CHAPTER IV

The Backgrounds of English

LANGUAGE tends to change for a variety of reasons. Moreover, this manifold tendency is operating all the time in all languages, though its operations are not perceptible save to a few persons who for one reason or another concern themselves with such matters and who may note changes in word usage and in the pronunciation of individual words (like *ration*, which used to rime with *passion*, but has comparatively recently acquired a pronunciation riming with *nation*) in the course of a lifetime. But such changes as these are really quite minor in relation to the changes in sounds and grammar which take place in the course of centuries—changes which would make the English of Geoffrey Chaucer, if we were happily able to resurrect him, seem very different from our own speech. That of Alfred the Great, who lived about as long before Chaucer as we live after him, would seem quite like a foreign language to all who had not made a special study of Old English.

THE PUZZLE OF LINGUISTIC CHANGE

Passage of time and geographical separation are but contributory factors and not direct causes of linguistic change. Thus, the English language as Americans speak it today has become differentiated from present Standard British English in certain respects in the relatively short time that has elapsed since its speakers became separated, though never really isolated, from the mother country. The English language of the seventeenth century suffered no sea-changes when it was brought across the Atlantic: the first Englishmen to settle in America spoke exactly as they had spoken at home. Nevertheless, changes have certainly come to pass. Those which have occurred in British English are, as a matter of fact, considerably more far-reaching and more fundamental than anything which has happened independently in Ameri-

can English—the treatment of *a* in words of the *staff, glass,* and *path* type, the treatment of *r* preceding a consonant or in final position, and the treatment of the penultimate syllables of polysyllabic words in *-ary, -ery,* and *-ory,* for instance. American English preserves the older British treatment; Standard British English has changed in all these respects since the seventeenth century.

We can say precisely what has happened in any sort of linguistic change, and we can also as a rule say where a particular change occurred and assign to it an approximate chronology. But we can supply a reason only for those changes due to analogy, which is a mental process, even though it need not be conscious. We can describe what happened as "rounding" when Old English *ā* (in *stān, hām, bān,* and so forth) came invariably to be pronounced [ɔ:] south of the Humber, and we can say, on the basis of spellings with *o,* approximately when this began to happen (late eleventh century). But *why* it happened, or, for that matter, why this later sound was subsequently raised to [o:], as in Modern English *stone, home, bone,* and so forth, we cannot really explain. We can only declare that certain tendencies toward changes of various sorts have prevailed at certain times and in certain places, and that these changes have accomplished themselves not haphazardly, but with the greatest regularity. We can usually map the exact course of such changes, and we can describe them systematically in phonological terms.

LANGUAGE FAMILIES

In the discussion of so-called linguistic families which follows, we must bear in mind that a language is not born, nor does it put out branches like a tree—nor, for that matter, does it die except when every single one of its speakers dies, as has happened to Etruscan, Gothic, Cornish, and a good many other languages. When we speak of Latin as a dead language, we are referring to a highly artificial literary language; but spoken Latin still lives in various developments in Italian, French, Spanish, and the other Romance languages.

Hence the terms *family, ancestor, parent,* and other genealogical expressions when applied to languages must be regarded as no more than metaphors. Languages are developments of older languages rather than descendants in the sense in which people are descendants of their forefathers. Thus, Italian and Spanish are different developments of an earlier, more unified language, Latin. Latin in turn is one of a number of developments, which include Oscan and Umbrian, of a

still earlier language called Italic. Italic in its turn is a development of Indo-European. Whether or not Indo-European has affinities with other languages spoken in prehistoric times and is hence a development of an even earlier language, no one is prepared to say with certainty; for, as we have seen, we are quite in the dark about how it all began.

Older scholars—and they were to some extent theorists—classified languages as monosyllabic, agglutinative, incorporative, and inflective, these being exemplified respectively by Chinese, Turkish, Eskimo, and Latin. The monosyllabic languages were supposed to represent the most primitive type—a notion which doubtless grew out of investigations into languages of our own Indo-European group, with their large number of monosyllabic roots. But even the earliest (middle of second millennium B.C.) records of Chinese, a monosyllabic language in its modern form, represent not a primitive but actually a late stage in linguistic development. It obviously cannot be inferred from such evidence as this that our prehistoric ancestors prattled in words of one syllable each.

