold on's feet
(6) I were better to be married of him
(7) Whom I left cooling of the air
(8) I wonder of their being here together
(9) He came of an errand
(10) only give me so much of your time
(11) if you suspect me in any dishonesty

Each of his five sequences with *on* requires a different preposition today. Likewise the five with *of* and the one with *in.* What prepositions would we use? Since Early Modern, each preposition has continued to narrow its range of denotations, with little overlap today. This telescoping has severely tightened word-choice. For example, "in" is almost always conveyed by *in* now, seldom by *of* or *on.* Shakespeare, however, could use any of the three in certain structures.

By Late Middle English, perhaps through the influence of French *concernant* and *touchant,* the number of prepositions was increased by at least four. They are formed from the participle. In chronological order of the written records, they are *touching, notwithstanding, during, concerning.* "Natives" like *in, of, on,* and *with* continued to dominate in frequency.

Transformations

The transformational rules are our next concern. Except for the inflections surviving today, almost all the remaining Old English ones died between 1500 and 1660. Besides the effect on **T-infl** and **T-agr,** one apparent consequence was a further tightening of word order. That is, the listener had to differentiate subject from object and the like by their position, rather than by their inflection. Renaissance surface structures can be characterized as luxuriant, enjoying the last moments of the syntactic freedom inherited from Late West Saxon.

Nouns

Let's consider **T-infl** and **T-agr** first, followed by **T-do, T-not,** and **T-contr.** Still more nouns adopted the *-s* plural through **T-infl,** largely finishing the process underway since Middle English. After Spenser, for instance, *treen* became *trees.* In Shakespeare we find *devils, fleas, hands, goats, knees, names, sins, words. Eyen* or *eyne, hosen,* and *shoon* were to hang on a while longer. Undaunted, *oxen* stolidly observed its closest friends changing. It was in no danger, as evidenced by Shakespeare's careful discrimination: "Oxen and wainropes cannot hale them together" and "like sheep and oxen." *Hose,* always plural, was later to lose the *-n.*

A certain stubbornness is seen in *month, pound,* and *year.* There are many structures like "five hundred pound a year, twelve month old, twelve year old." Needless to say, the last two persist today in certain distributions, where, unlike Shakespeare, we can't use the *-s* form. Usage requires "five-day week, seven-year itch, four-minute mile, ten-foot pole." Shakespeare preserved *deer, sheep,* and *swine* as "zero" forms for us. *Folk* and *kind* have an *-s* plural.

Determiners

T-agr continued the simplification begun after 1100, as applied to determiners, personal pronouns, adjectives, and verbs. Among the few remaining determiners, *a, an,* or *the* is selected for a [+ sing] noun. *The* can also be [− sing]. There's a little overlapping between *a* and *an,* especially before *h*. Consider the word *history*. Does it take *a* or *an?* In general, Shakespeare's practice is ours today, with *an* before vowels, and *a* before consonants including *h*.

Personal Pronouns

T-agr retained most of its Old English matching functions for personal pronouns, which hadn't changed much. Shakespeare's first person is somewhat regular. *My* usually precedes a consonant; *mine,* a vowel. He writes "my friend(s)" and "mine eye(s)," but also "my education" and "mine loved darling." *Mine* occurs in the only two structures where we use it, "She was mine" and "friend of mine," as well as in "brother mine."

In the Bible and often in Shakespeare, second-person number is discriminated. *Thou, thy, thine, thee,* and *thyself* are [+ sing]. *Ye, your, yours, you,* and *yourselves* are [−sing]. Note that Shakespeare writes "thou drunkard" and "ye choice spirits." Second-person singular requires *-est* for the verb, as in his "thou sayest." By the late Renaissance, *you* was frequently [±sing]. Nevertheless, many people were to distinguish number as late as the eighteenth century through "you was" and "you were."

The neuter *his* underwent aphaeresis. The ultimate result, *its* (analogized on the model of the genitive case of nouns with *-'s*), balances our *her* and masculine *his*. However, we find "My life is run his compass" and "nature cannot choose his origin" in Shakespeare. These structures don't seem to be personifications, which might require a masculine or a feminine pronoun. He often uses *it* as the genitive form, as in "of it own kind" and "to it own protection."

There was also a kind of group genitive. Consider Chaucer's title, "The Wyves Tale of Bath." In Early Modern the structure had become "The Wife of Bath's Tale." The original explanation may lie in the high frequency of the given noun unit, as in "the King of England," which constantly recurs in writings. How

natural for "the King of England's subjects" to be Shakespeare's genitive. "His and Jack's money" and "somebody else's" are comparable examples.

The reflexives complicate **T-agr** a bit, for they generally observe number. We find *ourselves* and *themselves*. Still, Shakespeare employs the older *ourself* dozens of times. He discriminates number in "Madam, yourself are not exempt" and "now show yourselves men." Among the interrogative and relative pronouns, *which* commonly refers to people, other life, and objects, as in the "beast which" and "tree which" of the Bible (E). Perhaps the most famous *which* is the biblical "Our Father, which art in heaven." Although Shakespeare sometimes chooses *who* for a structure like "who steals my purse," *which* is just as frequent. For instance, "This gallant which thou seest." In the accusative case, he uses *who* a little more than *whom*.

Adjectives

T-agr could no longer be applied to adjectives, once the apocopation of /-ə/ erased the last vestige of adjectival inflection. In Chaucer we observed *smale, sweete,* and *yonge*. In Shakespeare the spellings are our modern ones, so that even the orthography proves the loss of the second syllable. As for degree of comparison, we've seen the modern comparative since Old English *gēapre* in (D), through Wycliffe's *feller* (E), to "more subtil" (F). In Coote's subtitle we saw the superlative "most easie, short, and perfect." His problem still isn't resolved. Note the total number of possibilities in a comparable modern example, where both adjectives are superlative: *"easiest and most happy," "easiest and happiest," *"easiest and happy," *"most easy and most happy," *"most easy and happiest," and *"most easy and happy." In general, *-er* and *-est* continued from Middle English as the usual degrees in Early Modern. This fact explains Shakespeare's *extremest, honester,* and *rascalliest*. However, *more* and *most* were gaining in frequency. Along with his "more serious," we discover "more fast" and "more great." Doubling of the degree isn't uncommon in Early Modern, as in these words combined with *more: better, braver, corrupter, proudlier*. Shakespeare often couples the superlative *most* with the adjective forms *best, unkindest, worst, worthiest*. Today, of course, there's no longer a choice among *"most best," *"most good,"

and *best*. Doubling has died in standard speech. Adjectives of one or two syllables usually take *-er* and *-est;* the rest, *more* and *most*.

Verbs

Verbs, including the application of **T-infl** and **T-agr** to them, have been shown to be quite complex over the centuries. A few verbs didn't lose the inflectional *-n* of the infinitive until Late Middle English. Some past participles were to retain their *-n* into the eighteenth century—*sitten, spitten,* and *stridden,* for example. What can we say about *stridden* or *hidden* today? By the Late Renaissance, *-s* had generally replaced third-person *-eth,* except in the conservative Bible. However, the numbers of *-eth* in Shakespeare suggest that conservatism isn't necessarily the reason for its persistence. Consider Portia's speech for mercy, in which *droppeth* and *blesseth* occur with *gives* and *takes*. Throughout the Renaissance we find *hath* and *doth* not yet yielding to our *has* and *does*.

As usual, the largest alteration in verbs was in the vocalics, which continued to shift to the dental ranks. Several examples were given in the previous chapter. Shakespeare demonstrates dental victories in *crowded, dreaded, flayed, glided, helped, mow'd, waded.* To complete the history of the approximately 312 vocalic verbs of Old English, we can observe that only half of them survive today. About half of these are now dental. Here's a list of Middle English vocalics that were obsolete by Early Modern, with one example from each subclass: *agrīsen* (frighten), *bēden* (order), *limpen* (happen), *hęlen* (conceal), *quēþen* (from OE *cweðan*), *drazen* (draw), and *fangen* (seize). Some dentals died too, as in *dęren* (injure) and *wēnen* (hope). Among the other verbs, *wit* was finally beginning to fade. Shakespeare employs it only about a dozen times, as opposed to many occurrences of the ultimately victorious *know. Go* and the copula *be,* by contrast, enjoy high frequency.

So does *do*. Here's a biblical example: "We do you to wit." The shift in meaning, plus the archaism of *wit* today, may trouble us momentarily, before we translate the sentence as "We cause you to know." The popularity of *do* probably helped its modal homonym to die by mid-eighteenth century. Interestingly enough, the modal enjoyed its highest frequency until about 1660, when

T-do began to affect surface structure. This was an important new rule.

"Do" Transformation

Remember the **T-ne** result in Old English, as in "hē ne cwæð"? It's syntactically unchanged in a typical Shakespearean sentence, "It not concern'd me." To modernize Shakespeare, we need a *did*, for his usual structure is "Goes the King hence today?" Again, this contrasts with our "Does the King go hence today?" On the other hand, he anticipates the **T-do** rule in "Did not I tell thee yea? hadst thou not order?" In this remarkable sequence he uses the rule in the first question, but not in the second. The second one illustrates the pattern of negative questions until the Restoration, a pattern closer to Old English **T-ne** questions than to modern **T-not** ones. That is, Shakespeare's usual structure is "Went'st not thou to her?" rather than "Didn't you go to her?"

How did Shakespeare generate "Goes the King hence today?" Evidently **T-ques** could also permute the verb along with the tense:

NP **present** go hence today → **present** go NP hence today

T-aux and a **P** rule then produce his sentence. If a negative question is desired, the transformationally inserted *not* can either be inverted with the tense or else stay behind. The ultimate result then is either "Goes not the King hence today?" or "Goes the King not hence today?"

How do we derive our "Does the King go hence today?" First we create the deep structure and insert the words. Then the succession of **T** rules is applied:

The King **Tense** go hence today
The King **present** go hence today **(T-infl)**
present The King go hence today **(T-ques)**

At this point we attempt to apply **T-aux,** but we can't because *the* isn't a modal or auxiliary or verb. Therefore, we must apply **T-do,** which simply inserts *do* in front of **present.** A **P** rule later gives "Does the King go hence today?"

This single option contrasts with Shakespeare's three. He could use "Goes not the King hence today?" and "Goes the King not hence today?" The looming entry of **T-do** permitted his third option: "Does not the King go hence today?" Thus we see that word order has been significantly tightened since Early Modern. If we want a contraction, **T-contr** obligingly produces "Doesn't the King go hence today?"

Two Other Transformations

Actually, **T-not** and **T-contr** haven't changed much since Middle English. Remember the biblical passage, "Ye shal not eate" (F4)? **T-not** simply inserts *not* before the verb, whether or not there is a modal and/or an auxiliary:

Ye **present** eate → Ye **present** not eate
Ye **present** shal eate → Ye **present** shal not eate
Ye **present** have + -en eate → Ye **present** have + -en not eate

The first typifies Shakespearean structures like "I not doubt; I not deny." **T-aux** can't apply because there's no **suffix** + **verbal** sequence, and **T-do** hasn't really entered the language yet. Such structures are conducive to the multiple negatives found in much Early Modern prose. A Shakespearean couplet will illustrate:

> Is't not enough, young man,
> That I did never, no, nor never can?

If alert, we'll have noticed contractions like the *is't* above, as well as *aveng'd, concern'd, went'st,* and *on's* in previous examples. A quick glance through Shakespeare's plays will unearth *lookt, pourd, in's, on't,* and *to't,* among many others. He presumably wasn't trying to create new "words," but only to indicate his characters' speech. Thereby linguists are remarkably assisted in reconstructing Early Modern phonology. Clearly, apocope has been completed in the verbs just listed. When *abhorred* and *picked* are his spellings in such contexts, we guess that apocope hasn't occurred yet. Why doesn't **ag'd* occur as an adjective?

T-contr enjoyed expanded functions in Early Modern. In Old English we observed it in negatives like *nolde* and *nǣfre,* where

it was applied after the operation of **T-ne.** Writers then discovered its utility in helping them depict characters' dialect or just in phonetically spelling contracted words. The first record of a contracted modern modal negative is *mayn't* (1652), followed by *can't, an't,* and *ben't* over the next fifty years. *Ain't,* originally a telescoping of "are not," appears in a sentence by Fanny Burney, Dr. Johnson's friend, in 1778. Still stigmatized socially today despite widespread use in popular speech, *ain't* is also used for "am not" and "is not." *Isn't* is first recorded in 1895. *Won't* derives from "will not." All these contractions require the prior application of **T-not.** A related form is *ne'er* (or *e'er* and *o'er*), which illustrates syncope of the *v.* Having been a frozen compound for many centuries, *never* was an independent word in the lexicon by 1660, no longer requiring **T-not** for its existence.

Writers also utilized **T-contr** in positive structures. Shakespeare frequently used *that's, 'tis, 'twill,* and *'twere.* The middle two, archaic today, were replaced by which words? *'Twere* died. *It* was a real Renaissance liberal, appearing as *do't, in't, is't, on't, to't,* and *was't,* as well as the Scottish *ward* (were it). All these persist today in ordinary rapid speech, although they're no longer written so except as dialect. Overall, Renaissance rapid speech, as depicted in writing, made extensive use of **T-contr** for the first time in the long history of the rule. Frequency has probably risen since then.

Phonology

Now we turn to phonology to complete the generative-transformational model of Early Modern English. Let's first finish the history of the consonants, together with the "unshifted" vowels, before describing the Great Vowel Shift. Chaucer's /x/ changed to another fricative or died. A word spelled with *-gh* usually indicates the recent death of /x/ therein and publicly marks the grave. Consider *drouth* and *rough.* In the first, /x/ has become /θ/ or /t/; in the second, /f/. This fricative has vanished from *night.*

Two-consonant clusters continued their simplification. Each of /gn-, kn-, wr-, wl-/ has finally lost the initial phoneme. The event was quite late, as attested by words still spelled with all but the *wl-* cluster. Some terminal clusters were also simplified, the orthography again indicating the recency of the loss. What has

happened to /-mb/, in *climb?* Analogically, *b* was sometimes added to a spelling. For example, Middle English *crume* became Early Modern *crumb,* although /-m/ wasn't expanded to /-mb/. Dr. Johnson played safe, listing both versions in his *Dictionary* (1755). Or consider /-lk/, as in Chaucer's /fɔlk/. It changed to /k/, like *balk* and *walk.* What about *milk* and *help?* In Early Modern we find the beginning simplification of /nd/ to /d/. Thus *launde* became *lawn.* Later, *and* was to be reduced to /æn/, then /ən/. In context, the otherwise terminal /d/ is often dropped today, as in *land* and *garland.*

The short vowels in accented syllables changed relatively little after Middle English, except for some encroachment by /ə/, after its general conquest of unaccented syllables. In fact, neutralization wasn't completed by the late sixteenth century in all words containing /ə/ today. An easy way to prove this generalization is to observe words rhymed in poetry. In Richard Tottel's *Miscellany* (1557), we find *bloom-come.* Renaissance poets also rhyme *love* and *prove,* as in Ben Jonson's "Come, My Celia" (1606). Presumably *come* and *love* don't have /ə/ yet.

Consider Spenser's iambic couplet from *Amoretti* (1595):

> Leaves, lines, and rymes, seeke her to please alone,
> Whom if ye please, I care for other none.

Not only does *none* not yet have /ə/; the meter tells us that the *-e* has been silenced by apocope. On the other hand, Chaucer's /ʊ/ in *sonne* was already /ə/ in Early Modern *sun.*

A second important vocalic change was the unrounding of the high-front pair /ü:/ and /ü/. Soon after Late West Saxon, these became /i/ and /ɪ/, respectively. Later, the *y* was replaced by *i* in the spelling. Respellings like *din, dizzy, kin, kiss, thin,* and *sixty* conceal the former rounding. However, not all modern words with *y* once had a rounded vowel. *Gypsy* has undergone aphaeresis from the borrowed word *Egyptian,* and *sympathy* is a Latin loanword. The adverb suffix *-ly* is a respelling of *-līc.* Check the origin of *byre* (cow barn), *rye, tyke, type.*

As for the spelling of vowels, we've observed initial *v,* as in the biblical *vnto,* contrasting with *u* in other distributions. Later, *v* was to become a consonant. Since *j* and *i* weren't to be "separated" until the seventeenth century, with *j* the consonant and *i* the vowel, early Renaissance books employed the two letters

interchangeably. Some writers were to do so into Late Modern. The letter *w* is formed from "double *u*," hence its name.

Great Vowel Shift

The Great Vowel Shift, one of the most important changes in 1500–1900, wasn't to be completed in numbers of words until the eighteenth century. Consider these rhyming items in Pope, who died in 1744: *array, away, bay, day, lay, obey, say, sea, way.* Which one hadn't yet shifted? The spelling, as we've said before, often first identifies a word as being a possible GVS product. When we test Pope's items by pronouncing them, we discover /i/ in *sea.* As his other eight items have /e/ today, his pronunciation of *sea* must have been /e/ too. The *-ea* spelling of the word tells us that it once had a lower vowel. Actually, it was /æ/, if we recall Old English *sǽ.*

Let's use Spenser and Pope for two more tests from poetry. Spenser rhymed *gate, hate, late, retreat.* How do we know that the second vowel of *retreat* wasn't yet raised to /i/? Pope paired diphthongs in *wine* and *join.* We might jocularly remark "Now I know why so many people use the dialectal /jaɪn/ today." The answer, of course, is that Pope's pronunciation was /jaɪn/, which rhymed with the already shifted /waɪn/.

These examples should simplify the potentially complicated GVS for us. The set of long vowels simply moved forward and upward at least one "place," according to the chart below:

	front		**back**
high	i		u
	ɪ		ʊ
mid	e		o
		ə	
	ɛ		
	æ		ɔ
low		a	ɑ

We can start with /ɑ/, the low back vowel. Thus the original /ɑ/ in *name* moved up to /æ/, then became /ɛ/ before settling in as /e/. Now consider /æ/, the lowest front vowel. Old English /sæ:/ climbed the ladder to the very top, to /i/. Next

recall Chaucer's /hɛ:θ/, which also went to the top, to *heath*. Now his /sle:p/, which did too, in *sleep*. What if the word already had /i/, as in the second vowel of Chaucer's *inspired?* The /i/ of *-spir-* couldn't go any higher, or it would become a consonant. That is, when the tongue goes still higher, the air passage begins to be obstructed, as opposed to the nonnarrowing required for vowel production. The sound produced is then a glide, fricative, or stop, depending upon the degree of narrowing of the passage. Therefore, the /i/ in *inspired* became the diphthong /aɪ/. As long as the word wasn't borrowed after about 1500, the spelling usually informs us of the shift, as do all our examples above. The same is true for *sweet*. It's pronounced with /i/, but spelled with the "lower" *ee*.

The back vowels shifted similarly. Remember the Old English /stɑn/, which moved up to /ɔ/ in Middle English? It climbed up to /o/ in Early Modern, giving our *stone*. Next recall Chaucer's /rotə/, which went up to /ʊ/ or /u/, depending upon the dialect, for *root*. Or his /ʊ/ in *yonge*, which moved—whoops, it was neutralized to /ə/. Sorry, it was never a long vowel. As the /u/ in Chaucer's /šurəz/ couldn't go any higher without becoming a consonant, *showers* has the diphthong /aʊ/ today.

Authoritarian English, 1660–1800

Finally, we can turn to the last two periods of Modern English, Authoritarian English and Late Modern, when the lexicon experienced considerable expansion and changes of meaning therein. The main themes of Authoritarian English have already been introduced. There was continuation of the Renaissance emphasis upon learning, accompanied by a keen interest in language, particularly English. The swelling flood of grammars, some dictionaries, and spellers contributed to further "standardization" of the prestige dialect as developed from London Standard. Underpinning these was the classical requirement of logical, philosophical thinking.

Efforts after the Restoration were more and more directed toward deliberate regularization. Scholars attempted to describe English according to Ciceronian and other structures, as analyzed in the numerous Latin grammars. When English structures didn't match, as was naturally true for most, they were forced into a Latinate description. Logic was heavily used, with argu-

ments for systematic arrangement of the language, which, as we know, aren't relevant to language. For example, the plural form of *ox* isn't determined by logic. Nor do we sometimes forget and use a double negative because of disordered thinking. The particular structures simply evolved over the centuries, as we've seen.

Appeal to Authority

In short, the period 1660–1800 can be characterized as "the appeal to authority," as linguists often term it. When the grammarian found an English structure that he approved, he cited it as a model. If, in the very next line, he found one of which he disapproved, he might quote it as a warning of how even the best writers could stumble. His erroneous belief was that people like himself can and perhaps should dictate about language matters to other native speakers of the tongue who, alas, aren't qualified. Thus he made himself the arbitrary judge. This "appeal to authority" led to attempts to establish an English academy and to the development of a prescriptive attitude in grammars and dictionaries.

Academies

We need to look back to the two first academies. The Italian Accademia della Crusca was founded in 1582 to purify Italian. Considering that the language had derived from Latin several centuries earlier, the members must have had quite an ordeal in determining the non-Italian words. Nonetheless, they published a kind of national dictionary in 1612, based upon Tuscan speech. A fourth edition appeared in six folio volumes in 1729–38.