The older scholars also observed, quite correctly, that in certain languages, such as Turkish and Hungarian, words were made up of parts "stuck together," as it were; hence the term *agglutinative*. In such languages the suffixal elements are usually whole syllables having very definite meanings. The inflectional suffixes of the Indo-European languages were supposed likewise once to have been independent words; hence, some believed that the inflective languages had grown out of the agglutinative. Little was known of what were called incorporative languages, in which all sentence elements are combined into a single word; the elements have no independent existence, but can appear only as infixes.

The trouble with such a classification is that, though apparently objective, it is not really so, but is instead based on the out-of-date theory that early man spoke in monosyllables. Furthermore, the difference between agglutinative and inflective was not well defined, and there was considerable overlapping. Nevertheless, the terms are useful and widely used in the description of specific languages or even groups of languages.

A much more satisfactory and more objective classification of languages can be made on the basis of such correspondences of sound and structure as indicate relationship through common origin. Perhaps the greatest contribution of nineteenth-century linguistic scholars was

the painstaking investigation of these correspondences, many of which had been noted long before.

Such investigation indicated unmistakably that practically all of the languages of Europe (and hence of the Americas and other parts of the world colonized by Europeans) and some of Asia have in common certain characteristics of sound and structure and to some extent a stock of words which make it perfectly obvious that they have all developed out of a single language spoken in prehistoric times. This earlier language is usually called Indo-European.[1] What it was called by those who spoke it we have no way of knowing, nor do we know what they called themselves. We shall here follow the usual practice of referring to them as the Indo-Europeans, but it must always be borne in mind that the term has no racial connotations; it refers only to a group of people who lived in a relatively small area in early times and who spoke a more or less unified language out of which many languages have developed in the course of thousands of years. These languages are spoken today by approximately half of the world's population.

THE NON-INDO-EUROPEAN LANGUAGES

Before proceeding to a more detailed discussion of the Indo-European group, we may perhaps best delimit it by briefly noting those languages and groups of languages which are *not* Indo-European. Two important groups have names which reflect the Biblical attempt to derive all the races of men from the three sons of Noah: the Semitic (from the Latin form of the name of the eldest son, more correctly called Shem in English) and the Hamitic. The term *Japhetic*, once used for Indo-European, has happily long been obsolete. On the basis of many phonological and morphological features which they share, Semitic and Hamitic are thought by many scholars to be related through a hypothetical common ancestor, Hamito-Semitic; there are

[1] *Indo-Germanic* is not now much used except by German scholars. Its coinage was not due to German patriotism; it was intended to do no more than indicate what were thought to be the easternmost and westernmost limits of the geographical distribution of the languages recognized as belonging to the group. Another term, *Aryan*, has been used synonymously. Originally this term referred only to the Asiatic languages of the group. This is still the reference which it has in learned use, where its occurrence is now somewhat rare, *Indo-Iranian* and *Indo-Persian* being the preferred terms.

also those who believe in an ultimate relationship, impossible to prove, between Semitic and Indo-European.

The Semitic group includes the following languages: (Eastern) Akkadian, called Assyrian in the periods of the oldest texts, and later Babylonian; (Western) Hebrew, Aramaic[2] (the native speech of Jesus Christ), Phoenician, and Moabitic; (Southern) Arabic and Ethiopic. Of these, only Arabic is spoken by large numbers of people over a widespread area. Hebrew has comparatively recently been revived in Israel, to some extent for nationalistic reasons.[3] Ethiopic survives mainly in Geez, a Christian liturgical and learned language of Ethiopia, and in Amharic, which is used in state documents in that country. It is interesting to note that two of the world's most important religious documents are written in Semitic languages—the Old Testament in Hebrew (with large portions of the books of Ezra and Daniel in Aramaic) and the Koran in Arabic.

To the Hamitic group belong Egyptian (called Coptic after the close of the third century of the Christian Era), the Berber dialects of North Africa, and various Cushitic[4] dialects spoken along the upper Nile. Coptic is used in the liturgy of the Coptic Christian Church in Egypt, much as Geez is used in the Ethiopian Church and Latin in the Roman Catholic Church, but is not spoken elsewhere. Arabic became the national language of Egypt in the course of the sixteenth century.

Semitic is thus essentially Asiatic, and Hamitic North African. Hamitic is in no way related to any of the languages spoken by Negroes in central and southern Africa, the vast region south of the Sahara. These languages are usually classified into three main groups:

[2] Formerly—and incorrectly—called Chaldean, Chaldaic, or Chaldee. Though he should have known better, the foundations of modern linguistic science having already been laid in his day, Noah Webster thought that "Chaldee," which he believed to be the language of pre-polyglot Babel, was the ancestor of all languages. In his *American Dictionary of the English Language* (1828) he proposed a good many "Chaldee" etymologies which later and better-informed editors have quietly consigned to the wastepaper basket.

[3] Hebrew is not of course to be confused with Yiddish (that is, Jüdisch), a German dialect with many words of Hebrew origin which has come into being since the fourteenth century and has become a sort of international language of the Jews. Newspapers printed in Yiddish use Hebrew characters.

[4] Cush was a son of Ham.

Sudanese, extending to the equator, a large and highly diversified group of languages whose relationships to one another are difficult and in some cases impossible to establish; Bantu, extending from the equator to the extreme south, a large and well-defined group of related languages; and Hottentot and Bushman, remotely related languages spoken by small groups of quite primitive people—though the languages are in no sense primitive—in the extreme southwestern part of Africa. Hottentot and Bushman have no relationship to the other Negro groups, nor is it demonstrable that the Sudanese and the Bantu groups are in any way connected with each other.

Languages belonging to the Dravidian group were once spoken throughout India, where the earlier linguistic situation was radically affected by the Aryan invasion. These are the aboriginal languages of India. They are now spoken mainly in southern India.

The Indo-Chinese group includes Chinese proper and the languages of Tibet and Indochina. Japanese is totally unrelated, though it has borrowed the Chinese written characters and many Chinese words. Attempts to relate Korean to either Chinese or Japanese have not been successful. Ainu, the language of the aborigines of Japan, is totally unrelated to any other language of which we have any knowledge; it is now spoken by no more than a handful of people.

A striking characteristic of the Malay-Polynesian languages is their wide geographical distribution in the islands of the Indian and the Pacific oceans, stretching from Madagascar to Easter Island. The more or less moribund Australian native languages, spoken by only a few Australian blacks nowadays, have absolutely no connection with Malay-Polynesian, nor have the more than a hundred Papuan languages spoken in New Guinea and neighboring islands.

The American Indian languages constitute a geographic rather than a linguistic grouping, comprising many languages showing very little relationship, if any, to one another. It has been estimated[6] that at the time of Columbus' discovery only about a million and a half Indians occupied the huge area north of Mexico, with about forty million more in Mexico and Central America, the Antilles, and South America. A very important and widespread group of American Indian languages is known as the Uto-Aztecan, which includes Nahuatl, the language spoken by the Aztecs, and various closely related dialects.

[6] By P. Rivet, cited by Willem L. Graff, *Language and Languages* (New York, 1932), p. 427.

Aleut and Eskimo, which are very similar to each other, are spoken in the Aleutians and all along the extreme northern coast of America and north to Greenland. The isolation of the various groups, small in number to begin with and spread over so large a territory, may to some extent account for the great diversity of American Indian tongues.

Basque, a very intricate language spoken in many dialects by no more than half a million people living in the region of the Pyrenees, has always been something of a popular linguistic mystery. It now seems fairly certain, on the basis of coins and scanty inscriptions of the ancient Iberians, that Basque is related to the almost completely lost language of those people who once inhabited the Iberian peninsula and in Neolithic times were spread over an even larger part of Europe. Efforts to relate it to Etruscan, a language of which we know very little, to the non-Indo-European languages spoken in the Caucasus Mountains (not mentioned elsewhere here), and to the Hamitic languages have not been successful.

An important group of non-Indo-European languages spoken in Europe, as well as in parts of Asia, is the Ural-Altaic, which falls into two subgroups: the Ural, or Finno-Ugric, which includes Finnish, Estonian, Livonian, Lappish, and Hungarian, among others of less importance; and the very remotely related Altaic—though there are those who deny any such connection. Altaic includes various varieties of Turkish, such as Ottoman Turkish (Osmanli) and that spoken in Turkestan and in the Azerbaijan Soviet Socialist Republic, as well as Mongolian and Manchu.