Meanwhile, the French had developed their own academy by royal charter in 1635. Their dictionary of terms from the arts and sciences came out in two folio volumes in 1694. Suppose we contrast these two lexicographic achievements with the English situation, despite the linguistic misconceptions underpinning the idea of an academy. The first attempt in England to list all the words in the lexicon wasn't made until Bailey's dictionary of 1721. Yet English scholars, immersed in Greek and Latin philosophical grammars, were fascinated by these Continental academies. Paul Pellisson-Fontanier's history (1653), after sev-

eral French editions, was translated as *The History of the French Academy* (1657).

Among many writers, three of the greatest English literary figures of the time fervently advocated an academy to "fix" the language. Dryden, Defoe, and Swift urged its purification and refinement. In the prefatory letter to his play *The Rival Ladies* (1664), Dryden explained:

> I have endeavour'd to write *English,* as near as I could distinguish it from the Tongue of *Pedants,* and that of affected Travellours. Only I am sorry, that (speaking so noble a Language as we do) we have not a more certain Measure of it, as they have in *France,* where they have an Academy erected for that purpose, and Indow'd with large Privileges by the present King. I wish we might at length leave to borrow Words from other Nations, which is now a Wantonness in us, not a Necessity.

He undoubtedly had a hand in the Royal Society's decision a few months later to establish a committee for the purpose of improving the language. Nothing really happened therefrom.

In 1698 Defoe devoted one article of *An Essay on Projects* to a proposal for an English academy consisting of thirty-six members. To be approved by William III, the group would have broad intellectual powers. It would sanction usage, while exposing others' shortcomings: "They should preside with a sort of judicature over the learning of the age, and have liberty to correct and censure the exorbitance of writers, especially of translators." Like the French Academy, the proposed group would be "the allowed judges of style and language." Coinage without permission would be as criminal as the private coining of money. Of course, such a thought is naive. A writer, then or now, would hardly request an academy's permission to use a sprightly new word that he has just coined to evoke a desired image and meaning. After all, one characteristic of Renaissance authors, especially Spenser, was the introduction of Chaucerian terms and new compounds. Yet amid his call for refinement, Defoe lauded English as "the noblest and most comprehensive of all the vulgar languages of the world."

Swift presented the most powerful argument for an academy. His letter to the Lord Treasurer was published as *A Proposal for*

Correcting, Improving, and Ascertaining the English Tongue (1712). Assuming himself spokesman for "all the learned and polite persons of the nation," Swift complained about the extremely imperfect language. Latin, he said, "arrived at great perfection, before it began to decay." French was beginning to decline. By contrast, English was "not arrived to such a degree of perfection, as to make us apprehend any thoughts of its decay; and if it were once refined to a certain standard, perhaps there might be ways found out to fix it forever." Writers should be condemned, he felt, for using abbreviated spelling like *drudg'd, disturb'd,* and *rebuk'd* (as Shakespeare often did to indicate apocopated words in dialogue). Specifically, English should be reformed by a select committee. Following the French Academy's practice, the committee could utterly discard certain words, correct the many needing attention, and perhaps restore some obsolete ones. Swift's grand purpose, openly admitted, was a method "for ascertaining and fixing our language forever, after such alterations are made in it as shall be thought requisite."

A Proposal bore initial fruit. More than twenty persons of both parties were selected by the Lord Treasurer and Swift. Then the impetus slowed. Perhaps an academy would have been royally sanctioned had the sympathetic Queen Anne not died in 1714. George I, who ruled for the next thirteen years, was like William the Conqueror in that he knew no English. Unlike the Norman, he didn't attempt to impose his native German on the country. Anyway, objections were developing. People stopped talking about the academy after a time. However, such proposals from these three intellectual spokesmen and others intensified the drift toward authoritarianism in the English grammars of the day.

Grammars

Swift's comments about decay and the desired "fixing" of a language derived from a belief, shared by many grammarians, that there was once a logical, pure linguistic structure. The term they employed was *universal grammar.* Latin, they believed, fairly well preserved its original purity until its last years, though there was continuing controversy about its replacement by the "vernacular." Modern tongues like French and English, however, were beginning to decay. If the decayed portion could be

cut away, only the pure structure would remain. Then if there could be periodic judgment of likely future changes, the "harmful" ones being barred from entry into the language, the purity of English would be preserved forever. Indeed, all changes would probably be harmful. This naive, erroneous belief of some scholars was antithetical to Dr. Johnson's doctrine of usage, as we'll see. Clearly, it's also different from the generative-transformational theory. In the latter, language changes are recognized as inevitable but aren't judged, being described by rules of *particular grammar.* Almost everything written about English in this book is naturally from particular grammar.

Bishop Lowth

Let's consider three major eighteenth-century grammarians, the first of whom was easily the most prescriptive. Bishop Robert Lowth wrote *A Short Introduction to English Grammar* in 1762, and its twenty-two editions or so made it one of the most influential language books of the day. In his opening paragraph he commended the polish, refinement, and lexical enhancement of English. Still, "it hath made no advances in Grammatical accuracy." No effective method had been found to redress Swift's grievance in *A Proposal* fifty years before. Lowth insisted that English isn't irregular and capricious, but is easily reducible to a system of rules. Indeed, he emphasized the systematic qualities of the language in his 186-page book. Men mistakenly assume, he said, that they have a competent knowledge and skill, since they can't acquit themselves properly in the language. They need to study a grammar so as to be able to express themselves with propriety and "judge of every phrase and form of construction, whether it be right or not." Overall, he approached English rather philosophically, with Hebrew, Greek, and Latin his frequent models.

He condemned numerous constructions, thereby initiating the attacks upon them that would gather strength into mid-twentieth century. His tone was dogmatic, seldom including qualifying words like *possibly, perhaps,* and *may.* Thus he included Addison, the Bible, Dryden, and Swift in a long list of "improprieties" in prepositional usage of the sort we've seen in Shakespeare. Specific attacks are made on the double superlative

("most highest"); violations of the special degrees of *good, bad,* and *little* (Dryden is chastised for his *worser*); "you was" (poor Addison and Pope are lashed); and confusion between *sit* and *set, lie* and *lay, fly* and *flee.*

As these represent superseded forms for us, we may not be disturbed about Lowth's attacks. However, in his day they were still prevalent. His authoritative tone is echoed by a few purists today, and it is here that we can object. His statement that "Two Negatives in English destroy one another, or are equivalent to an Affirmative" is faulty logic. Nevertheless, continued, current injunctions against the double negative may eventually help kill that development from Old English. In a quick oversimplification, we can say that several such structures have been extracted by modern prescriptivists. These are cited as "thou shalt nots" that all writers must avoid at the risk of a failing composition or worse.

Lowth also told the writer or speaker what to say, besides what to avoid. For instance, *will* in the first person "promises or threatens; in the second and third Persons only foretells; *shall* on the contrary, in the first Person simply foretells; in the second and third Persons commands or threatens." Elaborate explanation and illustration accompany this Delphic pronouncement. *Hidden, holden,* and *bidden* are said to be correct; *hid, held,* and *bid* are corruptions. Which forms won?

He clucked his tongue at "custom" (usage): "We should be immediately shocked at *I have knew, I have saw, I have gave,* &c: but our ears are grown familiar with *I have wrote, I have bore,* &c. which are altogether as barbarous." Exposing one notable's slip, he ruled: "It ought to be *who,* the nominative case," rather than *whom,* the objective. *Who* refers to people; *which,* to things. Formerly, "they were both indifferently used of persons," as in "Our Father, which art in heaven." Pronouns following *be* forms should always be nominative, as in "I am he." The infinitive is the exception: "You took it to be him." Recognizing that a preposition often concludes an informal sentence, Lowth nonetheless recommended its placement before the relative pronoun as being "more graceful, as well as more perspicuous; and agrees much better with the solemn and elevated Style." His illustration-judgment is " 'We are still much at a loss, *who* civil power belongs *to*.' Locke. It ought to be *whom*." Thus nodding occasionally toward usage, Lowth can't be labeled as purely prescriptive.

Dr. Johnson

Dr. Johnson, the second grammarian whom we'll consider, was much more modern in attitude. His lengthy preface to *A Dictionary of the English Language* (1755) is a kind of miniature English grammar. When initiating the work in 1747, he had hoped to fix English pronunciation and spelling forever. But the years of lexical drudgery made him reject that goal as "the elixir that promises to prolong life to a thousand years." Of course, the spelling he chose completed the standardization of our orthography. He also rejected the idea of an academy, which, he prophetically hoped, would be hindered or destroyed by the spirit of English liberty.

Naturally he's much more famous for the dictionary proper. By far the most important contribution of his forty-thousand-word collection has been the 114,000 quotations used. His definitions primarily derive from analysis of these and many other actual examples of writing, especially from Spenser through the Restoration. There is occasional prejudice, like the definition of *oats* as the grain given to English horses but eaten by the people of Scotland. Such isolated instances don't militate against the *Dictionary*, one of the major linguistic events of Authoritarian English. As Johnson stated in the preface: "to collect the Words of our language was a task of greater difficulty: the deficiency of dictionaries was immediately apparent; and when they were exhausted, what was yet wanting must be sought by fortuitous and unguided excursions into books, and gleaned as industry should find, or chance should offer it, in the boundless chaos of a living speech. . . . I have much augmented the vocabulary." Thereby he rejected the procedure of previous lexicographers like Bailey, whose dictionary saw its sixteenth edition that same year (1755). They had mainly compiled their words from predecessors' lists.

Johnson's employment of illustrative quotations helped establish the "usage tradition," including levels of usage. The tradition overrides any "authority's" arguments from logic, philosophy, or classical origin. From his *Dictionary*, the line stretches to *A New English Dictionary on Historical Principles* (1884–1933), intended to record the usage of every word in extant English writing. More than six million individual word-slips for the half million entries collected during seven decades

underpin *The Oxford English Dictionary,* as it's now called. After subtracting all inflectional and derivative forms, we still have 240,165 main words in twelve huge tomes and a supplement. It has been the chief source in the preparation of this book. From it we skip to the usage-based *Webster's Third New International Dictionary* (1961), which cost more than $3,500,000 and required the labors of more than one hundred full-time specialists. It contains four hundred fifty thousand entries, of which one hundred thousand were not in *Webster's Second* (1934).

Noah Webster

Webster, the third grammarian, is a convenient bridge to Late Modern, in which we'll be mainly concerned with the American story. In terms of publication date, his *Dissertations on the English Language* (1789) belongs at the end of Authoritarian English, when the language was spreading around the world. In the book he followed the philosophical grammarian's traditions. There are ideas on language in general and on English in particular, some of which are highly modern. For example, he wrote that men learn their native language first and study about its structure later in grammars. Propriety in speech derives from "the general practice of the nation," or usage. The scholars' business is "to find out what the English language *is,* and not how it *might have been made.*" The lexicographer should locate the original quotation to justify a given definition, rather than appropriate it secondhand from some earlier dictionary.

However, there is a rambling down popular, speculative paths of the day. For instance, Webster believed that rude savages first make inarticulate sounds, or interjections. Repetition eventually converts the sounds to nouns, which form the beginning of man's language. How does Webster's explanation compare with those in our first chapter? In an appendix to *Dissertations,* he proposed spelling reform, which he later made use of in his great dictionary. Traditional views about the colonial grammar school had already been challenged in his "blue backed speller" (1783), in which he proposed to train American youth to believe in the independence of American schools, literature, and politics.

Late Modern, 1800–1900

Thus we reach Late Modern, or "the spread of English." It nicely complements Early Modern, when foreign words had flooded into the language. In the nineteenth century millions of English-speaking emigrants introduced their "foreign tongue" into Australia, North America, South Africa, the West Indies, and other distant lands. There were movements toward national dialects like American, as opposed to British. The chief alteration in the language was lexical, particularly the vocabulary of the American colonists and their descendants.

According to nationalistic Noah's *Dissertations*, British English was declining by then. "Within a century and a half, North America will be peopled with a hundred millions of men, *all speaking the same language*," which couldn't be dependent upon European English, "almost confined to an Island and to a few millions of people." His projection was right. The United States grew from an estimated 4,300,000 white and 1,000,000 black population in 1800, to a total 75,600,000 in 1900. He predicted a necessary, unavoidable "future separation of the American tongue from the English."

Webster's prediction was based considerably on the growing divergence in lexicon. There were "numerous local causes, such as a new country, new associations of people, new combinations of ideas in arts and sciences, and some intercourse with tribes wholly unknown in Europe." In his preface to the two volumes of *An American Dictionary of the English Language* (1828), on which his fame chiefly rests, he elaborated. British customs, he said, include hawking and hunting, nobility, and the like. Not indulging in these practices, Americans weren't using the words describing them. On the other hand, America's new government and overall situation required both new items and altered meanings for old items. He cited *land office, land warrant, plantation,* and *selectmen. Senator* required redefinition to fit the American meaning. For about a century he was right. As the American English lexicon expanded, there was a growing difference from British English.

Americanisms

Needless to say, the study of "Americanisms" could easily

consititute a separate book. *A Dictionary of American English* (1938–44; 4 vols.) and *A Dictionary of Americanisms* (1951; 2 vols.) are incomplete listings even up to their publication dates. We'll just sample the four principal sources. A first one is the employment of old words for new meanings, which Webster mentioned. A few more examples might be added to his: *corn* (Indian corn or maize), *fix* (bribe), *swan* (verb, meaning "declare"), and *turkey* (American bird). Whole categories might be added. Consider these shifts of meaning to designate American topography: *bluff, bottom, divide, notch, ridge.*

A second source is compounding. Again, the terrain is described by compounds like *backwoods, bottomland, canebrake, foothill, underbrush, watershed.* Note the simplicity of the words. Some are as picturesque as Old English items. A few language jingoists have stressed this creative quality as being characteristically American, whereas imaginative compounds exist in other tongues as well. Anyway, the colonists were mostly English. Here are items describing forms of life in the New World: *bluebird, bullfrog, devilfish, diamondback, doodlebug, ground hog, prairie dog.* The word *prairie,* borrowed from the French, was coupled with words like *chicken, rattler, rose, schooner, wolf.* Actually, a prairie wolf is a coyote. The colonists apparently thought the creature looked like a wolf, an animal well-known in England. Consider a dozen miscellaneous combinations, some of which are colloquial or slang today:

almighty dollar	flat-footed	salt lick
blue law	highbrow	strong-arm
bluenose	paleface	swimming hole
dugout	peace pipe	warpath

There may be idioms containing a verb:

be on the fence	kick oneself
give somebody fits	knock down and drag out
go in for	pull up stakes

A third source is borrowing. When the colonists took the Indians' lands, they took the Indian foods too, especially Algonquian. The names accompanied the theft: *hominy, pecan, persimmon, pone, squash, succotash.* Indian words were useful for the

strange forms of life: *chipmunk, moccasin, moose, (o)possum, (rac)coon, skunk, terrapin, woodchuck.* Native customs provided *papoose, powwow, squaw, tomahawk, totem, wampum, wigwam.*

Fellow colonists supplied other words. French borrowings from Indian tongues help illustrate the difference between French in the New World and that of Paris. Thus *bayou, caribou, coulee, lacrosse, lagniappe, levee* (embankment), *mackinaw, toboggan.* In general, such items didn't enter British English immediately.

Spanish loanwords, much more numerous, often went simultaneously, if not first, into British speech. After all, trade with the "Spanish Main" lands was quite intense from the Renaissance onward. Many items were borrowed only into American English originally, as in:

adobe	arroyo	mesa	ranch
alameda	fiesta	mustang	stevedore
alamo	frijoles	patio	tequila

There were also Germanic loanwords. The extensive contact with Dutch colonists, whose lands passed into American hands, led to many borrowings. Here are some:

boss (employer)	cruller	Santa Claus	spook
coleslaw	dope	scow	stoop (porch)
cookie	patroon	sleigh	waffle

A smaller source was German, as in these food-related words: *cranberry, delicatessen, hamburger, pretzel, smearcase, wienerwurst, zwieback.* Borrowing has continued from all these and other languages after 1900.

The study of borrowed place-names is another large topic. We'll simply generalize that the colonists used indigenous sources extensively, as the Anglo-Saxons had done from the Celts. Similarly, few Indian words that were not place-names have been taken. Consider *Alabama, Chicago, Dakota, Iowa, Massachusetts, Niagara, Ohio.* Dutch provided names of New York "islands" like *Coney, Long,* and *Staten.* French gave *Champlain, Huron, Lafayette, Louisiana,* and *St. Louis,* among many. Spanish supplied *El Paso, Florida, Los Angeles, Montana,* and *San Francisco,* of many. *American,* from "Amerigo Vespucci," was first recorded in 1578 to name the aborigines who inhabited

the continent. Most of the tens of thousands of place-names in the Hawaiian Islands came from Hawaiian, including *Hawaii, Honolulu* (protected bay), *Kona* (leeward), and *Mauna Loa* (long mountain).

A fourth source of Americanisms is coining, or neologism. It has occurred in English over the centuries and may be common to all languages, so that Americans are perhaps no more inclined to create brand-new words than any other people. Here is an American sample:

blurb	cahoot	caucus	chunky
debunk	gubernatorial	pep	skiddoo

This brief sketch of Americanisms may suggest the kinds of items that Webster included in the five thousand added to *A Compendious Dictionary of the English Language* (1806), the prelude to *An American Dictionary*. British reviewers were especially hostile, labeling many of the five thousand as "vulgarisms." They said that the general language shouldn't contain the words, which belonged only to the barbaric American dialect. In some ways they were right about the "American" qualities, once the value judgment is discarded. Most of the examples we've seen on the preceding pages haven't entered the general language. They're used in Miami and Los Angeles, not in London and Edinburgh.

American-British Spelling

Franklin, Webster, and others also proposed orthographic reform. Their motivation was linguistic, not some nationalistic desire to differentiate American spelling. Building upon Franklin's proposals, Webster wanted to complete the process of anglicizing French loanwords spelled with *-re*, as *chamber, disaster*, and others had already been. Uniformity required, he said, *meter* and *miter*. Some of his suggestions, never adopted, were later abandoned, despite their phonetic basis. For example, he urged the omission of silent letters. Which of these suggested respellings do we use today: *bilt, bred, brest, frend, giv, hed, ment, relm?* He unsuccessfully recommended a *k* for *ch*, resulting in *karacter* and *korus*.

Many other suggestions bore fruit. Introduced in *An Ameri-*

can Dictionary, they considerably dictate our modern practice. They explain whole sets of differences from British spelling. The *color - colour, rivaled - rivalled,* and *theater - theatre* sets have the largest number of members. Others have diminishing membership:

eon - aeon	defense - defence	willful - wilful
imbue - embue	program - programme	spelled - spelt
mold - mould	enroll - enrol	

To complete the spelling story, we can say that the English and American practices are now moving toward uniformity. Nineteenth-century linguistic nationalism has been discarded. The Oxford University Press, like other major publishers, has its own stylebook, *Authors' and Printers' Dictionary,* revised periodically. The Press naturally publishes many Americans' works. This book, with its underlying spelling philosophy, is increasingly influential in European English. In recent editions there has been a kind of leveling of differences. One notable example is the "y" type. The Press now uses *cider, cipher,* and *tire,* with *cyder* and so on cited as the "usual" British spelling. However, *pyjamas* and *syrup* remain. Another example is the "ise" type. The Press employs *-ize,* except in words like *advertise, apprise, chastise,* and others that are so spelled in *The Oxford English Dictionary.* Are these the American forms? There is a democratic choice of *ax - axe* and *story - storey.*

On the other hand, numerous single differences remain. For instance, *check - cheque, connection - connexion, curbstone - kerbstone, forever - for ever, gray - grey, gypsy - gipsy, peddler - pedlar, plow - plough, premise - premiss* (n), *toward - towards.* The difference in hyphenation stubbornly persists. The British frequently hyphenate prefixes, as in *a-kimbo, ante-room, bi-monthly, hem-stitch, hill-side, horse-flesh, mis-spell, north-west, re-echo.*

Influential American dictionaries are contributing to the informal move toward uniformity. *Webster's Third,* for instance, usually gives the British spelling as an alternate, as in "acknowledgment *also* acknowledgement." Sometimes there is a cross-reference, as when *Gipsy* is listed as a variation of *Gypsy.* The entry proper is under *Gypsy. A Manual of Style,* of the University of Chicago Press, recommends *Webster's Third* as the guide

for spelling. What should an American do? If writing for a British publisher, he normally must employ the English practice. If writing for an American reader, he probably should use ours, particularly since, strangely enough, some people may feel he's "affected." Neither practice is right or wrong; so we must be tolerant, as in all language matters. Meanwhile, many Britishers use some American spellings, while some Americans prefer British ones.

We've now sampled the growing lexical and spelling differences between British and American dialects until 1900, especially the Americanisms. Englishmen either didn't know some of the American words or else resisted their introduction into the British lexicon, while Americans weren't using some high-frequency British ones. Until about 1900 the two broad dialects seemed to be pursuing the usual path. Just as English and Frisian had eventually become mutually unintelligible, so it seemed that the "King's English" and "American" might face the same fate in the distant future. After all, the Anglo-Saxons and Frisians had been separated only by a narrow sea, with much intercommunication. Now an ocean separated two diverging "dialects." There was vast ill will because of the War for Independence, the War of 1812, and even the American Civil War. Economic competition intensified. Nationalists on both sides emphasized the political and linguistic differences.