The foregoing is by no means a complete survey of non-Indo-European languages. We have merely mentioned some of the most important groups and individual languages, along with some which are of little significance as far as the numbers or the present importance of their speakers are concerned, but which are nevertheless interesting for one reason or another. Louis H. Gray lists twenty-six linguistic groups and two isolated languages (spoken respectively in China and India by small groups of people), and comes up with a total of 2796 languages, of which 132 are Indo-European.[6] His figure coincides with that arrived at by the French Academy. But Gray rightly had no faith in such a count, for, as he points out, it is often impossible to reach agreement as to what constitutes a language: the line demarcating dialect and language is difficult to draw, and linguists do not always agree on where it should be drawn. Furthermore, depending largely upon one's

[6] *Foundations of Language* (New York, 1939), p. 418.

point of view, Old English, Middle English, and Modern English might be regarded as one, two (on the basis that the transition from Middle English to Modern English is somewhat less well defined than that from Old English to Middle English), or three. And there are yet further difficulties pointed out by Gray, who concludes that between 2500 and 3500 might be given as an estimate, but admits that such an estimate is "so rough as to be practically worthless."

EARLY STUDIES OF THE INDO-EUROPEAN GROUP

The concept of an Indo-European group of languages—and subsequently of other groups as well, for the work of the early Indo-Europeanists gave impetus to the study of non-Indo-European languages also—may be said to have grown out of British rule in India. It was this which was responsible for a wider knowledge of Sanskrit in Europe—to all intents and purposes a third ancient language with which to compare Latin and Greek. Latin had previously been supposed to be a degenerate form of Greek. Such an explanation of the correspondences between Sanskrit and the two hitherto-known ancient languages would obviously not do, for India was completely outside the sphere of Greco-Latin civilization.

For this new concept a remarkably versatile man, a veritable eighteenth-century "admirable Crichton," was largely responsible. A former member of Dr. Johnson's brilliant circle, Sir William Jones was at the age of thirty-seven judge of the supreme court of judicature at Fort William (Calcutta) in Bengal after a brilliant career as Orientalist, student of many languages, poet, classicist, jurisconsult, and public official in England, where he had withdrawn as a parliamentary candidate for Oxford University just before election day because of his sympathetic view of the American cause in the War of Independence and his opposition to the slave trade.

Shortly after his arrival in India, Jones founded the Bengal Asiatic Society. In a paper read before that group in 1786, he declared that Sanskrit bore to Greek and Latin "a stronger affinity . . . than could possibly have been produced by accident; so strong, indeed, that no philologer could examine them all three without believing them to have sprung from some common source, which, perhaps, no longer exists," going on to say that "there is a similar reason for that both the Gothick [that is, the Germanic] and the Celtick . . . the same origin with the Sanscrit."

Before Jones's time a good deal was known of many languages other

than the classical ones, and some attempts at classification had been made. A few Europeans, mostly missionaries, had even learned something of Sanskrit, and some of them had noted "affinities," or correspondences, with European languages. As early as the sixteenth century an Italian merchant in India pointed out in a letter the striking similarity of Sanskrit *deva-* 'God' to Italian *dio*; of *sarpa* 'snake' to his *serpe*; and of the numerals *sapta* 'seven,' *aštau* 'eight,' and *nava* 'nine' to his native *sette, otto,* and *nove.*[7] Had he been able to hear more of the ancient language known to and used by Hindu scholars of his day, he would have noted other such striking correspondences: in the numerical system alone, for example, *dvau/due* 'two,' *trayas/tre* 'three,' *čatvār-/quattro* 'four,' and *daśa/dieci* 'ten.'[8]

But Jones's clear statement, showing full realization of the relationship of Sanskrit to the principal European languages through a common ancestor, came just when the time was ripe for it, and may thus be said to have been the starting point for modern comparative linguistics. It only remained for more able, if less versatile, men to work out the details necessary to establish beyond any doubt the essential unity of that great group of related languages which we call the Indo-European family.

The first man to present a systematic comparison of a number of these languages was a German scholar with the unforgettable name of Franz Bopp. In 1816 Bopp published a brilliant study of the verbal endings of Sanskrit, Persian, Greek, Latin, and Germanic, and, in installments at intervals from 1833 to 1852, a huge comparative grammar in which he added Old Slavic and Lithuanian to the languages just named. He was later to add Armenian, Albanian, and Celtic. It was Bopp's work which got comparative Indo-European grammar on its feet.

But even before the publication of Bopp's first work the young Danish scholar Rasmus Rask had written a prize-winning essay on the origin of Old Norse in which he recognized the relationship of the Germanic, Hellenic, Italic, and Baltic groups and expressed the belief that Indo-Iranian might also be related. By the time his essay was published, in 1818, he had perceived that Armenian was Indo-European.

[7] Cited by Paul Thieme in "The Indo-European Language," *Scientific American,* CXLIX, 4 (October 1958), 63–74—a fascinating article intended for the general educated reader.

[8] The medial consonant of *daśa* was a palatal fricative approximating the final sound of German *ich.*

It was not until somewhat later that he admitted Albanian to the family.

Rask clearly perceived that sound shifts were regular, not sporadic, and recognized the Germanic sound shift later to be more expertly codified by Jacob Grimm in the second edition of his *Deutsche Grammatik* (1822).[9] This sound shift, to be discussed in some detail later, was in time to be associated with Grimm's name as Grimm's Law.[10] There is, however, some justice in the statement of Otto Jespersen that "if any man is to give his name to this law, a better name would be 'Rask's Law,' for all these transitions . . . are enumerated in Rask's *Undersøgelse* [the essay of 1818] . which Grimm knew before he wrote a single word about the sound shift."[11] In any case, the relationship of the Germanic languages, including English, to the other members of the group was now made perfectly clear.

Comparative and historical linguistics thus began with the study of the Indo-European family. Similar study of other languages and groups of languages was to come later, and as a result of the principles educed by the nineteenth-century Indo-Europeanists. Bopp, Rask, and Grimm had laid solid foundations for such studies, among them identifying the Germanic, Balto-Slavic, Celtic, Italic, Hellenic, and Indo-Iranian groups of languages as subgroups of Indo-European, along with the individual languages Armenian and Albanian. There followed in the course of the nineteenth century a series of brilliant and exciting studies, mostly by German scholars, which are still the basis of modern linguistic study. The story is best told by Holger Pedersen in his *Linguistic Science in the Nineteenth Century*.[12] The discovery and deciphering of writings in Tocharian and Hittite, found respectively in East Turkestan and Asia Minor, and the identification of the

[9] By *deutsch* Grimm meant 'Germanic,' not merely 'German.'

[10] It is an amusing irony that this great scholar, who never married, should be best known as the "author," with his brother Wilhelm, of a beloved nursery classic—the famous *Grimm's Fairy Tales*. The brothers Grimm, both of them fascinated by folklore, in the scientific study of which they were pioneers, were the collectors of the stories which have delighted generations of children despite the disapproval of modern child psychologists.

[11] *Language: Its Nature, Development and Origin* (New York, 1922), p. 43.

[12] Translated from the Danish by John Webster Spargo (Cambridge, Mass., 1931).

first of these languages as Indo-European in origin and the second as either Indo-European in origin or a language of common origin with Indo-European[18] are achievements of the twentieth century. They are also discussed by Pedersen in the work cited.

INFLECTION IN THE INDO-EUROPEAN LANGUAGES

All the Indo-European languages are inflective—that is, all are characterized by a grammatical system based on modifications in the form of words, by means of endings and vowel changes,[14] to indicate such grammatical functions as case, number, tense, person, mood, and the like. The older inflectional system is very imperfectly represented in most modern languages: English, French, and Spanish, for instance, have lost much of the inflectional complexity which was once characteristic of these languages; German retains considerably more, with its various forms of the noun and the article and its so-called strong adjective declension. Sanskrit is notable for the remarkably clear picture it gives us of the older Indo-European inflectional system; it retains much that has been lost or changed in the other Indo-European languages, so that its forms show us, even better than Greek or Latin can do, what the system of Indo-European must have been.