In the next chapter we'll see the end of this century-long widening. Some of the differences are now beginning to disappear in Present-Day English, although with a large residue that never really interferes with communication. Radio, television (with borrowed kinescopes and satellite televising), and greatly expanded commerce, tourism, and international politics are playing a fundamental role, something that diverging dialects hadn't experienced before. If they had, perhaps there wouldn't be as many different languages as there are today.

Activities

1. There follows a list of words included under *cha-* in Phillips's dictionary. Eight are Latin or French borrowings; one is "native." Find the etymology, together with the original Latin or French spelling of the loanwords, in order to compare it with Phillips's English spelling.

Which one of the nine words became obsolete? What's the modern spelling of the survivors?

chace—a warren, a tennis term
chagrin—care, heaviness
chalybeate—of the temper or quality of steel
chamberdekins—Irish beggars
chamois—a wild goat
champerty—term in common law
chaplet—wreath or head garland
chart—paper or parchment
chauncel—sacred part of a temple

2. "Mispronunciation" is a key way by which an intolerant person mistakenly judges a person's social class. What are the various pronunciations of *garage, potage, rouge,* and *siege?* Which are nonstandard?
3. Find at least two words containing each of the following Greek borrowings. How does the original meaning contribute to that of its compounded form?

archae-	bio-	idio-	morpho-	-phobe
auto-	chron-	logo-	-phil-	-phone-

4. Find a Scandinavian borrowing with *fj-*, as in *fjord*. Why is the spelling often *fiord?*
5. Loanwords from several languages not mentioned in the text have entered English. The following list contains one word from each of the following languages: Chinese, Hungarian, Japanese, Malay, Malayalam, Portuguese, Russian, Sanskrit, Singhalese, Tibetan, Tongan, and Turkish. Identify the language from which each word below was borrowed. Are any of the sources surprising?

anaconda	goulash	lama	tong
atoll	kimono	mammoth	yam
catsup	kiosk	taboo	yoga

6. In his "dictionary," Coote marks with an asterisk those words spelled "differently" because of their French influence. His first six are given below. What's the modern

spelling of the five survivors? Give at least one other meaning for each.

accomplish – finish	adiourn – deferre
accompt – recken	alledge – bring proofe
ascertaine – make sure	ambassadour – messenger

7. Find two *-ing* prepositions like *touching*. Compose a sentence with each word. What is their status as prepositions today?
8. Look in the King James Bible for three occurrences of each of the following words: *hath, doth,* and two other *-eth* verbs. Find two modern adjectives like *aged,* where apocope never operated after the vowel was neutralized to schwa.
9. Proving the Great Vowel Shift for a given word is often a breeze because of the spelling. Consider the following list. The vowel in one word didn't change because it was never long; in a second, because the word was borrowed too late. What are the two? Try to plot the vocalic shift "forward and upward" of each of the other eight: *bed, boot, brake, dame, food, foot, great, mouse, speak, write.*
10. Dr. Johnson originated the practice of including some of the quotations from which he derived his definitions. For some words he discriminated standards of usage. For example, the verb *womanise* is judged "Proper, but not used." Look up one of the many reviewers' attacks upon *Webster's Third* and evaluate the review from a linguistic point of view. Here are references to three:

    "Dig Those Words," *New York Times,* Sept. 10, 1961
    "It Ain't Necessarily Uncouth," *Chicago Daily News,* Sept. 9, 1961
    "A Non-Word Deluge," *Life,* Oct. 27, 1961

11. Find at least five place-names in the United States from each of the following sources: Indian, French, Spanish, and Dutch. What is the origin of the names of your state, of its two largest cities, of its longest river, and of its tallest hill or mountain?

12. Find at least three examples of the *colour* and *theatre* types of spelling. Check the classified section of a telephone directory. Is *theater* or *theatre* used? Find one other British example of each of the following types of British spellings: *aeon, embue, mould, defence, programme, enrol, travelled, wilful, spelt.*

Chapter Eight

Present-Day English (after 1900)

For the sake of convenience, 1900 is the arbitrary date chosen to mark the beginning of Present-Day English. By the time of Queen Victoria's death in 1901, America had decisively defeated Spain and emerged as a world power. Guam, the Philippines, Cuba, and Puerto Rico were new American possessions, into which English spread as a competitor to Spanish and other tongues. The British began the eventual crushing of the Boers in South Africa. That British victory expended the last of late nineteenth-century imperialism. The Union of South Africa was established in 1910. Overall, English replaced French as the diplomatic language, and it became the medium of international commerce. In view of the two world wars and the international flavor of other events like the formation of the United Nations, we shouldn't be surprised at the international influence on the twentieth-century history of modern languages. The further spreading of English around the world is a familiar theme to us, because Americans or other native speakers of English have settled permanently in countries like France, Italy, Israel, Thailand, and Japan.

Linguistic Nationalism

Two other familiar themes persisted during the early decades of the twentieth century: linguistic nationalism and widening lexical differences between American and British speech. Journalists like H. L. Mencken jingoistically extended the spirit seen in Webster. In fact, Mencken's title, *The American Language* (1919), is borrowed from Webster's title, *An American Dictionary of the English Language*. However, Mencken went con-

siderably further than Webster, predicting that the divergence of American would soon make it and English mutually unintelligible. The four major sources of Americanisms were continuing to pour words into the lexicon, although borrowing was not quite as extensive as in the Renaissance. By his fourth edition (1936), Mencken was even more bombastic: world events had denigrated England, he emphasized, while America was rising. Future scholars, he said, may find themselves studying English as a dialect of American. There is a slight basis for his prediction that American would overwhelm British English. We need only note the general British outrage at the Americanisms flooding into London English, despite some equally jingoistic efforts by the British to reject the Americanisms as degraded and barbarous.

British-American Lexical Differences

Mencken may have overemphasized the American tide, for the British dialect continued its own sturdy development after 1900. Joseph Wright's *English Dialect Dictionary* (1898–1905; 6 vols.) and *English Dialect Grammar* (1905) list British dialectal expressions seldom found in the United States. Wright collected some five hundred thousand word-slips in the process. His conclusion, nonetheless, was that pure dialect speech is rapidly disappearing from even the rural areas because of the spread of education and modern communications media.

Americans didn't begin their "Dictionary of American Regional English" until 1965. It is designed to collect the greater part of the expressions, pronunciation, and meanings of native American English speakers in a thousand local and regional speech-communities in fifty states up to the date of publication. The computer-based project is expected to store up to five million word-slips, many of which will naturally be repetitions of the same word. Unquestionably this important work will discover many American elements not occurring in British speech and writing.

The major scholarly listing of the diverging American lexicon began in the late 1930s, when the first books of the vast Linguistic Atlas of the United States and Canada began to appear. Other evidence can be found in works like *A Dictionary of American English* (1938–44) and *A Dictionary of Americanisms* (1951), which contain long lists of words often unknown to Britishers, or

at least their particular meaning unknown. By World War II, the British and American lexicons each contained many items unknown to the other, or, worse, known by a different meaning that might cause misunderstanding. At that time, it seemed that the two might be pursuing the usual path toward eventual mutual unintelligibility.

So much concern was voiced about the matter during World War II, that the United States Army published and distributed *A Short Guide to Great Britain* (1942). The pamphlet contains a list of 183 commonplace American words then unknown in England, along with their British equivalents, equally unknown in the United States. Other lists followed. England did the same for Royal Air Force cadets in *Notes for Your Guidance* (1942). It's still easy for us to compile such a list three decades later. Here's a sample:

apartment - flat	holdup man - raider
baby carriage - pram	ice cream - ice
beer - lager	line - queue
biscuit - scone	movies - flicks
carnival - fun fair	newstand - kiosk
cracker - biscuit	oatmeal - porridge
dry goods - drapery	overcoat - greatcoat
fall - autumn	potato chip - crisp
flashlight - torch	racetrack - race course
French fries - chips	subway - underground
gas - petrol	truck - lorry[13]

A contextual look at British-American differences might be worthwhile at this point. Consider the following beginning of a British newspaper story:

Text G:
A Sports Story
From the Harrow Observer and Gazette

(1) A goal in extra time gave Old Lyonians a well-deserved 3–2 A.F.A. Senior Cup win over Wood Green Old Grammarians at Enfield on Saturday.

[13] See the detailed list in Raven I. McDavid's edition of Mencken's *The American Language* (New York: Alfred A. Knopf, 1963), pp. 277–85.

(2) Opening exchanges were even with both sides taking a long time to settle down and when eventually J. Hines broke through for Lyonians it was only to be robbed as he was poised to shoot in front of goal.

(3) Clever inter-passing between the forwards sent P. Hines away on the right but his centre landed behind the goal and Wood Green began to take control with open football that made good use of both wings.

(4) These tactics paid when a free-kick was pushed out to the right and the winger beat his man and closed in towards the goal-line. Swain made a vain attempt to intercept and the ball was slipped to the centre forward, who had only to tap it over the line.

(5) Watson went close to replying for Lyonians from a good pass by Wilson and, with Grammarians' attacks foundering before strong play by Freund and Williams it was Lyonians' turn to get the upper hand.

(6) After Tremlett and Wilson had both been near the target, Lyonians began to play the ball too close for a time but a return to more open methods brought its reward before the interval, Watson levelling the scores during a scrimmage.

(7) Lyonians went into the lead early in the second half, a good movement on the right ending with a centre by P. Hines finding Wilson just in the right spot to score. . . .

Lexical Differences

Admittedly there's a special vocabulary associated with sports, not to mention that rugby is different from American football. On the other hand, note the seeming familiarity of every word in the passage. The general story is easy enough for us to follow; yet the particular meaning isn't always clear. The reason is that some of the words may have a different meaning for us. Let's skim through the passage. We don't hesitate over *extra time* (1), for the context gives us *overtime*. "Of goal" (2) is disconcerting, since we expect a *the* in the sequence. *Inter-passing* (3), with the British hyphenating of some prefixes, is simply *passing*. In "tactics paid" (4) we look vainly for the *off* to complete the idiom. *Towards*, as opposed to *toward*, has been mentioned before. "Watson went close" (5) would seem less quaint with *came* instead of *went*. "Strong play" might be clearer with

defensive inserted. *Target* (6) is contextually revealed to be the goal. *Interval* can be contextually interpreted as "halftime." Instead of "levelling the scores," what do Americans do? We may wish to delete *a* in "during a scrimmage," until we realize that an actual play is intended, not the practice play that we often describe with the same word.

Let's look at some other differences in meaning. In a London station we'll get a one-way ticket when we ask for a "single," intending "one person's ticket." If we're out of town and about to return to London, a request for a "return" will elicit a round-trip ticket. The sign "saloon" on a bus doesn't mean that drinks are to be imbibed there, but that this deck of the bus is for smoking, and the other is not. Drinks are sold at a "public house," clipped to form *pub*. On a train, they're sold in a "parlour car." What's the meaning of "no naked light is to come near petrol tanks"? "Open flame" rather than "electric bulb" is meant, because the gasoline (petrol) tanks can be observed from our hypothetical train window. Luckily, we didn't try to "book passage" in the "parlour car."

Prepositional Differences

Such lexical differences are superficial. After all, deep-structure rules are the same for all dialects of a language. Let's consider the differences in terms of our generative-transformational model. The choice of items to be inserted in the deep structure then becomes the major cause of our momentary difficulty with the sports story. The most troublesome form class is the preposition. In certain phrases the preposition may require an article before the noun; in others it may bar one. The reason is still a little mysterious. "In front of goal" (2) illustrates an omitted article, whereas the same "in front of" requires *the* before *goal* for us.

A broad explanation may be that prepositions were narrowing their range of meaning at the time when the English began colonizing the New World. Remember the substantial overlapping of Shakespeare's prepositions? After separation of the two broad dialects, perhaps prepositions simply pursued a slightly different path. Consider this current British structure: "He went to hospital." They don't use *the* before *hospital;* we must. The

same is true for their "at college or table, by government, in future, out of window."

By contrast, we have several idiom-like phrases, where insertion of the article either changes the meaning or makes the structure nonsensical. For instance, "at church, home, school," "in bed, service, sight," "on foot, leave," and "by air, cable, phone." Contrast "at the church" with * "by the air." We say "on file"; the British say "on the file." If an article is to be included, the American preposition must be *in* instead of *on*—thus "in the file," never * "on the file."

The British say "I am in health." For us, the difference isn't the omission of an article, but of an adjective like *good.* * "I am in a health" and * "I am in the health" aren't acceptable, since *health* is [- count] for us. Conversely, "He made good use of his time" is the British structure, which is acceptable to us, as well as "He made use of his time." Sometimes the preposition may vary. Americans say "different from" or even "different than" when the purist isn't looking. The British use "different to." They go "in their way" to London, while we go "on our way." Their *apothecary* is *in* Oxford Street; our *drugstore* or *pharmacy* is *on* that street. They omit *to* before a pronominal object, as in "Give it me." They use *to* when the direct object is a proper noun, as in "Ian wrote to Mary." That structure is also acceptable to us, although "Ian wrote Mary" may be a little more frequent. The deep structure of the latter is ambiguous. It derives either from **NP + Aux + V + Prep-Phrase,** with **T-del** deleting *to,* or from **NP + Aux + V + NP.**

Transformational Differences

Transformational differences between British and American speech primarily involve **T-pro, T-infl,** and **T-agr.** Consider "They bring in him," which is the final British surface structure. We use a revised form of **T-pro** to invert the personal pronoun, for the obligatory result "They bring him in." Contrarily, the British employ the same **T-pro** to change the deep structure "Sarah catches up with him" to "Sarah catches him up." For us, there's a constraint against use of the rule to produce this inversion. Such variations in the positioning of pronouns are somewhat complicated.

Differences in **T-infl** occur with a few dental verbs, as seen

in the *-ed* or *-t* spellings. For the British, this small group takes /t/ in the preterit and past participle, as opposed to the American /-d/. The verbs are the following:

burnt	dreamt	knelt	learnt
smelt	spelt	spilt	spoilt

The cleavage isn't complete. We probably say /nɛlt/ more than /nild/ for what is often spelled as *kneeled*. However, our past participle of *lean* is seldom pronounced /lɛnt/.

T-agr gives verbs the matching feature [- sing] of certain group-nouns, which are [- count, - sing] for the British. For instance, "Government find, Northampton Team are, the village decide." These aren't yet collective nouns for us, except for dejected coaches, who say "My team aren't very good," meaning that the members play as inept individuals rather than as a team.

Phonological Differences

As for phonology, we've probably observed that British intonation often begins on a higher pitch on the first syllable, drops, and then rises at the end. In general, they've lost secondary stress on the penultimate syllable of a polysyllabic word. Examples are *altérnately, coróllary, labóratory, miscéllany*. Which syllable do we stress? The British usually syncopate words ending in *-ary, -ery,* and *-ory,* producing a final syllable of /tri/. Thus /sékrətri, bǽtri, ləbɔ́rətri/.

The last syllable is often neutralized, as in /mɪlkmən/, as opposed to our /mɪlkmæn/. The London borough *Southwark* is /səðərk/, contrasted with the frequent American /saʊθwərk/, although both dialects stress the first syllable. Note the several phonemic differences between the two pronunciations. The British /ə/ is generally familiar to us in *-shire*. We're more likely to use /aɪ/, as in *Oxfordshire*. Conversely, they normally employ /aɪ/ in words with *-ile*. This contrasts with our schwa in the final syllable of words like *fragile, futile,* and *sterile*. They may be moving toward /ə/, as in their /ǽjəl/.

On the whole, the chief pronunciation difference involves a handful of rather high-frequency words. Here are five: their /š/ in *schedule,* the /i/ in *been,* the /ɛ/ in *ate,* the /æ/ as the first vowel in *garage,* and /vítəmən/. Initially, we might point an ac-

cusing GVS finger at *vitamin,* guessing that it didn't shift for the British but did for us. However, the moment we recall the date of Casimir Funk's great discovery, in 1912, we know that his compounding of the Latin *vita* (life) and *amine* occurred long after the shift. The spelling simply misleads us. Actually, Funk's original spelling was *vitamine.* The *-e* was discarded after the discovery that most vitamin compounds aren't amines.

Such variant pronunciations can cause momentary difficulty in communication. Suppose a Londoner says /ɔɪv gɔt mi·ǽt/. Without a context, we may puzzle a little before interpreting the structure as "I've got my hat." There still may be misunderstanding. *My* is pronounced /mi/ by many Englishmen, so that we must be alert, or we'll mistakenly interpret the structure as *"me hat." Or suppose an Australian says he's going /tə·dáɪ/. Again, a homonym confuses us. He's only going *today,* not "to die." Still, all such differences are trivial when viewed in relation to the overall language.

Dialects

The differences between British and American English raise the question of regional and social dialects. As was said at the beginning, it's a gross oversimplification to speak of a British or an American dialect anyway, except in the broadest regional terms. Because the Middle English dialects have continued historical development in Britain, differences in speech are ordinarily more acute within England, Wales, and Scotland than among the various American regions today.

Like the earlier American Atlas, a *British Survey of English Dialects* is being published to show the regional varieties. The University of Leeds collected the data in 1951–60. Informants were asked 1,322 questions, of which 387 are phonological, 205 syntactic, and 730 mainly lexical. The questions relate to the farm; farming; animals; nature; the house and housekeeping; the human body; numbers, time, and weather; social activities; and states, actions, relations. For example, the informant is asked about the "farmstead":

1. _____ is the place where you keep the animals that grunt.
2. _____ is the place where you keep the birds that lay eggs for you.
3. _____ is the place where you keep the birds that coo.

The informant's answer is often *pigsty, hen-house, dove-cote,* respectively. Would an American's answers differ?

American Regional Dialects

Now let's consider the American dialects. They've developed primarily from the original settlers' speech, according to region. The differences have been considerably narrowed by the American propensity for moving, together with the spread of education, books, and various communicational media. In fact, these are probably worldwide forces now helping to make more uniform the otherwise diverging dialects of particular languages.

Obviously, American dialects can't coincide with the descendants of the Middle English ones. Nonetheless, Southern and London speakers primarily settled New England and the central Atlantic coast. Northern and West Midland speakers moved into western New England. Many Welshmen and indentured people populated the areas south of the Ohio River. A class-conscious Southern aristocracy, composed primarily of wealthy tobacco planters, maintained close connections with England. There were also "poor whites" in the back areas, together with the Negro slave majority. The speech patterns of all three social groups somewhat blurred together over three centuries. In the Eastern states, three broad belts of dialects developed—Northern, Midland (North and South), and Southern, as shown in Figure VI.

In areas like southeastern Pennsylvania all the way to the Susquehanna River, large German migrations in the eighteenth century considerably affected the speech. Many Ulster Scots migrated westward. Their routes often followed the topography—that is, rivers, mountain gaps, and well-watered valleys. The migrators came mainly from the western areas of the three dialectal regions named above. The Blue Ridge, the first of the mountain barriers, was crossed in the eighteenth century, primarily by the Ulster Scots. As the population of the inland areas of the Carolinas and Georgia expanded, the speech there was soon influenced by cultural centers like Charleston and Savannah.

In the early nineteenth century, Pennsylvanians moved to the Great Lakes and the Mississippi Valley. More European immigrants arrived, particularly in the Middle West, after the

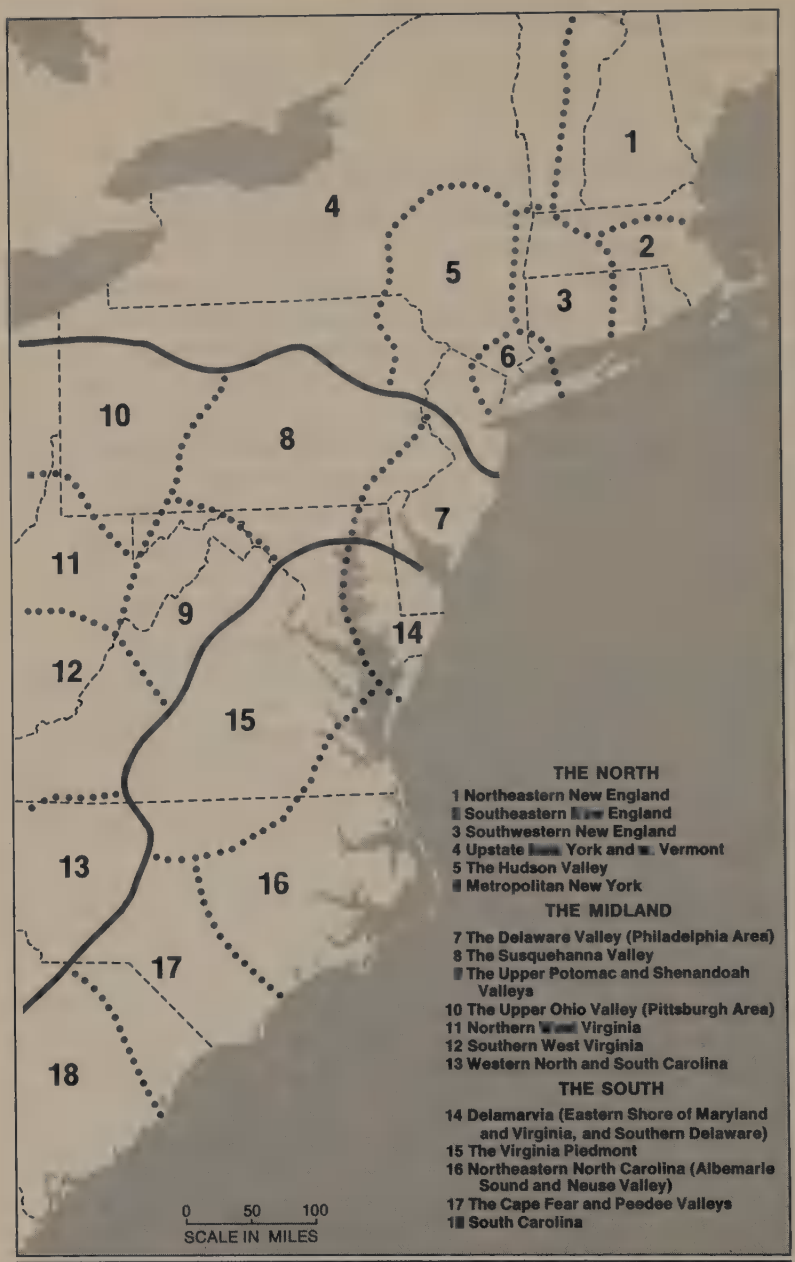


Figure VI:
The Speech Areas of the Eastern United States

Indians were killed or ejected. In general, Northerners followed the Mohawk Valley to the Great Lakes, since the Erie Canal was not opened until 1830. Midlanders and Southerners migrated west on rivers like the Miami, Wabash, and Illinois. There was a dotted line of settlements all the way to the Pacific by 1850, when California gained statehood. The Pacific Northwest was settled mainly by speakers of Northern and Midland. We must remember that these dialects are named for American regions; the British dialects, for British regions.

Let's glance quickly at the three American dialects of the Eastern states. Each has internal social variations, partly because each regional dialect is composed of poorly educated speakers, younger and somewhat educated speakers, and cultivated college graduates.[14]

Still, there are clear lexical distinctions:

Northern	Midland	Southern
darning needle	snake feeder	snake doctor, mosquito hawk
pail	bucket	bucket
spider	skillet	skillet, frying pan
swill	slop	slop
whippletree	singletree	singletree
you (plural)	you, you-all	you-all

The number of examples where there's a different word in each of the three regions is limited, and there are exceptions even in the above list. For example, *frying pan* is used in urban New York areas. Even so, we can visualize possibilities of horrified misunderstanding. What if the newly migrated Bostonian asks his Atlanta neighbor for the loan of a spider? Or what if, for some curious reason, the resettled Virginian asks his Buffalo neighbor whether snake doctors are prevalent in the area? The native of Buffalo may think that his new neighbor owns a pet copperhead.

[14] The following sketch of the three dialects is taken principally from McDavid's excellent chapter "The Dialects of American English," in Francis's *The Structure of American English*. Carroll E. Reed's summary, *Dialects of American English* (Cleveland: World, 1967), is the chief source about other American speech. The detailed data can be found in various continuing publications of the Linguistic Atlas of the United States and Canada, such as Hans Kurath and Raven McDavid's *The Pronunciation of English in the Atlantic States* (Ann Arbor: University of Michigan Press, 1961).

An ex-Providence citizen might evoke the same bewilderment in Savannah by commenting on the number of darning needles in the air. The cause is a familiar one—the same word used with a different meaning.

There are numerous other lexical distinctions among regional dialects. For example, New Yorkers stand "on line" and live *in* Manhattan, never *on* Manhattan, which is an island. In the Northern dialect the people eat Dutch cheese or pot cheese and johnnycake. In Pennsylvania, these foods are called smearcase and pone or corn pone. In Northern, one is sick to the stomach. In South Midland and Southern, the people eat "light bread." In Midland, there are blinds (shades), green beans (string beans), and pokes (paper bags). In Southern the cows low; the people tote things and eat snap beans.

Contextually, none of these creates any real problem. Without adequate context, misunderstanding can arise. Suppose a Baton Rouge boy tells his Toledo girl friend that he'll pick her up in the evening, right after dinner. When he arrives about 1:00 P.M., he may be shocked to find that she may not be home. She eats dinner about 6:30, with "afternoon" continuing for a while longer before "evening" begins. He eats "dinner" at noon, when "evening" begins and runs until sunset. Such variations aren't consistent within even a single dialectal region. In parts of southern Pennsylvania, for instance, the A & P clerk puts one's groceries in a toot.

Naturally the three eastern dialectal areas don't extend neatly across the country. Roughly, Michigan and the Upper Midwest and on to the Pacific Northwest, with some skips in between, share many Northern features. Exceptions in Washington, Oregon, and Idaho are frequent. Most people in these areas apparently use the non-Northern "back porch, cornbread, corn on the cob, dragonfly." Midland words may be expanding west of the Mississippi, often at the expense of Northern ones. California speech is generally somewhere between Northern and Midland, perhaps more like the speech in Illinois than like that of the "mother" New York City or Pennsylvania. Purely Southern dialectal terms are quite rare on the West Coast.

Transformational differences among American regional dialects are quite restricted. They're primarily confined to **T-agr** and **T-infl.** In Northern and Midland, **T-agr** gives [+ acc] to the

pronoun in structures like "It wasn't me." **T-infl** and a **P** rule provide *dove* as the preterit of *dive* in Northern. For some speakers of South Midland, the possessives may be *ourn* and *yourn* in "That book is _____." The previously discussed *"might could," where *might* is intended as an adverb, is used in the "Deep South" by many people.

Phonological differences also exist. *Hoarse-horse* and *mourning-morning* are homonyms from the Potomac northward and westward through Pennsylvania. In Northern, *with* has /ð/, and *grease-greasy* have /s/. Elsewhere, we find /wɪθ/ and /z/. In eastern New England, New York City, the Hudson Valley, and Southern, /r/ is lost before consonants, as in *garbage, farther,* and *horse.* In New England and down into New York, the linking *r* is found in structures like / hi sɔ́rər / (he saw her) and / əmérəkərits / (America eats). In Midland and Southern, /y/ occurs in *due, new, Tuesday;* and the *hoarse-horse* and *mourning-morning* pairs are homonyms. *Bleat* and *Mary* have /e/ in Southern.

American Social Dialects

Metropolitan influences and the social variations within a single region complicate the three broad belts of regional dialects. Cities like Boston, New York, Charleston, and San Francisco have extensive cultural effects on adjacent areas. Within metropolitan areas, class distinctions are often rather rigid. Part of the cleavage is a person's speech, which seems to be considerably determined by education, whether in a city or in rural parts of Nebraska. As the child of a disadvantaged family proceeds up the educational ladder toward college, he becomes more aware of his dialect. Heretofore, he may not have noticed the possible /d/ in his pronunciation of *then,* or a [+nom] for the second pronoun in *"They come at Jack with they blades." The social, perhaps consequent rejection was silently effected by the listener. When the child observes the standard form in "advantaged" colleagues' speech or is simply criticized by one of them, he attempts to drop the socially unacceptable item. Today one of the major problems in the United States is the proper attitude toward what we'll term nonstandard social dialects.

Negro Speech

Some people erroneously generalize Negro speech as a nonstandard social dialect. In fact, Negro speech isn't a separate dialect of American English at all, and certainly many blacks speak standard English. It's as impossible to characterize black speech as it is to characterize white, although the Southern dialect is probably used by most blacks in the United States. The many exceptions are those blacks brought up in northern metropolitan areas like Chicago, Detroit, and New York. Over the telephone, a black's Southern dialect can seldom be distinguished from a white's Southern dialect, assuming equal education and culture. Unfortunately, many blacks have been deprived of educational and cultural opportunities. When their parents speak nonstandard English and they themselves have been deprived of the chance to associate with speakers of standard English, their structures may frequently omit the copula. The omission is, of course, a matter of usage, not a sign of inferiority. Still, the omission is nonstandard and may be a feature of the speech of both deprived whites and blacks.

We might consider some sample dialogue from Warren Miller's explosive novel about Harlem, *The Cool World.* It presumably reflects the Negro novelist's understanding of characteristic elements of his uneducated Negro characters' speech. Note the omitted *be* forms:

For a while any way it clear	(before adjective)
She a big woman	(before nominal)
I in a big hurry	(before prepositional phrase)
But now I here	(before adverb)
It the truth	(after filler "it")[15]

We see immediately that all five examples are in present tense. Data collected from some disadvantaged Negro children in Florida confirm the omission of the copula for present. The children use it, however, for past tense. The deep structure of the five examples is clear enough. **T-del** or some other rule has obliga-

[15] From the structures selected by Beryl Bailey, in "Toward a New Perspective in Negro English Dialectology," a lecture presented to the Linguistic Circle of New York in March, 1965.

torily deleted *be* when [+ pres]. The surface structure of the five is standard English otherwise.

The problem isn't quite so simple. **T-del** is seldom used even optionally in these sequences in standard speech. Why is it required here? The answer may be related to two others. One is the substitution of *ain't* for negative forms of *have* or *be*. For example, a Miller character says, "I been thinking about that. An they ain't no plan." The other answer involves the use of **T-ques** and **T-do.** Consider this deep structure, "he **present** say that now." We then apply three T rules in sequence:

T-ques: present he say that now
T-adv: when **present** he say that
T-do (since **T-aux** won't work): when do **present** he say that

The result is "When does he say that?" Note the widening of **T-adv** in Present-Day English. If the adverb is one of place, it's changed to *where* and permuted to the front of the sentence. If it's an adverb of time, like the sequence above, *when* is permuted.

There is no *when* or *where* in our Miller examples. Still, other data suggest that *"When he say that?" may be the nonstandard sentence of deprived blacks, not "When does he say that?" **T-ques** and **T-do** simply aren't employed in the creation of such nonstandard structures, and [+ pres] of *say* in the above sentence is **ø.** The data aren't wholly consistent. Some blacks from Detroit and Washington, D.C., use sequences like "They be fighting and going on." *Be* isn't deleted in this present-perfect structure, and [+pres] of *be* takes **ø.** Past-tense modals turn up in sentences like "He might be home now" and "Should they be bad kids?"

Nonstandard Speech

Much more research needs to be done on *be, do,* and *have* structures. We do know that sequences like "My brother he at school" and "She the first one started us off" bring down social disfavor upon the speaker. They aren't incorrect. They just aren't accepted socially, for whatever reason. The speaker of a nonstandard dialect is sometimes unaware that his speech is looked down upon by intolerant persons. The poor jokes about the uneducated New Yorker's *"Toity-tree boids" have sometimes alerted him to some fellow New Yorker's reaction when he

uses /d-/ in *they* and /t-/ in *thing*. Metropolitan social stratification is partly based upon such phonological choices. The Linguistic Atlas lists many other features that aren't fashionable at the moment. Some would have been appropriate centuries ago but now, arbitrarily, aren't. If at a White Sox game a personnel officer sits next to someone who says *"He like to done it" or *"He done finished it," the verb sequence communicates the intended tense and meaning. If the officer is later approached for a job by the same person, a reminder about the player who *"done finished it" may cause automatic rejection.

Why do sequences like *"done finished" become socially unacceptable? Linguistic questions asking "why," as was said at the beginning, are usually unanswerable at present. For instance, the elegant Prioress listens without interruption to the Miller's "four-letter words" on Chaucer's fictional pilgrimage of about 1400. She is in no way offended, whereas those same words have so degenerated that only special dictionaries list them today. What we might do is note the value judgment ultimately underpinning any society's decision that a banker's dialect may be the prestige one. A laborer is somehow expected to discard his own speech in favor of the banker's, at least in more public situations in which he may be communicating with a dentist or an attorney. The leaders of society needn't be careful of their speech, which sets the temporary standard.

How does one know when he is using a nonstandard expression? If we've been unwittingly employing it, we won't reject it unless we observe its absence in the prestige speaker's language, or unless he tells us to avoid it. Shaw's *Pygmalion,* transformed into the musical comedy *My Fair Lady,* has illustrated the problem to audiences around the world. The aristocratic phonetics professor, Henry Higgins, sets out to make a lady of Eliza Doolittle. Her social dialect is cockney, which he wishes to change to his London Standard in order to help elevate her social status. He fails because it is quite difficult to change one's speech patterns.

The difficulty is compounded by the fact that what is acceptable and what is unacceptable vary significantly with the situation and the audience. For example, questions in informal speech are comparable to the nonstandard ones considered earlier. If a friend sees us in the hall and asks "Hungry?" are we offended? He has even deleted the *you*. "That your sister?" is acceptable

too, whereas a measured "Is she your sister?" might even be a little formal. As for recognizing nonstandard features, we're fortunate in having a detailed list. Professor Raven McDavid has discovered certain pronunciations and inflections to be social "markers." Each can be handled by pattern drills that make the speaker aware of the socially acceptable equivalent:

(1) omission of a weak syllable preceding the stressed one, as in *fessor* and *porter* for *professor* and *reporter*
(2) frequent front-shifting of the primary stress, as in *fínances, ínsurance, pólice*
(3) heavy stress on a weak final syllable, in *accidént, elemént, evidénce*
(4) confusion between /θ/ as in *thin* and /t/ in *tin,* /f/ in *fin,* /s/ in *sin*
(5) confusion between /ð/ in *they* and /d/, /v/, /z/
(6) vocalic homonyms in *bird-Boyd* and *curl-coil*
(7) "zero" for the noun plural, in "Two boy came"
(8) "zero" for the noun genitive, in "Mr. Brown hat"
(9) analogizing of the pattern of /-n/ of *mine,* to form *ourn, yourn, hisn, hern*
(10) analogizing of the compound reflexives, in *hisself, theirselves*
(11) use of *them* for [+ nom], in "them books"
(12) compound demonstratives, in "these-here dogs, them-there cats"
(13) analogizing of inflected comparisons, in "wonderfullest time, lovinger child"
(14) double comparisons, in "more prettier, most ugliest"
(15) unorthodox person-number agreement of the present of *be,* used with all persons
(16) unorthodox agreement of the past, in "I were, we was"
(17) unorthodox agreement of the present of verbs, in "I does, he do"
(18) omission of /-ɪŋ/ in the present participle, in "He was open a can"
(19) omission of /-t, -d, -əd/ of the past, in "It burn a hole yesterday"
(20) omission of /-t, -d, -əd/ in the past participle
(21) omission of *be* before a predicate noun, in "Chester my best friend"

(22) omission of *be* before a predicate adjective, in "They ready"
(23) omission of *be* before a present participle, in "I going"
(24) omission of *be* before a past participle, in "The train window broken"
(25) omission of *have* before *been,* in "They been drinking"
(26) substitution of *been, done,* or *done been* for auxiliary *have,* in "He done been finished"[16]

Several observations can be made about this checklist. All twenty-six listings are acceptable in humor or mimicking. Some are historical (9–10, 13–14), having been replaced as the language changed. Structures similar to (19) frequently occur in speech. In /ɪt bərn daʊn/, juxtaposition of the /d/ in *down* requires omission of the /-d/ in *burned,* as is true for some other consonants. Twelve of the twenty-six involve *be* or auxiliary structures. Several are acceptable in the right context. For instance, in informal speech to a friend, one could begin the conversation with "I been thinking." We simply remember to insert the auxiliary *have* in a more formal situation or when we're writing the sentence. Likewise, the written past of *burn* has *-ed* in (19).

Bilingual Speech

The already complex problem of nonstandard social dialects is complicated for anyone whose first language is not English. Unknowingly, he substitutes elements of his native language. Thus the phonemic substitutions of a Spanish speaker can be predicted from a comparison of the two sets of sounds. As Spanish doesn't have /ɪ/, a native Spanish speaker will use the other high-front vowel. *Sit* becomes /sit/. An ignorant if not intolerant interpretation might be *"I seat down." Lacking a native /æ/, he will probably use a nearby vowel. *Bag* becomes /bɑg/ or /bɛg/, giving the possible interpretation of *"Where's the bog?" Lacking /ə/, he may use /ɑ/ or /ɔ/ for the vowel in *but.* His /u/ substitutes for our /ʊ/. There are comparable

[16] This list is almost verbatim from McDavid's "A Checklist of Significant Features for Discriminating Social Dialects," in Eldonna Evertts's *Dimensions of Dialect* (Champaign, Ill.: National Council of Teachers of English, 1967), pp. 7–10.

consonantal replacements. We've talked earlier about his /j/ for /y/ in some structures. Indeed, /c/ and /š/ may also be employed. The result may be /cip/ (*sheep*), /jɛlo/ (*yellow*), and /šes/ (*chase*).

The problem is basically a foreign-language one, resolvable by pattern drill and laboratory work. Some of the structures are as nonstandard as any produced by a native speaker of Present-Day English, and yet they technically aren't English. Great Britain is experiencing similar problems in the speech of Pakistanis and Indians who have moved there, in addition to problems from cockney, a social dialect. A Puerto Rican or a Punjabi adult probably first studied English under a nonnative speaker, before moving to an area where English is the national tongue. For the child, early bilingual education would seem to be the best solution, where his teacher has been specially trained in bilingual, bicultural matters. Thereby he becomes fluent in English, systematically learning a standard dialect, which will then be employed when he studies subjects in that language.

Necessary Lexical Variations

Yet there will be some differences in every speech community. Otherwise, we must recast our definition of the term *language* as man's creative possession endowing each individual with a unique idiolect. One person will always know some words that his neighbor doesn't know. There are several kinds of necessary lexical variations, of which we'll consider only three.

The first is obvious, involving the technical words of a discipline. For instance, entomological dictionaries list almost a million scientific names for insects, usually constructed from Latin or Greek elements. Few of these words can be included in even the largest unabridged dictionaries. In the life sciences, a separate, generally accepted name is required for each of the myriads of species, not to mention the almost limitless variations within some species. One reason why the "Ivory Tower" scientist may not be liked is his inability to move from his specialized vocabulary and thoughts to more relevant matters in the human world. The technical words in his idiolect, shared by his London or Moscow counterpart, actually belong to the International Scientific Vocabulary, as adapted to English structure. Most other English speakers don't know them.

A second kind of variation is illustrated by Eric Partridge's *A Dictionary of the Underworld* (London, 1961). Though 817 pages long, this listing is admittedly deficient in American words. It's a historical collection of the special words of "crooks, criminals, racketeers, beggars and tramps, convicts, the commercial underworld, the drug traffic, the white slave traffic, and spivs." (A spiv is someone who lives by his wits without working.) What do *bindle stiff, bird dog, buzz-man,* and *dentist* mean? The vocabulary changes fast. After all, thieves' argot supposedly permits private communication through words possessing agreed-on meaning. Suppose the *fuzz,* an expression dating back to at least 1929, learns the criminal's intended meaning. The criminal usually drops the word or changes its special denotation. However, *weed* has been used as a synonym for marijuana since about 1918, although it often means a cigarette today. *Hemp* (Indian hemp, bhang, hashish) has an 1870 entry in *The Oxford English Dictionary.* Now it means marijuana. What about *pig?*

A last kind of necessary difference can be illustrated by a person living along the Mexican border, where he must periodically use words from the contiguous foreign language. Consider only his foods, forgetting his clothing, shelter, and customs. He's accustomed to using and hearing words like *arroz, chili, enchilada, frijoles, guacamole, taco, tamale, tortilla, tostada.* But all of them appear in *Webster's Third,* and the foods named are fairly common in metropolitan restaurants around the country. Aren't they simply loanwords that some people don't know or at least use very seldom? Possibly, but he also uses *carne, chalupa, pescado,* and *piña,* which aren't in *Webster's Third.* Many English speakers don't know any of the four, originally borrowed into Spanish from Nahuatl, a Uto-Aztecan tongue. He may enjoy *jalapeños,* now commercially canned and available in some cities outside the Southwest. The fact helps account for the inclusion of the word in *Webster's Third,* even if many Americans still don't know the pepper. Delicatessens and food-specialty stores help introduce the names of foreign foods to English.

Narrowing of Dialectal Differences

Despite the existence of these three special lexical variations and of regional and social dialects, dialectal differences remain

superficial. They may be somewhat narrower in the future. For one thing, the international spirit after World War II has apparently stopped the diverging. Heretofore, the general history of languages has been the eventual separation of dialects into mutually unintelligible tongues. We don't know how many languages there were in 449. Probably there were not nearly as many as the three thousand to six thousand estimated today, some of which have millions of speakers. If we count only native speakers, Mandarin Chinese comes first with 460 million, followed by English with 250 million. Hindustani has 160 million; Spanish, 140 million. Russian is fifth, with 130 million. In order, there follow German, Japanese, Arabic, Bengali, Portuguese, and French. Italian is the twelfth largest, with 55 million.