Traces of this inflectional system which survive in varying degrees in other related languages led the early Indo-Europeanists to the discovery of the relationship of languages as widely separated geographically as Icelandic and Sanskrit. Once one understands and makes allowances for the regularly occurring sound changes, the relationship of the personal endings of the verb in the various Indo-European languages becomes perfectly clear. For example, the present indicative of the Sanskrit verb corresponding to English *to bear* runs as follows:

SINGULAR
1 bharā-mi 'I bear'
2 bhara-si 'thou bearest'
3 bhara-ti 'he beareth'

[18] Those who hold to the second of these alternatives, notably the late American Hittitologist, E. H. Sturtevant, hypothesize a parent language called Indo-Hittite, from which both Indo-European and Hittite stem.

[14] As in Modern English *boy–boys; who–whom–whose; walk–walks–walked–walking; man–man's–men–men's; sing–sings–sang–singing.*

	PLURAL
1	bharā-mas 'we bear'
2	bhara-tha 'you bear'
3	bhara-nti 'they bear'

The only irregularity here is the occurrence of *-mi* in the first person singular, as against *-ō* in the Greek and Latin forms to be cited immediately below. It was a peculiarity of Sanskrit to extend *-mi*, the regular first-person ending of verbs which had no vowel affixed to their roots, to those which did have such a vowel.[15]

Leaving out of consideration for the moment differences in vowels and in initial consonants, compare now the present indicative forms as they have developed from Indo-European into Greek and Latin, with special regard to the personal endings:

GREEK	LATIN
pherō[16]	ferō[16]
pherei-s	fer-s[18]
pherei[17]	fer-t
phero-mes (Doric)	feri-mus
phere-te	fer-tis
phero-nti (Doric)	feru-nt

[15] This vowel (for example the *-a* suffixed to the root *bhar-* of the Sanskrit word cited) is called the thematic vowel. The root of a word plus such a suffix is called the stem. To these stems are added endings. The comparatively few verbs lacking such a vowel in Indo-European are called athematic. The *m* in English *am* is a remnant of the Indo-European ending of such athematic verbs.

[16] In Indo-European thematic verbs the first person present indicative had no ending at all, but only a lengthening of the thematic vowel.

[17] The expected form would be *phere-ti*. The ending *-ti*, however, does occur elsewhere in the third person singular, for instance in Doric *didōti* 'he gives.'

[18] In this verb the loss of the thematic vowel is exceptional. The expected forms would be *feri-s, feri-t, feri-tis* in the second and third persons singular and the second person plural respectively. Compare *legō, mittō, scrībō*, and other verbs of the third conjugation, all of which have the thematic vowel throughout, for example *legis, legit, legitis*, and so forth.

Comparison of the personal endings of the verbs in these and other languages leads inevitably to the conclusion that the Indo-European endings had to be as follows (the Indo-European reconstruction of the entire word is given in parentheses):

-ō, -mi	(*bherō)[19]
-si	(*bheresi)
-ti	(*bhereti)
-mes, -mos	(*bheromes)
-te	(*bherete)
-nti	(*bheronti)

Note now in Gothic and Old English the Germanic development of these personal endings:

GOTHIC	EARLY OLD ENGLISH
bair-a	ber-u, -o
bairi-s	biri-s
bairi-þ	biri-þ
bairi-m	bera-þ[20]
bairi-þ	bera-þ
baira-nd	bera-þ

Germanic *þ* corresponds as a rule to Indo-European *t* (see p. 93). Leaving out of consideration such details as the *-nd* (instead of expected *-nþ*) in the Gothic third person plural form, for which there is a soundly based explanation, it is perfectly clear that the Germanic

[19] An asterisk before a form indicates that it is a reconstruction of what can be assumed to have existed on the basis of comparative study. Since Indo-European was spoken only in prehistoric times, all forms cited as existing in that language are necessarily reconstructions; the same is true of cited forms of any language in a prehistoric stage, for instance Germanic and very early Old English. The asterisk is also placed before a form assumed to have been current during the historical period though not actually recorded. Some of the forms from Greek and Latin in the two preceding footnotes might also have been preceded by asterisks, though labeling them "expected forms" in contrast to the forms which are actually attested gives sufficient notice that they are hypothetical and thus satisfies the claims of scholarly integrity.

[20] From the oldest period of Old English the form of the third person plural was used throughout the plural. This form, *beraþ*, from earlier **beranþ*, shows Anglo-Frisian loss of *n* before *þ*.

personal endings correspond to those of the non-Germanic Indo-European languages. A complete comparison of the Germanic languages makes possible a reconstruction of the proto-Germanic endings in the same way that the Indo-European forms have been reconstructed. As has been seen, no guesswork is involved in such reconstruction.