The trend toward dialectal separation seems to have been reversed. English, one of the five United Nations languages, is used by Americans, Britishers, Egyptians, Indians, South Africans, and many other people at the UN. Sometimes speeches are broadcast around the world. The UN has its own radio station in New York. Imagine the general dialectal "equalizing" that indirectly results when a Californian hears the English of a New Yorker, a Rhodesian, or a Burmese on the UN station. International radio and television carry the actual speech of a London longshoreman to America, and our speech back to him. Widespread use of movies and television in Canada, Great Britain, and the United States further contributes toward the growth of some uniformity. The flood of tourists also does its part, as do world commerce and international politics. Movement toward greater language uniformity may even be accelerating a bit.

Generalized Present-Day English

Since English is now becoming a little more uniform around the world, we might take a final look at it, apart from dialects and specialized vocabularies. A generative-transformational description of the language since 1900 would fill volumes. All we can do is point to a few conclusions and possible trends.

We can start by defining English as the twentieth-century outgrowth of everything developed and illustrated in our book up to this point. The kinds of changes observed since Old English are apparently still going on. Any native speaker should be able

to give numerous illustrations. Think of the Great Vowel Shift, for instance. Old English *ðē* became /ði/, before /ə/ generally replaced the vowel. Terminal schwas were often apocopated in the past. If /ə/ is lost in *the,* the whole word is lost, since the vowel constitutes the syllabic peak. Many British structures don't use an article. Is this the beginning of the end for *the?* Probably not, but the possibility can't be wholly rejected. One problem is our closeness to the language. We can seldom perceive at close range the systematic quality in what are seemingly isolated changes. Remember that the leveling of inflections, the neutralization of unstressed vowels in words of at least two syllables, and the GVS took centuries.

Deep Structure

What has happened to deep-structure rules since 1900? Certainly, almost nothing. They've changed little since 1100. The chief alteration was the expansion of the **Auxiliary** rule. The progressive passive developed in Modern English, producing "is being acted." The same **T-pass** can now create "will have been being acted," when applied to the complete deep structure. The meaning of the sequence is clear enough. Will this be a new expansion in the variety of modal-auxiliary structures? As for the modals, we've seen a reduction from the Middle English peak of ten down to *can, may, must, shall, will.* Today *get* is used as a defective modal. We "get beaten, get going, get married." We don't *"get go" or *"get have gone," comparable to "can go" or "must have gone." Will it become a regular modal? Or will combinations like *getaway* and *getup* mark the future grave of its unsuccessful attempt?

Along with the expansion of the **Auxiliary** rule, we've witnessed a steady tightening of word order. The deep-structure rules permitted the good Alfred to generate this sequence of prepositional units: **NP + Tense + V + NP + Prep-Phrase + Prep-Phrase + Manner + Place + Time.** However, he probably never said anything as complicated as "Jane encountered the boy from the west side of Boston suddenly on the campus at noon." Every optional prepositional slot is filled, creating a somewhat cumbersome sequence. We have to read it slowly, in order to associate "from the west side" with "the boy," and "of Boston"

with "from the west side." With what is *suddenly* associated? Obviously the word order is tightly controlled.

Such multiple phrases may lead to greater ambiguity or even misunderstanding. In the absence of inflectional signals, even a single phrase can be vague. Consider "Henry defended the law against royal prerogative." Without knowing the deep structure, we can't be confident of the meaning. Did the law prohibit royal prerogative, or did Henry oppose it in common law? The pairing of prepositional units leads to vagueness in "I will give her mortgages on the houses." "I will give mortgages on the houses to her" is a bit awkward, but clearly differentiated from the other meaning, "I will give (to somebody else) her mortgages on the houses."

Or try "They dismount to shoot, then remounting instantly ride away." Which is instant, the remounting or the riding away? By choosing the order "instantly remounting" or "ride away instantly," we can specify the intended meaning. Misunderstanding easily results from "The quarterback rejoined the team after a day's absence on Saturday." If he really came back on Saturday, the sentence needs recasting as "The quarterback rejoined the team on Saturday," with the phrase "after a day's absence" coming before or after it. These four sentences aren't presented here to challenge the expanded Present-Day use of prepositional phrases. We simply note that the tightened word order may result in vagueness or misunderstanding when there are multiple phrases. The speaker or writer must be careful not to misplace them in the structure.

We've also seen the steady rise in importance and frequency of conjunctions and prepositions. *As* has continued to be a prolific combiner. In Middle English there were compounds like *asfast, asmuch,* and *assoon.* Today we add a second *as* to these, giving connecting expressions like *insofar as*. Now the word may be giving us a new preposition, *as far as*. Structures like "as far as X is concerned" have been in the language for some time, but only recently have we started omitting "is concerned." Will *as far as* join the high-frequency club, or will it remain somewhat in the prepositional background like *during* and *concerning?* To this point, the history of our language has demonstrated that few high-frequency prepositions have been added since Old English. The "natives" still do most of the connecting of structures.

Lexicon

Despite this particular stability, the lexicon has continued to change. Items already in the language have sometimes shifted in frequency. The loanwords *litchi nuts* (Peking Chinese) and *sake* (Japanese) were recorded in English as long ago as 1588 and 1687, respectively. As those foods are now served in many restaurants and affluent homes, their names have a much higher frequency.

Items already in the language have sometimes changed in meaning. Let's demonstrate with a set of semantically related items. *Privy* (latrine) was used at least as early as 1375. It's now somewhat dialectal, particularly as a name for bathrooms without a means of automatic flushing of wastes. *Latrine* was used by 1642 for special rooms in barracks, camps, and hospitals. It's still employed in that context. *Chamber pot* was first recorded in 1705. It persists today, with some degeneration. *Water closet,* an expression for the small room where there's the modern equivalent of a chamber pot, entered the language about 1755. It's widely used in England. *Bathroom* came in about 1780 and still generally prevails in the United States. Lord Byron used *toilet* in 1819; today it may be losing out to *bathroom.* The London *Times* used *commode* in 1851 to name a small article of furniture enclosing a chamber pot. In America, *commode* seemed to be defeating *toilet,* until *commode* also started to degenerate. *Bathroom* may be starting to lose ground to *powder room,* at least for women. The sophisticated dowager who asks the hostess at a party "Where's the john?" or "Where's the toilet?" would have made a better impression had she inquired sweetly about "the little girl's room."

Euphemisms, a term for softer and more agreeable expressions, expand the lexicon. The garbage collector becomes a sanitary engineer in his union contract, or he won't work. The salesman is an agent. The laborer is a blue-collar worker, perhaps aspiring to be a white-collar worker until shirt colors changed. Now that men seldom wear white shirts, will the former white-collar man become a green-collar worker? Will the color of men's collars be changed back, to save all the semantic confusion? When the color of blackboards became green, the word didn't become **greenboard.*

Two familiar kinds of additions have expanded the language:

borrowings and varieties of coinings. The magnitude of the overall expansion can be demonstrated by the 331-page supplement to *The Oxford English Dictionary*, added in 1933 to the original twelve volumes. It lists words recorded only after February 1, 1884. Here are three from French: *garage* (1902), *limousine* (1902), *camouflage* (1917). Did World War I have anything to do with the borrowing of *camouflage?* Would we expect *jalopy* to have probably come from the French? Some others are *tango* (Spanish, 1913), *intelligentsia* and *Bolshevik* (Russian, 1914, 1917), and *robot* (Czech, 1923). One of the newest and financially most influential borrowings from Russian is *sputnik* (traveling companion, 1957). Note that it doesn't have the usual *-ck* pattern of English spelling. Another recent borrowing is *fedayee* (Arabic, literally one who offers himself for his native land). The plural is *fedayeen*. The word is seldom used in the singular, and then the form is sometimes *fedayeen*. *Sukiyaki, tempura,* and *teriyaki* have recently come from the Japanese, along with the foods they name. Such words belong to the general English lexicon, not being Americanisms or Canadianisms. They may be heard in London as well as in Honolulu and Vancouver.

Coining has greatly accelerated since World War II. Space exploration alone has added many words to everyman's lexicon. They are often compounds from existing elements. A few examples are *blast-off, countdown, feedback, to lock on, payload,* and *payoff*. Some have reduced the frequency of established items. The spaceship of Buck Rogers and Flash Gordon is now a *spacecraft* or a *space vehicle*. Formations with *space* are common, as in *-medicine* and *-station*. Formerly technical words like *apogee* and *perigee* are gaining in general usage. *Uptight,* from the prolific *up,* is a new compound. "Do your own thing" is as picturesque as some Old English combinations.

Blends, another kind of coining, are like compounds, in that they're words created from two or more existing elements. However, the elements are metamorphosed into a wholly new form. Did *dandle* originate from a slip of someone's tongue about 1530, when he meant to say either *dance* or *handle?* Naturally we don't know. Blending has contributed new words and is a useful process today. What are the elements probably composing *aniseed, chortle,* and *splatter?*

Acronyms, created from the initial parts of a sequence, are another fruitful source of new words. Here are some examples,

which many of us employ sometimes without knowing their origin:

AWOL—absent without leave
NATO—North Atlantic Treaty Organization
radar—radio detecting and ranging
snafu—situation normal all fouled up
sonar—sound navigation ranging
univac—universal automatic computer
veep—vice-president
WAC—Women's Army Corps
WASP—white Anglo-Saxon Protestant

Transformations

The chief recent alterations in the transformational rules seem to be in **T-infl** and **T-adv.** Analogy continues to operate, with more nouns and verbs being "regularized" in inflection. Remember the element *-ful?* It was added to *hand, cup,* and some other words long ago. For a time the plural suffix was inserted medially, giving *handsful.* However, "infixing" has never really been an English process. *Cupful* and *handful* are becoming somewhat frozen, no longer as easily divided for insertion of a suffix. Our usual inflection is accomplished by terminal addition, even if it's "zero." Will *cupfuls* eventually drive out *cupsful?*

Combinations like "brother-in-law, lady-in-waiting, maid of honor" are resisting the process. They are far from being frozen compounds at this point. Although we often hear "brother-in-laws," the form may not be quite acceptable. Will *"maid of honors" be the eventual plural? Certainly "maid of honor's" is the singular possessive. Other possible regularizations include the long-defiant borrowings *cacti* and *radii.* Will *cactuses* and *radiuses* drive them out? The Latin loanword *datum* seems to be thoroughly mixed up at the moment. Because people always want more than one datum, *data* may become the base, adding ø for the plural. Or will analogy operate, making **datas* the plural?

There may be a movement toward "zero" forms for plurals of number-words. Consider *dozen* as a kind of base in these structures:

Mary found two dozen Easter eggs.
Mary found two dozen(s).

In the first, *dozen* must take **ø**. It usually has *-s* in the second, but **ø** is acceptable. What if *hundred, thousand, million, billion,* or *trillion* is substituted in the second sentence? They often take **ø** in speech, with *-s* perhaps a little more frequent in writing. What about the future? Or let's try *heathen.* As long as it's a collective noun made from an adjective, as in "They're the heathen," **ø** is required. When an individual plural, it usually takes *-s*. Thus "He converted three heathens." Will **ø** eventually dominate? Will more American nouns become collective, having **ø** in the plural, like the British "The state have ruled in the matter"?

We've mentioned the added functions and greater frequency of **T-adv.** The expansion may be related to the more extensive use of sequential prepositional phrases. Here's a simple example of the required transformational order:

He **past** be at home then
past be he at home then (**T-ques**)
Where **past** be he then (**T-adv**)

Note that "at home" is wholly replaced by *where.* The product of the transformation must be "Where was he then?" There's no possibility of *"Where was he at then?" The rule itself, in a sense, makes this structure nonstandard. Or consider a deep structure of "He **present** be -ing go to town." The whole of "to town" is replaced. The product becomes "Where's he going," once **T-contr** is optionally applied. It can't be *"Where's he going to?"

"When" structures are derived in the same way. "Genghis Khan **past** die in 1227" becomes "When did Genghis Khan die," after obligatory application of **T-do.** If there are prepositional phrases of both place and time, **T-adv** can effect three results:

Where did Genghis Khan die in 1227?
When did Genghis Khan die in Asia?
Where and when did Genghis Khan die?

The adverb, of course, need not be a phrase. **T-adv** changes "I am going home" into "Where am I going?" Likewise, "I'm going now" gives "When am I going?"

Phonology

In phonology, all the processes that we've observed since Old English seem to be operating today. Consider simplification. Remember the loss of the /h-/ clusters, with only /hw-/ remaining in stressed position? Will /h-/ die in *what, when, where,* and *why* too? The leveling of certain /-nd/ clusters may be continuing. Even if still nonstandard, */wənərfəl/ is not uncommon in an unstressed position like "She's a wonderful gírl." Medial /nt/ is apparently simplifying to /n/ in words like *center* and *sentence.* Some foreign-language speakers, a little unsure of their English, pronounce an "nt" word as it's spelled, as do some careful enunciators. As a result, /sɛntər/ and /sɛntəns/ remain in competition with /sɛnər/ and /sɛnəns/.

Neutralization, often followed by syncope, continues. The schwa is still gaining in frequency. What has been syncopated in the common /váɪləns/? Is there syncope in /əproprəéšən/? If not, what has happened? Think of the British and New Yorker practice of omitting /r/ before consonants, as in *hard* and *horse.* When *hard* and *hod* are homonyms, has there been simplification of /-rd/? Interestingly enough, recent investigations of standard speech in the boroughs of New York indicate the possible restoration of /r/. A New Yorker can then be socially stigmatized by his /gɑbɪj/, rather than /gɑrbɪj/ or /garbɪj/. Some common elisions, among many, are /æsprɪn/ and /æmbləns/.

Other processes continue. There is metathesis in pronunciations like /pʊrti/, as opposed to /prɪti/. Clipping, or extreme apocope, occurs in *eco, gas, poly sci,* and *rhino.* What other names of animals have been clipped? In *specs,* from *specifications,* we find a small chain reaction. *Specification* is [+ sing]. The clipped form isn't **spec* for singular and *specs* for plural. It's [− sing] only, a logical change according to draftsmen, who always include more than one specification in any drawing.

Thus we reach the 1970s. The conclusions and possibly even the trends sketched over the past several pages may suggest some of the things that may happen to English in the future. Our book has shown that we can't predict the actual alterations. The only certainty is future change. The English of 2072 won't be the same as our language today. In the last chapter, we'll weave some of the major changes into generalizations about language in gen-

eral and ours in particular. These generalizations have underpinned the whole book. In that sense, as well as in their applicability to many situations outside the classroom, they constitute the most important part of the book.

Activities

1. Name two nouns that don't take an article when co-occurring with each of the following prepositions: *at, by, in,* and *to.* Examples are "at sea, by mail, in prison, going to court."
2. Find the American equivalent of these high-frequency British words: *beer, bonnet, calendar, guard, keeper, lounge suit, milliard, roundsman, page.* Try to find the British equivalent for these American expressions: *doctor's office, glue, lifeguard, lumber, orchestra seat, radio, scholarship, school recess, television.*
3. The British use /aɪ/ in the last syllable of *fragile.* What are two other words with *-ile?* What does the suffix mean?
4. Look up the history of the original settlers in your area. For example, New Englanders settled along the Great Lakes, Scandinavians went a little farther to the west, and Germans and Czechs went into Texas. Don't forget the Indians or any French, Spanish, or Dutch colonists. Are there any lexical effects upon the speech of your area today because of the language of its early settlers? What about transformational and phonological effects?
5. Lexical variation is the chief difference among the various American dialects. What word or words in your dialect do you use to name the following items taken from the Linguistic Atlas questionnaire: a small and narrow valley, a channel cut by a stream of water in a field or across a road, the best room in your residence, good (or bad) weather, and 10:45? What do the following words mean: *croker sack, cuppin, french harp, house stoop, spouts, sugar bush, whet seed?*
6. Argot and slang change quite rapidly. For example, in the Addenda to Partridge's *Dictionary* in 1961, there are four synonyms for a marijuana smoker: *herb, tea-hound, viper, weed-head.* More recent words relating to drug use

like *acid, grass, hemp,* and *trip* aren't listed. What other words used commonly in the public media seem to be recent additions to the drug argot? Make a list of some basic terms of the jazz musician.

7. Make a list of euphemisms for socially "offensive" words like *corpse, die, the poor.*
8. Collect three combinations like *brother-in-law.* Then look up their plurals. Do any have plurals made terminally? What's the plural of the number-words *ream* and *score?*
9. Make a list of common acronyms like *radar,* clippings like *zoo,* new compounds like *countdown,* and blends like *chortle.*

Chapter Nine

Language Attitudes

Certain generalizations can be made about the ways that language attitudes relate to human interaction and to practical application of linguistic principles and data to such interaction. Armed with these generalizations, we will be more enlightened about language matters and perhaps more tolerant toward other people as a result of that enlightenment. Twelve of these generalizations follow, with elaboration and illustration of each. They are adapted from Garland Cannon and Sumner Ives's "Some Generalizations about Language."

1. The term *linguist* no longer refers only to someone who has learned a great many languages. At present, it is used more often for someone whose field of study is linguistics —the study of language and of one or more individual tongues. There are several branches, such as historical linguistics and dialectology, and a few combinations with other fields, such as psycholinguistics and sociolinguistics.

A dictionary often includes more than one meaning for a given word. Such is the case for *linguist*. Many people use the term to mean "polyglot," unaware that *linguist* also belongs to the technical vocabulary of one who analyzes language professionally. Indeed, it names the analyst. The first question often asked a linguist is "How many languages do you know?" The questioner is often disappointed if the linguist doesn't speak at least four or five languages. What the linguist actually does is

work with data from one or more particular tongues, according to his purposes.

If he's a historical linguist, he might be concerned with the study of English from 449 to 1972. Since few people born earlier than 1870 are alive today, his data consist principally of written records. If he's a dialectologist, he can learn about Kentish speech today by consulting the *British Survey of English Dialects*. The Linguistic Atlas of the United States and Canada can tell him about Northern features of a Bostonian's speech in the late 1930s. Another linguist may require information from psychology, as in investigations of how a child learns his native language. If such data are needed often and in quantity, he should properly be a psycholinguist, someone trained in both psychology and linguistics and in their interrelationships. Sociolinguistics is especially important today, for sociologists are making discoveries about man's social nature. Purely linguistic information about peoples' dialects may be comparatively unrevealing unless enriched and extended by direct information about the human being doing the speaking. The combination of discoveries might then be made helpful to that human being.

2. A linguist may have no superior innate ability to learn languages, but his training enables him to learn them more efficiently.

So far as we know, no person has a mental block against or a special gift for learning a foreign tongue. His success or failure is often the result of motivation. If his motivation is poor, he probably won't master the language. If he intends to specialize in French imports, he can be expected to learn French rather quickly. He may realize that one effective method is simply to move to a native French area, where he can't use his own language. There he must understand and use understandable French. Otherwise, he may be unable to obtain food, clothing, and shelter. A linguist usually has comparable motivation, because some of the data may be required in his later research. His training leads him to concentrate first on listening to and interpreting native speakers of the language that he is studying, then on speaking to people in that language, and finally on reading and writing. He

knows the relationships among the components of language in general. The knowledge helps him perceive the particular, important points within the one he chooses to learn and describe. Systematic attention is given these components.

3. No language seems to be intrinsically easier or more difficult than any other. A child learns the language of his social environment, and the variety that he learns is the dialect of his environment. Children seem to learn all tongues at about the same rate. On the other hand, an adult is likely to learn some languages more easily than others. For example, an adult speaker of English can master Dutch more readily than he can master Hungarian. A Hungarian can become fluent in Finnish more easily than he can in English.

Children seem to have no trouble in acquiring their native language. All tongues are apparently of equal ease or difficulty to a child who is learning to speak, since he knows no other language that might interfere with the acquisition process. To be truly bilingual, a child should learn two languages simultaneously as his native ones, beginning about the age of eighteen months. He usually acquires his mother's dialect, slightly modified by playmates' speech. If his mother's dialect is Virginia Piedmont or cockney or London Standard, the child's will be too. Location doesn't seem to affect the rate of acquisition. He learns at about the same rate, whether in the city or the suburbs, whether in an affluent home or in a ghetto apartment. The only real necessity is someone to speak to him—to provide the speech models and to encourage his efforts at imitating and then responding to the speaker. The presence of books, newspapers, and the whole cultural world are generally irrelevant at this stage. Television may assist the acquisition a little, although all the child can do is listen rather than respond. If he does answer a question asked by a character on television, there's no one to encourage and gently correct him.

The adult, however, may be handicapped by his knowledge of his native structures. One's speech habits are generally fixed by about the age of ten or twelve years. When a person begins

studying a second language as an adult, he's likely to fill in "gaps" in this second tongue with data from his native speech. Certainly we've all observed people who speak a foreign language with many "errors." These errors are actually their native structures, presumably substituted when the speakers simply didn't yet know the particular foreign structures. The answer to that problem, of course, is simply more practice with the new structures.

On the other hand, our knowledge of a first language can sometimes facilitate the learning of a second. An English speaker can learn Dutch fairly easily, because English and Dutch derived from the same tongue, Low German, which branched into those two dialects by A.D. 499. Modern German, which derived from High German, should be a little harder for the native English speaker. Learning Persian should be much harder for him. Germanic and Indo-Iranian, the predecessor of Persian, developed out of the Indo-European family. The Iranians' forebears left the Indo-European homeland by at least 2000 B.C. Consequently, major similarities between Persian and English aren't plentiful. If the English speaker studies Hungarian, he doesn't even have these similarities on which to build. Finno-Ugric, from which Hungarian and Finnish developed, was apparently never in the Indo-European linguistic line. Or if so, the time was so remote that all traces of common descent have disappeared. Hungarian has an elaborate, large system of inflections for case; English has few case inflections, of a rather different kind.

4. Each language is a system of symbols by means of which persons can communicate. To discuss the actual communication process, we can conveniently divide the system into three interacting and somewhat overlapping components. In popular terms, these are the phonology, grammar, and vocabulary. In speech, a sentence is manifested by units of the phonology. If there is a writing system for the language, a sentence may be represented by units of this orthographic system.

Basic to the definition of *language* are the system and symbolism within the particular tongue. Suppose that each of the

infinite number of potential sentences of a language has a unique structure, rather than being a member of one of a few basic structures. Obviously, we could never master the language, not even as our native tongue. The symbolism within a language is equally recurrent. A given word has only a limited number of meanings at a given time. The particular meaning intended is often conveyed by the context. This system of symbols wasn't devised by a linguist to give him something to do or to fill in an otherwise empty spot in one's school curriculum. The purpose of language is communication. The more we discover about it, the more we know about the whole marvelous process by which man learns and uses language with his fellowman. He seems to be the only species with that capacity.

Most people know that a language has three components. When someone speaks, we hear the phonemes, or basic sound-units. When someone writes, we see the graphemes, or orthographic units, representing the phonemes. Our interpretation of either requires knowledge of the grammar, or syntactic component. This provides the basic structures of the language, conveyed by words from the lexicon. The words are transmitted as recurrent sounds from someone's mouth along sound waves into someone's ears. Because of such interaction and overlapping, it isn't always easy to separate the components cleanly.

5. After a community begins to use writing, somewhat different conventions for written and spoken expression develop. Thus when we learned to convey ideas in writing, we weren't simply mastering the transcription of speech. We were learning the use of a related, but somewhat different, mode of expression. The term *writing* is itself a little confusing, in that it's employed both for transcribing and for more artistic creation.

Many of the world's languages still don't have writing systems, although those studied in schools usually do. The major problem in achieving literacy derives from the fact that writing isn't speech. It's a visual representation of speech, with relatively fixed conventions like spelling and punctuation. Pronunciation of a word can vary considerably, even within a regional dialect.

Decisions as to what constitutes mispronunciation are, therefore, difficult to make. Variation from the established spelling of a word, however, constitutes an error, a point on which everyone can agree. More importantly, writing isn't talking with a pencil. The casual "Hungry?" and "I been thinking 'bout it" may be acceptable in informal speech, but on paper the structures should be fleshed out to "Are you hungry?" and "I have been thinking about it." Nor is *"He kneel down yesterday" acceptable in writing, although in the spoken sentence the suffix /-d/ is omitted because of the following /d/ of *down*. In short, we must master a somewhat different means of communication, if our written product isn't to be rejected. This writing isn't the kind found in Shakespeare's dramas, which constitutes another meaning of the word.

6. Each language reflects the needs of some community in thought and expression, and the language changes over a period of time. The change is not necessarily improvement or deterioration. Our English today is quite different from Old English. The English of Alfred's day was, so far as one can know, adequate to communicate the speaker's wishes, just as ours is today. Great literature has been composed in each linguistic period. Literary value isn't resident in a language; it results from the way the language is used.

Occasionally we hear someone characterize a given people as primitive, measuring them against some shadowy, technological standard. What he has actually done is to make a value judgment, based upon his tastes and experience. Does a people's state of advancement rest upon the capacity to discuss nuclear physics or to utilize the knowledge to create a bomb for razing a city? If this is so, then some tribe living along a New Guinean river could be considered primitive. But like all other men, they have language, which is neither backward nor new. For instance, they have separate words for the myriad of fish caught in their traps. Many of the words have no English equivalent, since we've never seen the particular fishes and have no need for terms naming them. The people have terms to differentiate the fresh-

water fishes from the saltwater variety, as well as the poisonous from the nonpoisonous, and then the edible from the inedible, nonpoisonous ones. They probably don't have a separate word for the automobile yet. The rapid development of hydrofoils and similarly powered land vehicles may mean that the people will soon have to put together something like "that curious thing that contaminates your air and sinks your catamaran if you don't look sharp."

Their language is adequate for their needs. In the abstract, it's no better or worse than any other tongue. The lexicon is sufficiently elastic to add a whole new set of words describing the varieties of snow if the New Guinean weather and setting were to change to that of Alaska. In turn, while peeling off layers of clothes, the Eskimo can just as easily add words that describe monsoons, jungles, and catamarans.

As a people's needs change, so does their language. Few if any linguistic changes are predetermined by any person, organization, or nation, and then systematically effected. Existence of a French Academy or a Spanish or an English version probably has little or no effect upon the ongoing alterations in sounds and structures. Whether the purpose is to block or accelerate some change, or to drop a word from the lexicon or to insert one that some "regulator" has created or wishes to borrow from another language, the effort seldom succeeds. Linguistic change inexorably takes place over the years. The lexicon, phonology, and syntax are just different after a while. A speaker of Old English would undoubtedly not be able to communicate with us, if he could be resurrected and brought forward in time. Most of his lexicon has been replaced and greatly expanded by Latin and French borrowings. Too, the "native" survivors have usually changed in sounds and meaning, and English word order today is much tighter.

Scholars once uncritically theorized that linguistic change was neither for the better nor for the worse. Actually, there's probably no more possibility of proving this thesis than of disproving the view held by some laymen that English has badly deteriorated in one period and has improved in some other period. We can insist, however, upon the adequacy of one's language to communicate one's wishes, so far as we can ever know. There's no evidence for hypothesizing a less adequate language for the uneducated, unsophisticated Londoner of A.D. 900 than for the

elegant Renaissance courtier or for us today. If we boast about our larger lexicon, the West Saxon could point to his greater array of inflections. If we beam at our greater variety of modal-auxiliary structures, he could note his freer word order. English today simply doesn't represent a decay and regression in the language since the Renaissance. Representing different moments in the history of our tongue, the two are chronologically rather than qualitatively separated.

People do feel affection for their favorite words and pronunciations, rejecting others' variations. Purists have condemned, for example, the use of *-wise* to form new adverbs *(dollarwise, truthwise)*. Apparently they're unaware that *-wise* (manner, fashion) formed numerous combinations in Old and Middle English. Each was originally two words, before being permanently conjoined. This freezing helps disguise the presence of the suffix in *anywise, crosswise,* and *leastwise*. Certainly personal rejection of *-wise* as a current suffix is one's prerogative. Anyone who detests the element can even avoid the word *likewise,* if he wishes, but he shouldn't mistakenly try to stop others from employing *-wise* otherwise.

Prescriptivists have also condemned the supposedly recent emergence of *like* as a conjunction. However, that use dates back at least to the Renaissance. We might smilingly wonder whether the adman who created the notorious cigarette commercial didn't borrow something from 1530: "Ye have said lyke a noble lady ought to say." Actually, many distinguished writers have used *like* as a conjunction since the Renaissance, even if the adman's creation is somewhat less than literary. A cigarette manufacturer has shrewdly capitalized on them cancer-prone individuals which prefer good taste to good grammar. Apparently *like* was a respectable conjunction in the Renaissance, when there weren't any purists to attack it. It may be gaining in frequency today.

Purists have attacked other historical developments of the language. They may be surprised to know that there was some confusion between the predecessors of the near-homonyms *lie-lay* and *sit-set* in Old English, with no end to the problem in sight. "It is I—it is me" is another pair of structures that have been freely substituted for centuries. A few decades ago, the chief goal in English of many American school boards was to develop the student's capacity to make a kind of artificial dis-

tinction between *will-shall, lie-lay, sit-set, likely-apt,* and to avoid double negatives, split infinitives, dangling participles, and run-on sentences. The teacher was sometimes thereby forced into a noncompassionate, prescriptive mold by such local and sometimes even national attitudes. The fact, of course, is that some of these structures are changing. One wastes his time in attacking alterations to which he and an eighteenth-century prescriptivist object, because the changes will probably take place anyway.

A closely related view involves written literature. If a survey of all extant Old English literature reveals no Chaucer, Spenser, Shakespeare, or Milton, we still shouldn't make a value judgment about the language of the period. In view of the ravages of the Danes in the eighth and ninth centuries and of the German V-2s in the twentieth, the manuscripts of a dozen epics like *Beowulf* may have been lost. Perhaps not even allusions to their titles survive in extant manuscripts, much less to the authors' names. At least we have knowledge of the names of Greek dramatists like Phrynichos and Pratinas, even if we know little more about them or their work. Anyway, language eminence can't be demonstrated by some super literary hierarchy like the one consisting of the four authors mentioned above who wrote from about 1340 to 1674. It's impossible to make a serious comparison of diverse artists of different periods and then to make an overall evaluation in the abstract. Are Spenser's sonnets better than Milton's pastorals?

What can be demonstrated is the creation of excellent written literature in each of the usually accepted periods. The literature of Alfred's day, great or mediocre, didn't result from some quality resident in the language. Cynewulf and Cædmon would probably have been fine poets, regardless of their dates, homeland, and native tongue. Had they been Chinese living at the time of Confucius, the cultural and intellectual influences would naturally have been different. It's still likely that their literary genius would have manifested itself. A writer uses language as his tool, much the way a sculptor shapes clay or a composer blends tones into chords.

Even if an artist is of the quality of James Joyce or Dylan Thomas, he can't vary too much from the norm of his linguistic medium. If he does, then he obscurely talks to himself, because his medium is unrecognizable to others. This charge might be made against some of the most experimental moderns. The lin-

guistic norms of any period are sufficiently elastic to permit almost unlimited creativity, as long as the artist stops short of excluding others' understanding. As Robert Frost once told a poet noted for his obscure verse, "If you have a secret, keep it to yourself."

Any aesthetic, moral, and intellectual values in an artist's work don't result from the happenstance of the author's living at a time of especially pregnant language. The values derive from the way the language is used, in oral or written form. After all, only man can translate into speech his unique sensations and sensibilities. The capacity of language to convey such human feelings and compassion is an important attribute, beyond the basic one of communicating facts in an objective way.

7. So far as we know, every language has been used for artistic expression, including those that have no writing system. Oral traditions of song and story may be rich and sophisticated. Intricate poetic forms have been used by men who were not literate. In Iceland, as late as the tenth century, the law of the community was usually not written down, but carried in the memory of a trusted individual. Artistic expression and record-keeping are probably facilitated by writing, but they aren't entirely dependent on it.

Anthropologists haven't discovered a single tribe that doesn't create and enjoy literature. With highly developed themes about life and death and with cultural harmony, the songs may be revered by the tribe. Indeed, the fluent storyteller often has an exalted status precisely because of his literary productions. The native rhythms are often quite intricate. Sometimes the creations become essentially a daily newspaper. Thus the Copper Eskimos enact all important happenings in a dance-song, daily performed for the tribe. We seldom know about such literature because it hasn't been put into writing, much less English orthography. *Beowulf* is a notable exception, in that it was orally transmitted for many generations before being written down for posterity about the tenth century. Celebrating heroes of the fifth and sixth centuries, the epic includes a complex record of kings and deeds.

The record was apparently carried in the memory of the original writer and then transmitted by successive storytellers until it was eventually put into writing.

Ballads like "Barbara Allan" were transmitted orally for a long time before finally being transcribed. The rich mythological tales of the Icelandic epic *Edda* must have been carried in the memory of the storytellers until they were committed to writing after 1250. One of these tales, "Sibyl's Prophecy," is one of the greatest of all Germanic poems. Did those nonliterate Icelanders have better memories than we do? Certainly man has an amazing capacity to remember when he must. Do poems suffer from oral transmission, either in literary quality or in fidelity to the original? A conceited guess is that artistic expression and historical records may be more accurately and permanently preserved by writing. At least few if any Shakespearean scholars attempt to commit to memory the dramas that they revere and criticize. So, too, few mathematicians are likely to discard their tables of sines and cosines in favor of memorization.

8. To say that a language is phonetic or unphonetic is to confuse speech and writing. The term *phonetic* refers only to language sounds. An alphabetic writing system may represent the sounds more or less consistently, and thus may be described loosely as more or less phonetic. English phonology has changed rather dramatically, but the orthography hasn't changed comparably. Thus our alphabet no longer consistently represents the sounds of spoken English. Several unsuccessful attempts to reform English spelling have been made in the past few hundred years. As a result, the letter *i* may now represent a diphthong, as in *bite,* or a monophthong, as in *bit.*

Phonetic, the adjectival form of *phoneme,* refers to speech. If the graphemes of an alphabetic spelling system represent the sounds of the language with consistency, it might loosely be termed "phonetic." That is, every occurrence of a given grapheme must always stand for a particular phoneme, and every occurrence of that phoneme in speech must be spelled by that grapheme. Consider *c,* as in *call, cell, charlatan,* and *chase.*

Need we go further? Although English consonants haven't changed much since A.D. 449, *c* represents /k, s, š, c/ in the above four words, respectively. Vowel spellings are much more inconsistent. Try these items: *mat, mate, met, father, meet, mitt,* and *might.* Six phonemes and one diphthong are represented by *a, e, i.* The diphthong is in what word?

Should we condemn our spelling system and then construct a truly consistent one? The seven spellings above do record the phonological history of those words. They indicate that the vowel sounds haven't really changed in *mat, met, mitt,* and *father,* but that the others have. Why hasn't the orthography of the "changed" words been altered accordingly, so that it will represent the shifted vowels? Actually, some consonant graphemes have been changed. Vowel spellings, however, have hardly been touched since the Great Vowel Shift that caused most of the changes beginning in Middle English. Both Britishers and Americans have proposed reforms, notably Franklin's and Webster's attempts in the late eighteenth century. Webster urged the omission of silent letters, resulting in forms like *bilt* and *relm.* Was he successful? Some of his other suggestions are chiefly responsible for the differences between British and American practices today, like *honor-honour, rivaled-rivalled,* and *theater-theatre.* Overall, English spelling is essentially that of 1500.

9. Every linguistic performance utilizes resources from the grammatical and the semantic components of the language. It's manifested by units from either a phonology or an orthography. Any serious theory of meaning, or of linguistic communication, must consider the interaction of these two components with the phonemic or orthographic units. Ideally, there should also be consideration of such additional information about the language as is feasible and relevant in the local situation.

Consider this sentence, "Are these three apples góod?" What information is communicated? Let's sketch the syntactic part first. The sentence derives from the deep structure "These three apples are good," as transformed into interrogative order. The primary stress on *good,* as opposed to its possible placement

on any of the other four words, greatly narrows the meaning. We know that the apples exist, as opposed to having existed in the past, in which case *were* might have been employed. They're here, rather than there, or we would have used a word like *those* instead of *these*. The quantity is three, and the objects are apples. None of these lexical points is really at issue. The emphasis is on *good,* as opposed to adjectival alternatives like *bad* or *green*. The *-s* suffix of *apple* informs us that the subject is plural, requiring the same number for the demonstrative, cardinal, and verb. *"This one apples, this three apples, these one apples" can't be accepted. Nor can we say *"Is these three apples good?"

Certain semantic information is needed for the selection of the adjective modifier, which must match its noun. *"These three apples are virtuous" is unacceptable, because apples don't have the quality of virtue. Men do. The semantic component is further involved in specifying the particular meaning of each word, from the total possible meanings of each. *Good* often means "well-founded," but our context rejects this possibility. What are the three different meanings of the word in a "good dictionary, good doctor, good time"? It means "suitable or edible" when co-occurring with *apple*. Each of the other words in our sentence poses a comparable choice of meanings.

Both the syntactic and the semantic components provide information for the orthography—that is, the written form in which the sentence appears. There are no variant spellings for the five words, and a question mark must conclude the structure. If someone said it to us, dialectal variations are possible. If the speaker intends the plural form of *this* but accidentally or habitually makes it a homonym with *this,* we'll hear the nonstandard *"Are this three apples good?" In short, serious consideration of linguistic communication must recognize the sometimes inseparable interaction of syntax, lexicon, and phonology. When we study the system of English, including its written representation, we must include all three components. A bit of dialectal information might also be useful to students of language, especially when it applies to the local population. Someone from New England might usefully learn that his use of *tonic* for "soft drink" isn't paralleled in Virginia or Florida. What does *tonic* mean there? If he lives in a community of widely ranging incomes and education, he should be aware of the social stigma that is often arbitrarily attached to the use of structures like *"done finished."

10. The use of the term *grammar* to designate a part of the total language system prevents its use as a general term for matters of usage. These may include variations in pronunciation and vocabulary, as well as in inflections and syntax. The English language is a general system; yet whenever it's spoken, some dialect is used. Each dialect has at least two dimensions—regional and social. In the expression of intended meaning, one dialect seems to be as efficient as any other, within the needs of the group that uses it.

There are several, sometimes confusing definitions of *grammar*. Suppose a person remarks self-consciously "I'd better watch my grammar." He is referring to the way he uses his language in front of someone who he thinks is an authority, who'll evaluate his speech by mysterious but definite standards. If we hear "I can't find my grammar for the next class," a textbook is intended. The word also means "the study of the classes of words, their inflections, and their functions and relations in the sentence." These three and other meanings suggest *usage* to be a better term for describing specific variations within a language.

In general, language changes by the doctrine of usage. Thus if all English speakers begin pronouncing *either* with /aɪ/, never using the monophthong /i/ again, usage will eventually render the latter pronunciation obsolete. Certainly this hypothetical change won't be effected because purists may feel that the word is more elegant when uttered with /aɪ/, as President Franklin D. Roosevelt did. What happens is that people unconsciously change their language as they use it. If everyone adopts *spacecraft* instead of the earlier *spaceship*, lexical usage will kill the latter. If we all elect *dove*, never using *dived* again, the former inflection will naturally dominate.

Syntactic alteration is always the slowest of all. Usage could enthrone *get* as a modal, like *may* and *shall*. Presently *get* is a badly defective modal, if one at all. We can't say *"got go" or *"got have been heard," in contrast with "might go" or "should have been heard." All we can do is "get clobbered, get lost, get started." Or if everyone starts electing "All he can do is study," instead of "All he can do is to study," the two-verb combinations presently usually connected by *to* might begin their eventual path

toward obsolescence. Only usage can dictate such syntactic changes.

Usage, of course, involves data from individual native speakers. Each time someone speaks, he utilizes two kinds of dialects. Normally he can be expected to use the phonology, lexicon, and syntax of his geographical region, collectively termed his regional dialect. Simultaneously his speech identifies his social dialect, the aspects of which can be quite numerous. Consider age. The eighteen-month-old baby may call his father *da-da,* hardly an expression geographically identifiable. Sex may inhibit the male from describing things as *darling* and *precious,* word choices more often found in female speech. Social and economic background plays a major role. The son of an affluent father who has taken him around the world twice and has showered every financial and cultural advantage upon him can be expected to develop a somewhat different speech from someone without this experience. That is, a person's social class has a great deal to do with his language. The extent of his formal or self-education is especially influential. It is constantly reflected in what he says and writes.

Despite the influence of age, sex, social class, and education, one point seems clear: the dialect of any one person is apparently as efficient for communicating his meaning within his peer group as that of anyone else in his own peer group. When trying to imitate another group's speech, the phony always stands out from his opening words. Our generalization that dialects are equal in the abstract parallels the larger generalization that languages are similarly equal in the abstract. There's no reason to expect or even think that some tongue or dialect should ever bow to another.

11. A child's language is a reflection of his social environment. A dialect can be changed deliberately only through a great deal of effort. Hence, linguistic usage is one of the primary marks associating an individual with a group. Furthermore, judgments about a social group are often assigned to the usages associated with it. For these reasons, complete rejection of the usages may be resisted, for it implies rejection of group identification. Mastery of an additional dialect increases one's freedom

of movement within the social structure. There seems to be no connection between race and dialect, except as race may influence social discrimination or grouping.

In an area of considerable social stratification, like New York or San Francisco, a person's dialect directly reflects his social class. Even when there are more mobility and leveling of class lines, a person's dialect will still reflect his social environment, but won't be as reliable in indicating whether he lives "on the wrong side of the tracks." Presumably everyone adopts the speech patterns of his mother and his overall environment, often unaware until almost adulthood that these patterns identify him socially. The patterns reach all the way into syntax, in what we can loosely call "levels of usage." Three generally recognized levels of usage are vulgate, informal educated, and cultivated educated. To what level does a simple structure like "I put it down yesterday" belong? Probably to all three, since there's little opportunity to vary the inflection, agreement, and the like. Consider these two sets of examples:

(a) *vulgate:* I ain't got no change
(b) *informal:* I don't have any change
(c) *cultivated:* I have no change

(d) *vulgate:* She don't know nothin'
(e) *informal:* She doesn't know anything
(f) *cultivated:* She knows nothing

In structure (a), the double negative and *ain't* are nonstandard. Note that *ain't* means "have not" here. When it means "am not, are not, is not," it's still disapproved of by many people but isn't nonstandard, as in "I ain't going." The Linguistic Atlas findings reveal that many educated people use *ain't* in speech, especially in structures like "ain't I." Another double negative, plus failure to maintain person-number agreement, identifies the speaker of (d) as belonging to the disadvantaged.

Compare (c) and (f). The cultivated speaker has been taught to be perceptive in his language choices. He can move easily to informal. In fact, in his home he'll ordinarily use informal speech. His writing will be cultivated when the occasion requires. The informal speaker is also perceptive about his language, being

especially careful to avoid nonstandard usages. These may stigmatize him in his own group and will almost certainly do so in the cultivated group that he usually, whether conscious or unconscious, seeks to impress. Generally, he adheres to informal speech, not really commanding the cultivated level, despite a desire to do so. The vulgate speaker commands only the one level.

Of course, there's nothing intrinsically wrong with the double negative. Chaucer and Shakespeare employed it constantly for emphasis. Nor is there anything "incorrect" about using *ain't* to mean "have not," failing to maintain person-number agreement with *be* and the verbs, or omitting noun suffixes as in *"two boy." Public taste simply rejects these structures today; it may embrace them tomorrow. Meanwhile, people must publicly conform to that taste, if they don't wish to face social rejection. A decade ago the "four-letter words" were quite obscene, so degenerated that only speakers of vulgate supposedly used them. Now, although some people are still somewhat appalled by these words that were apparently proper in Chaucer's time, we sometimes find them in the speech and even in certain writings of the cultivated. Will the words become socially respectable? Who knows?

People often deliberately seek to "improve" their language by trying to imitate the dialect of the higher social class. If so, why are there still such wide social variations in speech? Linguists have fairly well determined the above-mentioned structures and a few others as marking nonstandard, lower socioeconomic speech. Why not simply distribute the list to every person in the country, with an explanatory note about the arbitrary stigma presently attached to these structures? Society might promptly choose a new list by which to help place people socially. There are two serious reasons why the distribution might not help much, although certainly it could do no harm.

The first is the extraordinary difficulty in changing one's dialect once it's formed. Certain structures will be so much a part of a person's language that they may unconsciously enter into his speech when he is talking rapidly. That is, when we talk spontaneously, we're unable to preselect every word of every sentence in advance. What if one conditions himself to talk slowly, preselecting every single word he says? Such an attempted psychological restraint upon his language might crush out all the creativity and individuality that make language a purely human

attribute in the first place. In one sense, deciding to make permanent alterations in one's speech is like requiring the left-handed child to become right-handed. We know of the consequent psychological repercussions. By contrast, if the left-handed child is lovingly encouraged toward ambidexterity, great improvements will be made. At least baseball pitchers always fear the switch-hitter.

The second reason is that the speaker may resist permanent rejection of his social environment. After all, judgments about his language imply the same judgments about his peer group. The vulgate speaker who decides to abhor the double negative will hear it in his mother's speech. She will be a constant model for him unless he rejects her. He literally must shut himself off from his parents and whole peer group. Even if he succeeds in changing his dialect permanently, he may himself be rejected by his old friends. Nor does mastery of informal educated speech insure his acceptance by that group.

The preferable solution is mastery of a second social level of speech for use in that group, while retaining the original level for use in one's original environment. Social mobility is unquestionably improved for those who command two levels. Another conclusion is that race and language don't seem to be related, beyond the unfortunate fact that race is still sometimes a factor in acceptance or mobility. When it is, the Negro vulgate speaker may thereby be kept from hearing the informal or cultivated dialect frequently enough for either dialect to have any effect upon his speech.

12. Variations in language may be stylistic as well as dialectal. These are the adjustments a person makes to his audience, to factors in the immediate occasion, and to his mode of expression, which may be speech or writing. A customary set of choices made on this basis constitutes a functional variety of the language. These functional varieties are resources. A person's ability to choose the variety appropriate to the immediate situation is one measure of his command over the resources of the language.

Whoever commands all three dialectal levels is in an enviable position. Few people are sufficiently versatile to do so. Com-

manding two levels is important, for a person can then probably pass the major linguistic test that may confront him several times daily. The test is complex, requiring harmonious choice of language motivated by three social factors. They are the speaker-hearer relationship, the occasion itself, and the mode of expression. If the speaker and hearer are social and professional equals, this factor may be insignificant. If the hearer is our employer, someone who can affect our future, we're handicapped unless we can adjust our dialect to his.

The occasion considerably affects the relationship between the speaker and the hearer. Whether it's a formal wedding, a job interview, or a chess game, the occasion helps determine the words and structures chosen. As one example, the groom is congratulated and the bride is given best wishes; she isn't congratulated. Again, although an interviewer may do most of the talking, he'll probably be watchful for vulgate structures and other behavior that he judges unsuitable. In a chess game, the technical names of the pieces and of the moves naturally dominate the conversation.

If the mode of expression is speech, there's usually a chance to correct a misunderstood meaning or a mispronunciation. We can employ smiles, gestures, personal appearance, and other nonlinguistic factors to enhance our persuasiveness. When writing is the mode of expression, the danger of violation of spelling and punctuation conventions appears. We have to observe these conventions, however poorly the graphemes may represent the actual sounds of a word, or however overlapping the conventions for commas and semicolons may be. Since there's no opportunity to correct anything once the final draft is transmitted, all revisions should be carefully made.

Now consider the reader. He may be in a bad mood, he may receive the communication at an inconvenient moment, or he may be tired or hungry. Such unpredictable factors significantly influence his interpretation and reaction. In addition, the questions of occasion and author-reader relationship arise again. What reactions might be expected from these two opening paragraphs from letters of application:

> Jackie Nolan, who lives down my block and who empties wastebaskets nights on your floor, tole me he heard you got a job opening for a countant. I ain't had much work in this line,

but I studied about counting in P.S. 34. My teach said I was good in it. So . . . .

Jack H. Nolan, Jr., Esq.—a night attendant in the firm of Doran, Rolan, and Smith, Certified Public Accountants—has informally apprised me of the possibility that your firm may have a position for an accountant which does not entail extensive professional experience. May I be so bold as to submit my application herewith, happily presuming that there is, indeed, an opening and that you are now entertaining applications for it from candidates like myself?

Right, the personnel officer throws both letters in the wastebasket, unless public-relations policy requires a polite, form-letter rejection, signed by his secretary in his name so as to save time. The first applicant isn't educationally qualified. The second may not be either. He is so self-consciously attempting the cultivated level, which he clearly doesn't command, that he may be rejected as rapidly as the first applicant. His letter is much too formal, even pompous.

The point is simple: the effective speaker or writer commands a flexible set of choices within at least two social levels. Thereby he can marshal appropriate combinations for most situations encountered. Ideally, he's never doubtful of his language, but is confident of its unlimited creativity and power to persuade anyone in any situation. Of course, probably no one has perfect, persuasive powers of language. Still, a person's ability to make the appropriate choices in varied situations, which we'll term a functional dialect, is one measure of his command of the resources of his language.

Activities

The twelve generalizations above apply to situations far beyond our factual history of English. For example, they provide a partial rebuttal to laymen's unfounded views about language that we hear on every hand. Suppose someone announces: "! is a bad word. It deals with the body and so is unclean and immoral. It's an expression that no decent person knows. Cree Indians use words like that all the time. That's why they're Crees." How do we rebut?

The speaker's intolerance isn't disguised. His information about the Cree use of words to name parts of the body is accidentally correct. However, Crees consider the items useful, without social stigma whatsoever. The English item that has socially shocked him has undoubtedly been in the lexicon for many centuries. For the time being, public taste condemns it. He has every right to avoid the temporarily degenerated word himself, but we can laugh at his ridiculous effort to make it logically "bad." His logic would also eliminate *mouth, heart,* and *leg.* How would he answer the doctor's question of "Now where's the pain?" According to his logic, is he a "decent person"? What he needs, of course, is understanding, both human and linguistic.

There follows a list of comparable laymen's assertions. Each contains just enough truth to make it deceptive, but is basically untrue. Explain the fallacy and rewrite each assertion to make it linguistically valid. The first five are deliberately simplified, in that the premises are stated. Some of the later ones are rather subtle and complex.

1. Language is composed of words and letters. Since speech comes from these, we should pronounce words as we spell them.
2. Language is logical. For example, two *no*'s make a positive. Therefore, double negatives must be avoided.
3. A form wrong on one occasion is always wrong. As an illustration, "I been thinking" is always wrong. Conversely, a correct form is always correct. Thus *shall* is used in first person; *will,* in the others.
4. Certain forms are normal and fixed. Any deviation is abnormal and thus to be corrected. We must eliminate the verb-form *dove* from the language.
5. Language will continually deteriorate unless the educated and cultured exercise their responsibility of keeping up standards.
6. John is stupid. He says, "He don't." He's probably rude and immoral too.
7. The Hupas (speakers of an Athapaskan language in northern California) are ignorant, lazy, and uncultured. That fellow is a Hupa, and he can't read or write his language.

8. You English speakers are insensitive tactually. We Eskimos have a series of different words for what you inaccurately call "snow."
9. Frenchmen are illogical. They say "Je ne veux pas d'eau."
10. Arkansas Ozark is a vulgar corruption of the real English language.
11. English is decaying badly. Just imagine: "Something tastes good *like* it should."
12. English is a great language; Zulu, poor. We have more than half a million words; Zulu has only a few hundred.
13. English is really a Romance tongue. Most of the words are from Latin or French.
14. Spanish is really easy. I'm a New Yorker, a long way from Mexico and certainly poor at languages. Even I learned Spanish.
15. Your Chinese is really a difficult language. We even have a saying, "It's Chinese to me."
16. Our Mandarin Chinese is the most important language. We have 460 million speakers. Your English is poorer. You have only 250 million.
17. German is ugly and guttural. At least it's a scientific tongue.
18. I don't want to study Twi, a language spoken in the Gold Coast. It's not a poetic tongue. Nor is it beautiful like Latin, or musical like Italian.

Appendix A:
Phonological Symbols

stops	p-b				t-d		c-j	k-g
fricatives		f-v	θ-ð		s-z		š-ž	
lateral					l			
apical					r			
nasals	m				n			ŋ
glides				y	h	w		

vowels	front	central	back
high	i ɪ		u ʊ
mid	e ɛ	ə	o
low	æ	a	ɔ ɑ

Appendix B:
Deep-Structure and Transformational Rules Of Present-Day English

Deep-Structure Rules

Sentence → Sequence + Intonation

Sequence → (Conjunction) Noun Phrase + Predicate Phrase

Predicate Phrase → Auxiliary + Verb Phrase (Place) (Time)

Verb Phrase →
- *be* + Predicate
- Verb + Sentence[1]
- Verb + Predicate
- Verb (Noun Phrase) (Prep-Phrase) (Prep-Phrase) (Manner)

Auxiliary → Tense (Modal) (*have* + *-en*) (*be* + *-ing*)

Predicate →
- Adjective
- Adverb

Noun Phrase → (Determiner) Noun

Determiner → Article

Lexical Insertion

Words are then inserted for the deep-structure elements. Here's an example of a deep structure where all the optional elements are used, as expressed in the abbreviations employed in the book:

Conj + **Art** + **N** + **Tense** + **Modal** + *have* + *-en* + *be* + *-ing* + **V** + **Art** + **N** + **Prep-Phrase** + **Prep-Phrase** + **Manner** + **Place** + **Time**

This sequence then becomes

And the boy **Tense** may *have* + *-en be* + *-ing* meet the girl of Italian extraction from Boston secretly on campus at noon

When some of the following transformational rules are applied to this generated sequence, the result can be "And the boy might have been meeting the girl of Italian extraction

from Boston secretly on campus at noon" or a phonemic representation of speech.

Transformational Rules

Three sets of examples are developed below, so as to show the necessary order of application of the fifteen **T** rules. These are (1–7), (9–11), and (12–15).

(1) **T-not:** adds *not* to certain verbal structures:

this boy **Tense** can eat at noon → this boy **Tense** can not eat at noon

(2) **T-infl:** gives number to a noun, and past or present tense to the predicate:

this boy **Tense** can not eat at noon → this boy **pl past** can not eat at noon

(3) **T-agr:** makes the determiner and *be* or modal or auxiliary or verb agree with the subject:

this boy **pl past** can not eat at noon → this **pl** boy **pl past** can not eat at noon

(4) **T-ques:** makes a declarative structure into a question:

this **pl** boy **pl past** can not eat at noon → **past** can not this **pl** boy **pl** eat at noon

(5) **T-adv:** replaces the adverbial prepositional phrase of time with *when* and permutes it to the front of the sentence; does the same with a "place" phrase, which becomes *where:*

past can not this **pl** boy **pl** eat at noon → when **past** can not this **pl** boy **pl** eat

(6) **T-contr:** contracts certain word-sequences:

when **past** can not this **pl** boy **pl** eat → when **past** can n't this **pl** boy **pl** eat

(7) **T-aux:** permutes the tense-marker and verbal:

when **past** can n't this **pl** boy **pl** eat → when can **past** n't this **pl** boy **pl** eat ("When couldn't these boys eat?")

(8) **T-to:** conjoins two or more verbal sequences by adding *to* while deleting surplus elements:

Jo **present** not like X; Jo **present** swim → Jo **present** not like to swim

(9) **T-ing:** conjoins two or more verbal sequences by adding *-ing* while deleting surplus elements:

the girl **present** not like X; the girl **present** swim → the girl **present** not like -ing swim

(10) **T-adj:** moves the adjective in front of its noun headword:

the girl **present** not like -ing swim; the girl **present** be good → the good girl **present** not like -ing swim

(11) **T-do:** adds *do* when **T-aux** can't operate:

the good girl **present** not like -ing swim → the good girl do **present** not like -ing swim ("The good girl does not like swimming")

(12) **T-pass:** changes active-voice structure to passive:

I **past** see Ed then → Ed **past** be -en see then by I [+ acc]

(13) **T-pro:** moves pronouns in certain distributions and/or adds needed case:

Ed **past** be -en see then by I → Ed **past** be -en see then by I [+acc]

(14) **T-del:** makes certain deletions:

Ed **past** be -en see then by I [+acc] → Ed **past** be -en see then

(15) **T-perm:** permutes certain kinds of units forward to certain positions:

Ed **past** be -en see then → then Ed **past** be -en see

These fifteen rules produce many structures. For example, if **T-aux** or **T-do** follows certain outputs, there are other results:

if applied after **T-ques**	(4) –	"Could not these boys eat at noon?"
T-to	(8) –	"Jo does not like to swim"
T-ing	(9) –	"The girl does not like swimming"

T-adj (10) – "The good girl does not like swimming"
T-pro (13) – "Ed was seen then by me"
T-del (14) – "Ed was seen then"
T-perm (15) – "Then Ed was seen"

Appendix C:
Glossary

acronym A word created from the initial parts of a sequence, as in *radar* from "radio detecting and ranging" or CARE from "Cooperative for American Remittance Everywhere."

Adjective (A) One of the two kinds of Predicates that can follow the copula *be,* as in "Jack is busy."

affix A prefix or suffix that cannot be used without combination with one or more other forms. Thus *dis-* and *-er* can be combined with a base form into *disclaimer. Claim* is a free form and so can be used without combination.

alphabetic Writing consisting of a set of letters or characters representing the individual vowels and consonants, or sounds, of a language.

alveolar A sound produced by the articulation of the apex of the tongue against the alveolar (or tooth) ridge, as in /t/.

aphaeresis The loss of one or more initial sounds of a word, as in *'bout* in the rapid-speech utterance of "I'm 'bout to go." This syllabic loss also illustrates aphesis, as contrasted with the example of *adder* from *nadder,* in which both syllables are retained.

apocope The loss of one or more terminal sounds of a word, as in the change of the former disyllabic *bite* to a monosyllable. Loosely, a clipping like *gym,* from *gymnasium,* might be called apocope.

argot Words composing part of the lexicon of certain groups —e.g., of the underworld. The words supposedly convey a special, private meaning to speakers familiar with the argot, as in *fuzz* to mean "police."

Article (Art) The most important kind of Determiner. It can be definite (*the*) or nondefinite (*a* or *an*).

Auxiliary (Aux) The part of the Predicate Phrase containing the Tense and possibly a Modal and/or (*have* + *-en*) and/or (*be* + *-ing*).

base form The form of a linguistic element before it is inflected and / or affixed. For example, *nerve* is a base, before it becomes something like *unnerve* or *unnerved.*

blend A word created from two or more existing elements, as in *chortle* from *chuckle* + *snort.*

borrowing A word transplanted from one language into another. *Allah,* originally borrowed from Arabic, is now an English word.

centum language A language in which the Indo-European palatal stops did not change to palatal or alveolar fricatives.

clipping The extreme reduction of the initial or terminal part of a word, as in *bus* from *omnibus,* or in *eco* from *economics.*

cognates Words similar in sounds and structure derived from some common ancestral word. English *nine* and German *neun* are cognates, derived from Indo-European **enewen.*

coining A new word invented or made up, like the American English *blurb.*

comparative linguistics A division of linguistics concerned with the genetic relationships among languages of a common origin, as in the comparative study of Italian, French, and Spanish.

competence, linguistic The speaker-hearer's total linguistic knowledge of his language, which accounts for all speech or writing produced in that language by the given speaker-hearer.

compound A combination of at least two free forms, to which one or more affixes may be joined. The compound *gentlemanly* consists of *gentle* + *man* + the suffix *-ly.* The process by which this word is formed is called *compounding.*

copula The word *be* that links the Noun Phrase serving as the subject to its Predicate, as in "He is friendly" or "He is here."

cuneiform Writing consisting of wedge-shaped characters.

declension The ordered list, or paradigm, of all the inflectional endings of a given noun, pronoun, or adjective base.

deep structure The basic elements of a sentence that provide the meaning of that sentence, as originally generated by deep-structure rules. For example, the deep structure *he* + **Tense** + *be* + *here* underlies and conveys the meaning of an eventual surface structure like "He was here."

degeneration The historical adding of a negative or moral judgment to a word. For example, Old English *cnapa* (serving boy) has become *knave* (rogue or dishonest person). Some "four-letter words" have degenerated. This is the reverse of elevation.

descriptive linguistics A division of linguistics concerned with the analysis of one or more languages at a particular stage in their development. It is the basic division because both historical

linguistics and comparative linguistics must begin with a description of some language at some stage in its history.

Determiner (D) The optional, initial part of a Noun Phrase, as in *the* in "The boy ate quickly."

dialect A regional or social variety of a language as used by one or more speakers of a language who share certain features of pronunciation, lexicon, and grammar. In the abstract, any dialect is as good as any other.

dialectology The study of the various dialects of a language, often resulting in a linguistic atlas (a set of maps recording the variations in syntax, lexicon, and phonology of a given region).

diphthong A combination of two vowels or glides that are produced so swiftly as to constitute one syllable, as in the pronunciation of *-i-* in *ride*.

elevation The historical removal of a negative or moral judgment from a word. Old English *cniht* often meant boy or servant, whereas its descendant *knight* describes someone honored by his sovereign.

embedding Placing a structure within and subordinate to another structure. For example, the sentence "He wanted the Chinese to intervene" can be embedded in "It is suspected," producing "That he wanted the Chinese to intervene is suspected."

euphemism An expression regarded as softer and more agreeable than the more direct one, as in *pass away* for *die*.

form classes The various grammatical classes into which the various words in a language fit. Some English form classes can take inflections, like nouns and verbs. Others, like prepositions and conjunctions, cannot.

free form A linguistic form that can be used without combination with one or more affixes, as in *boy* and *girl*.

fricative A sound produced when two opposed parts of the air passage do not quite touch. However, the narrowing causes turbulence or a rubbing sound because air comes through the narrowed passage under pressure. The English voiceless fricatives are /f, θ, s, š/.

generalization The widening of the meaning of a word. *Mill* once signified a place where grain was ground into meal. Today there are textile mills and steel mills, in which no grinding is done.

glide A sound produced when two opposed parts of the air passage are wider than they are for the articulation of a fricative,

but narrower than for the non-narrowing that produces vowels. The English glides are /y, h, w/.

grammar The total abstract apparatus and theory needed to explain the phenomenon of language in general and the data of all the particular languages, and to account for the mental activity required in the speaking and understanding of any particular language. (See **universal grammar** and **particular grammar.**)

Great Vowel Shift (GVS) The shift of certain English vowels beginning in Late Middle English and continuing into the nineteenth century. The GVS explains why many English spellings today do not represent the actual vowel sounds in the words, as in *bite* and *book.*

Grimm's law The systematic explanation by the fairy-tale collector Jakob Grimm of the shifts in certain Germanic consonants that did not occur in non-Germanic consonants. Thus English *tooth* is consonantly different from Greek *odontos,* although both words derive from Indo-European **edont-*.

hieroglyphic Writing consisting of characters pictorializing the thing represented by the word they stand for.

historical linguistics A division of linguistics concerned with the earlier stages of a given language, as in the study of Middle English to illuminate aspects of Present-Day English.

homonyms Words that sound alike but have different meanings, as in *bare-bear.*

hybridization The process by which a new word is formed from elements originally belonging to at least two different languages. Thus English *answer* and French *-able* were joined to create the hybrid *answerable.*

ideographic Writing consisting of characters pictorializing actions or abstractions represented by the word an individual character stands for.

idiolect An individual's unique total pattern of the pronunciation, lexicon, and grammar of his language. The pattern is different from anyone else's idiolectal use of that language. The differences among individual speakers are quite trivial and seldom impede communication.

inflections Modifications in the form of words, usually achieved by adding a suffix to the word base, to convey grammatical relationships like number, case, and gender. *Boys* is the plural, inflected form of *boy.*

inkhorn terms Certain polysyllabic words that Renaissance translators attempted to introduce into English from Greek and Latin. The words often described scholarly pursuits and attitudes.

intonation A composite term for the stress, pitch, and juncture of a language. Although broadly phonological, these three kinds of phonemes contrast with the sounds of the language, since we can't hear them but do hear the sounds.

labial A sound produced by the articulation of the lower lip against the upper lip, as in the labial nasal /m/ or the labial voiced stop /b/. When the lower lip articulates against the teeth, a labiodental like /v/ is produced.

language A method of communication that is simultaneously oral, symbolic, systematic, arbitrary, recurrent, social, purely human, noninstinctive, and adequate for man to make a linguistic response to any experience.

lexicon The list of all the words of a given language at a given time. No one speaker can ever know all the words.

loanword A somewhat specialized kind of borrowing, although all borrowings ultimately become loanwords if they remain in the language. Thus after French *Paris* was borrowed into English, the French pronunciation of *-ris* was naturalized to fit English sound values, causing Frenchmen to shudder when English speakers pronounce the name of their capital.

meaning (See **degeneration, elevation, generalization, specialization.**)

metathesis The inversion of sounds in a word, as in the development of *bird* from *brid.*

Modal One of a few optional words of the Auxiliary that immediately follows the Tense, as in *can, may, must, shall, will.* The Modal precedes the optional (*have* + *-en*) and / or (*be* + *-ing*). It can be defective, as in "He got married," which contrasts with other modal structures in which the verb is not inflected ("He should marry").

monophthong A single vowel sound, as in *lip,* contrasting with a diphthong like that in *ride.*

nasal A sound produced when the air goes up through the nasal cavity and out the nose, rather than through the mouth. The English nasals are /m, n, ŋ/.

nonstandard dialect A social variety of a language that is not accepted by the leaders of the speech community. This variety is

often spoken by the uneducated, underprivileged members of that community. It is entirely adequate for communication within the groups who speak it.

Noun (N) A word like *Jack, Syria, desk, courage, girl, rice. Jack* is a proper noun, and *Syria* is a place-name. Neither word can be preceded by a Determiner. *Desk* is a concrete noun; *courage* is a nonconcrete one. *Girl* is a count noun; *rice* is a noncount one. The latter four words are common nouns, as contrasted with *Jack* and *Syria*.

Noun Phrase (NP) The word or group of words employed as the subject of a sentence or as certain functions within a Predicate Phrase. *He* and *the race* are Noun Phrases in "He lost the race."

palatal A sound produced by the articulation of the front of the tongue (just behind the apex) against the palate, which is the immovable bone plate just behind the tooth ridge.

paradigm The ordered list of all the inflectional forms of a declension or a verb conjugation.

particular grammar The specific set of procedures and rules governing the actual data of a given language, making it different from all other languages. All languages, however, share the organizational principles of universal grammar.

performance, linguistic A particular corpus of speech or writing produced by a given human being. Because the corpus derives from that person's competence, an analysis of a particular performance is of little scholarly value by comparison with the projected analysis of his linguistic competence.

permutation Inversion or change of the lineal order within an ordered set of elements.

phoneme A unitary, or basic, sound feature in a language, usually written within slant lines. For example, the phonemes of *knight* are /naɪt/. There are also intonational phonemes.

phonological One of the three components of every language. Loosely, it can also be defined as the study of the sounds, intonational features, and their distributions in one or more languages. (See **aphaeresis, syncope, apocope, metathesis, clipping.**)

phonological rules (P rules) A set of rules belonging to particular grammar that govern the assignment of the particular phonetic interpretation to the word sequence of a given surface structure. **P** rules supply the proper pronunciation of that surface structure.

pitch The relative vibration frequency of a human voice that causes our ears to hear a given sound sequence as "higher" or "lower" than some other sequence. As men do not customarily speak in a monotone, a simple sequence like "I'm going home" will likely have a higher pitch on one of the words, primarily depending on the speaker's stress.

place-name A kind of proper noun naming a geographical locality.

Predicate Phrase (PP) The word or group of words containing the Auxiliary and a Verb Phrase.

preterit The past tense of a verb or of the copula *be*.

progressive constructions Certain English *-ing* verbal structures, as in "I am reading" and "He had been loafing."

protolanguage An unattested ancestral language from which one or more later languages are assumed to have developed.

reconstruction The scholarly process in linguistics by which known forms of a language are used to project unattested forms of that language at some earlier date from which the known forms later descended.

rules The kinds of rewrite instructions that provide for the change of a certain element or sequence into some other element or sequence. (See **deep structure, transformation, lexicon, phonology.**)

rune An alphabetic writing used by certain Celtic and Germanic peoples from about the third to the thirteenth centuries. Both the Celts and the Anglo-Saxons wrote in runes.

satem language A language in which the Indo-European palatal stops became palatal or alveolar fricatives.

semantic One of the three components of every language. It is a set of rules that assign a semantic interpretation to each syntactic description generated by the syntactic component. Native speakers of a particular language share the major lexical rules, which select the meaning of particular words in contexts.

specialization The narrowing of the meaning of a word. Old English *dēor* once signified an animal; today it specifies deer. This is the reverse of generalization.

Standard English The language of educated, cultivated members of the English speech community, as contrasted socially with the nonstandard speech of the underprivileged in that community.

stop A sound produced when two opposed parts of the air passage make contact, completely blocking the breath stream. Some English voiceless stops are /p, t, k/.

stress The accent or relative degree of loudness of the syllables in a sequence. In the word *away*, the second syllable receives the stress, contrasting with the unstressed *a-*.

surface structure The ultimate form of a sentence as used in actual communication, derived from the deep structure of the sentence by the application of transformational rules.

syncope The loss of one or more medial sounds of a word, as in the pronunciation of the word *vegetable* as /vɛjtəbəl/, in which the /ə/ after /j/ has been lost.

syntactic One of the three components of every language. It generates an infinite number of syntactic descriptions of sentences in a given language, each of which is composed of individual words. The descriptions are then acted upon by the semantic and phonological components. Thus the syntactic component is the basic one.

tonal language A language in which a given word has an intrinsic pitch or tone that dictates the meaning of the word. For example, a falling tone for Thai *kaw* means "nine"; a low tone for *kaw* gives the meaning "to be old." Yet the sounds of the two words are identical.

transformational rules (**T** rules) A set of rules belonging to particular grammar that change a deep structure to a surface structure. For example, **T-pass** changes the active-voice structure "Jack saw Bill" to the passive "Bill was seen by Jack."

tree A graphic presentation of the deep structure of a given sentence. A tree diagram shows the exact derivation of each constituent of the sentence.

universal grammar The principles determining the form of the grammar of any language. Shared by all men, the principles select the specific form of the language that a baby learns from his mother.

usage A term for describing the specific phonological, lexical, and grammatical variations within a language. When variations are not acceptable socially, they are considered nonstandard and help stigmatize the speaker.

velar A sound produced by the articulation of the dorsum (or back) of the tongue against the velum (the soft palate, just be-

hind the palate), as in /g/. When the velum is lowered, the air passage is closed and a nasal is produced, as in the velar nasal /ŋ/.

verb, dental The kind of English verb that adds a dental /t, d/ to form its inflections, as in *learn-learned.*

verb, vocalic The kind of English verb that does not add a dental /t, d/ to form its inflections. This kind usually has an internal vowel change, as in *hide-hidden.*

Verb Phrase (VP) The part of the Predicate Phrase containing either the copula (plus one or more other words) or a verb (which may be followed by other words). In "I might be happy" and "I went home later," the last three words of each sentence constitute the Verb Phrase.

Verner's law Karl Verner's explanation of the exceptions in Grimm's law. That is, Germanic voiceless fricatives occurring between vowels became voiced unless the preceding vowel was stressed. The English word is thereby *seven* rather than something like **sefen,* as Grimm's law would otherwise predict.

vowel A sound produced when the air passage is not narrowed, as contrasted with the three degrees of narrowing that produce stops, fricatives, and glides. Some English vowels are /i, e, æ/.

writing The visual representation of language. Writing should not be confused with language, which existed long before man began to represent it through marks on stone or paper.

zero form A language form that does not add the inflectional element that other forms like it do add. For example, *sheep* is pluralized by adding **-ø,** unlike *boys.* A "zero" verb like *cut* is comparably fitted into the paradigm of verbs like *eaten* or *learned* by the addition of **-ø.**

Appendix D:

How Did Language Begin?

By Mario Pei

There are many theories, all unproved, as to how language began. Most picturesque among them is the "bow-wow" hypothesis to the effect that men began to speak by imitating the natural sounds they heard, or thought they heard, around them.

The barking of a dog would strike the ear of the leader of a small band of primitive humans. It would sound to him like "bow-wow," and as he tried to imitate it he would convey to the others, by pointing to the dog and repeating "bow-wow," that the creature that made that particular sound should henceforth be referred to as "bow-wow." Too simple? Yet consider how often children spontaneously fasten upon some utterance produced by one of their number and use it to designate him, pointing to him in derision as they do it.

The scientific name for this process is onomatopoeia, or "name-making." Less scientific but easier to pronounce and spell is "echoic word." You echo what you hear. If the fall of a big tree in the forest sounds to you like "crash," that is what you use to designate that type of sound. The noise produced by a bee may sound like "hum" or "buzz." Words like "click," "wham," "bang" all seem to be of echoic origin.

Different breeds of dogs bark in different ways, or the same sound may be differently interpreted and echoed by various human beings. This would account for "bow-wow," "woof-woof," "yip-yip," and "arf-arf" all appearing in the same language. If you have many languages, the differences may be far greater.

English has perhaps more echoic words than any other civilized tongue. Is this because we are more primitive and elemental? Or because our language runs more to monosyllables and avoids endings? Or because we make greater use of comic strips, where the picture largely tells the story but sound effects have to be graphically portrayed?

Yet the echoic word does not have to be monosyllabic, or even repetitive, especially if the sound it portrays strikes the ear as composite. Among early echoic words that are not repeti-

tive are ancient Sanskrit *chish-chá,* denoting the "whiz" of an arrow in flight followed by the sound of its impact, and *kikirá,* to denote a palpitating sound like our "pitter-patter" of the heart.

Both Greek and Latin had plenty of echoic words, a few of which, like the Latin *murmur,* have been passed on to us. The Roman grammarian Quintilian describes Latin as poor in such sounds, but the facts don't seem to bear him out. Both Latin and Greek, however, are languages given to endings that denote specific parts of speech, so that many of what must have been originally one-syllable echoic words appear as two- or three-syllable nouns, verbs, and adjectives. For example, one of the Latin words meaning "to bark" is *baubari,* where the *-ari* ending is merely an infinitive suffix; the Roman "bow-wow" was evidently *bau* or *bau-bau,* though it does not appear by itself in the literary records that have come down to us. "To neigh" in Latin is *hinnire;* here again the Roman speakers must have used the root *hinn-* when they wished merely to imitate the neighing of a horse.

The Greeks did better in leaving us records of bare echoic roots. To them the croak of a frog was *koax,* but that of a raven was *kro.* The grunt of a pig (our "oink") was *gru,* but the squeal of a small pig was *koi.* The bleating of a sheep (our "baa") was to them *beh,* and it is a joke among linguists that as the Greek sounds changed during the course of centuries, the same written word came to be pronounced *vee,* which does not at all sound like a sheep, thus proving that animal language remains the same though human speech changes. Even the speakers of Sanskrit have left us their idea of "splash" as represented by *bal, p-hal,* or *p-hat.*

It is fairly evident by this time that different groups hear the same sound in different fashions. What to us is the "smack" of a kiss is to Spanish speakers *muá.* The "snip-snip" of a pair of scissors sounds like *krits-krits* to the modern Greeks, *su-su* to the Chinese, *cri-cri* to the Italians, *riqui-riqui* to the Spaniards, *terre-terre* to the Portuguese. Our "bang" of a pistol may come out as *bam, pam, pan,* even *tau.* The "crash" of a tray of plates and cups falling to the floor is *kling* to the Danes, *krats* to the Finns, *chir-churr* to the Hungarians, *hua-la-la* to the Chinese, while the comic strip "wham" of someone sitting down suddenly and very hard is *pan* in French, *cataplúm* in Spanish, *catrapuz-bum* in Portuguese, *patatrac* or *patapunf* in Italian. Even the

ringing of a phone, for which we have no echoic word (unless we accept "ring" itself, or "tinkle") may come out as *dringh* in Greek, *drin* in Italian, *kili* in Finnish, *tlim* in Portuguese.

But along with this diversity there are widespread resemblances. The Latin *murmur,* which has come down to most Western languages, appears in very similar form in Armenian, Lithuanian, Greek, even Sanskrit, where *marmarah* means "noisy." Our "gurgle" is *glut-glut* in Latin, *glu-glu* in Italian. Even our "slap" has a close equivalent in Latin—*stlop(pus)*—from which, interestingly, Italian and Spanish derive their words for "shotgun": *schioppo, escopeta.* The "ho-ho-ho" of the Jolly Green Giant is *khokhot* in Russian, *kakhat* in Sanskrit.

Animals have proved their superiority over humans by achieving an international language within their respective species. There is no convincing evidence that the braying of a donkey, the cackling of a hen, the quacking of a duck, the mooing of a cow is any different in China or the Soviet Union from what it is in the United States. But there can be vast differences in human reception and rendition.

Most standardized, perhaps, among animal sounds is the cow's "moo" (it may, of course, be spelled *mu,* and French has an interesting variation—*meuh*). Second in standardization is the cat's "meow" (here the spelling runs from Italian *miao* and French *miaou* to German and Rumanian *miau*). But Japanese has *nya-nya,* and Arabic has a double form, *nau-nau* for ordinary meowing, but *maw-oom* for the cat's voice in the mating season. This unanimity does not extend to the cat's "purr," which Spanish imitates as *arro-arro-arro* and French as *ronron.*

The bleating of the sheep gives rise to two renditions, one with *b-,* the other with *m-*. Greek, Latin, English, Spanish, Italian, Russian, and Vietnamese favor the first (Russian has *bya-bya,* Vietnamese has *be-be*); German, Rumanian, Chinese, and Japanese favor the *m-* sound (*ma-ma* or *me-me*). French and Arabic use both.

The crowing of a rooster and the cackling of a hen have only one element that is internationally common, an initial *k-* sound, often repeated elsewhere in the word; everything else is different. Corresponding to English "cock-a-doodle-doo" is the French *cocorico,* Spanish *quiquiriquí,* Italian *chicchirichì,* German *kikiriki,* Rumanian *cucurigu,* Russian *kukareku,* Arabic *ko-ko* or *qee-qee* (*q* in Arabic is a very guttural *k,* pronounced as far back

in the throat as you can get it), Japanese *kokekkoko,* and Vietnamese *cuc-cu.*

For the hen's cackle, English does not have a real echoic word. (We make up for it by imitating a turkey's "gobble-gobble," which few other languages bother about.) Here French has *cot-cot,* Rumanian has *cotcodac,* Italian has *coc-cote,* Arabic has *qa-qa,* Chinese has *ko-ko-ko,* Japanese has *kukku,* Vietnamese has *cuc-tac.* Even Latin has *co-co* or, with greater repetition, *co-co-co-co.*

We lack an imitation of the horse's neigh, which Italian portrays very graphically with *ih-ih-ih-ih-ih.* Rumanian has *hi-hi-hi,* Arabic *hem-hem,* Japanese *hi-hin,* and Vietnamese *hi.* On the other hand, we have a donkey's "hee-haw." Here French has *hi-han,* Italian and Chinese share *i-o,* German and Russian share *i-a,* Rumanian has *i-hau,* Arabic uses *ham-ham* or *hee-hee.*

There is considerable internationality in the duck's "quack-quack." French uses *couac-couac* or *coin-coin* (the latter sounds like "kwan-kwan"). Spanish has *cuac-cuac,* Italian has *qua-qua,* German shares *quack-quack* with us, Russian has *kva-kva,* Vietnamese has *cac-cac.* But Japanese begins to diverge with *ga-ga;* Rumanian carries it on to *mac-mac,* Arabic to *bat-bat,* Mandarin Chinese to *ya-ya,* and South China's Cantonese to *ap-ap.*

No language seems to have a real lion sound, though many use our own "grrr" for any kind of growl or roar. Arabic, because of some contact with lions in their native habitat, uses *u,* which is a prolonged *oo.* Vietnamese has no lion sound, but with plenty of tigers in the land the tiger's roar is imitated as *ham-hu* or *gam-gu.*

Animals have imitation words where they are well known to

the people. The Nutka Indians of Alaska imitate the sound of a whale as *'hw* (constriction of the throat, strongly uttered *h,* sound of *w*). Eskimo tribes prefer *peu-wu.* Closest to our "oink" for a pig's grunt is French *oui-oui,* which means that in France the pig is forever saying yes. Quite remote are Russian *khru-khru* and Rumanian *guits-guits.* The "peep" and "chirp" of chicks and small birds are imitated as *pio-pio* in Italian, *piu-piu* in Rumanian, *cui-cui* in French.

The dog, who contributed one of his names to the "bow-wow theory," has the most far-reaching divergences, due perhaps to different breeds but also to the fact that it was probably the first animal domesticated by man (the dog was the sole domestic animal of the North American Indians). Corresponding to our assorted "bow-wow," "woof-woof," "yip-yip," and "arf-arf" we find French *oua-oua* (pronounced "wah-wah"), Italian *bu-bu,* Spanish *guau-guau* or *jau-jau* (pronounced "how-how"), Rumanian *ham-ham* (with *a* of "father"), German *hau-hau* or *wau-wau,* Russian *vas-vas* or *vaf-vaf,* Arabic *'au-'au* (constrict the throat at the start), Vietnamese *gau-gau,* Turkish *hov-hov,* Chinese *wang-wang,* and Japanese *wan-wan.* Even ancient Sanskrit had *bhuk-bhuk.*

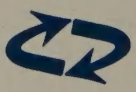
Kindred to echoic words are interjections, those exclamatory sounds which we use to express pain, pleasure, surprise, disgust, annoyance, joy, sorrow, or simply to call someone's attention. These come closest to the natural, spontaneous sounds made by animals. Some are surprisingly international, others surprisingly different. As a sample of the first, we find in the ancient Sanskrit of the Vedas all of these familiar forms: *a, ha, haha, ahaha, he, hai.* But some can undergo amazing changes in meaning in the course of time. The Latin *bua* is described as "a sound made by infants to denote what they are drinking." The same word is used by children today in Italy, but it means "to hurt," "to ache," "to have a sore spot," or "to be ill." To call someone's attention at a distance we generally use "Hey!" The ancient Greeks used *eia,* the Romans *eho,* the modern Italians, particularly in Rome, *ao.*

I once asked a girl who was completely trilingual, having been brought up in New York, Paris, and Havana in equal measure, whether she had ever gotten her three languages mixed up. She thought and thought, then brightened up. "Yes! One time, on Varadero Beach in Cuba, someone stuck me with a pin,

and I yelled 'Ouch!' instead of '*Ay!*'" On a French beach, she should have yelled "*Aie!*" or "*Ouille!*" In Italy it would have been *"Aio!"* in Hungary *"Jaj!"* (pronounced "yoy"), in Finland *"Boi!"* in Japan *"Itai!"*

We indicate disgust by using "blah" or "aak" ("phooey" is a recent borrowing from Yiddish). In Spanish it's *huy* or *uf,* in Italian *uffa,* in French *fi, pfutt,* or *zut,* in German *pfui.* In a good many of these, there seems to be some imitation of the sound of spitting. But Latin used *pro.*

Sorrow used to be indicated by "Alas!" but this, save for the initial *a,* is not an echoic word. It comes from Old French *ah, las!* —"Oh, weary (me)!" Our real international exclamation of sorrow is, historically, the "woe" of "Woe is me!" This has wide range, from Latin *vae* of *vae victis* ("Woe unto the conquered!") to Welsh *gwae,* Gothic *wai,* Armenian *vai,* Old Persian *avoi.* But Latin had, side by side with *vae,* also *eu, eheu,* and *ei* (the last often combined with *mihi*—"Woe is me"). Greek used *pheu,* which is perhaps linked to the German *pfui.* This *pheu* has come down into the college yell of Italian university students: "*Pheu, pheu; baru!*"—"Alas, alas; way down in the dumps!" But the old sorrowful connotation is altogether lost, and the mournful Greek words have been turned into a happy rallying cry. Thus do the centuries work their ways on language.



domains
in language and composit

harcourt brace jovanovic

0-15-312356-