COGNATE WORDS IN THE INDO-EUROPEAN LANGUAGES

Words of similar structure and similar, related, and in many instances identical meanings in the various languages of the Indo-European group may be recognized, once one knows what to expect in the way of sound-shifting, as cognate—that is, of common origin (Lat. *co* plus *gnātus* 'born together'). Thus all the roots just cited (*bhar-, pher-, fer-, bair-, ber-*) are of common origin, all being developments of Indo-European **bher-*; so, for that matter, are the thematic vowels and the personal endings, though the untrained observer may sometimes find it difficult to recognize the relationship. For cognates, as we have seen, do not necessarily look much alike: sound shifts have occurred in the various languages of the Indo-European group (these languages may also be referred to as cognate) which may make related words as unlike in sound as *father*, Sanskrit *pitā*, and Irish Gaelic *athir*[21]—all developments of Indo-European **pətēr*. Sometimes, however, there is sufficient similarity—for example between *maharaja*, ultimately Sanskrit, and Latin *mājus rex* 'great king'—to be apparent even to the untrained observer.

The most frequently cited cognate words are those which have been preserved in a large number of Indo-European languages; some have in fact been preserved in all. These common related words include the numerals from one to ten; the word meaning the sum of ten tens (*cent-, sat-, hund-*) in various quite dissimilar-looking but nonetheless quite regular developments; words for certain bodily parts (related, for example, to *heart, lung, head, foot*); words for certain natural phenomena (related, for example, to *air, night, star, snow, sun, moon, wind*); certain plant and animal names (related, for example, to *beech, corn, wolf, bear*); and certain cultural terms (related, for example, to *yoke, mead, weave, sew*). It is interesting to note in passing that cognates of practically all of our taboo words—those monosyllables which

[21] Indo-European *p*, which corresponds to Germanic *f*, was lost completely in Celtic.

pertain to sex and excretion and which seem to cause great pain to many people—are to be found in other Indo-European languages. Historically, if not socially, these ancient words are just as legitimate as any other words.

One needs no special training to perceive that our *one, two, three* are akin to Latin *ūnus, duo, trēs*; to Greek *oinē* 'one-spot on a die,' *dyo, treis*; to Welsh *un, dau, tri*; to Gothic *ains, twai,* **þreis*; and to Dutch *een, twee, drie*. Comparison of the forms designating the second digit indicates that non-Germanic (as in the Latin, Welsh, and Greek forms) *d* corresponds to Germanic (English, Gothic, Dutch) *t*. A similar comparison of the forms for the third digit indicates that non-Germanic *t* corresponds to Germanic *þ*, the initial sound of *three* and *þrir* in English and Icelandic. Allowing for later changes, as in the case of *þ*, which became *d* in German (*drei* 'three'),[22] as also in Dutch, and *t* in Danish, Norwegian, and Swedish (*tre*), these same correspondences come to light perfectly regularly in other cognates in which the consonants in question appear. We may safely assume, for reasons unnecessary to go into here, that the non-Germanic consonants are older than the Germanic ones. Hence we may accept with the greatest confidence (assuming a similar comparison of the vowel systems) the reconstructions **oinos*, **dwo*, **treies* as accurately representing the Indo-European forms from which the existing forms have developed. The comparative linguists have of course used all the Indo-European languages as a basis for their conclusions regarding correspondences, not just a few such as are cited here.

INDO-EUROPEAN CULTURE

On the basis of these cognates, which must not be confused with loan-words, we can infer a good deal about the state of culture attained by the Indo-Europeans before the various migrations began, probably during the third millennium B.C. or even somewhat earlier. This culture was not contemptible; it was in fact considerably more advanced than that of some groups of people living today. As we have seen, they had a clear sense of family relationship and hence of the family organization, and they could count. They made use of gold and perhaps silver as well; copper and iron were not to come until later. They drank a honey-flavored alcoholic beverage whose name has come down to us as *mead*. Some genius among them had thought up the principle of the wheel, as words corresponding to *wheel, axlē,* and *yoke* testify;

[22] German has *t* from earlier *þ* in a very few words, for instance *tausend* 'thousand.'

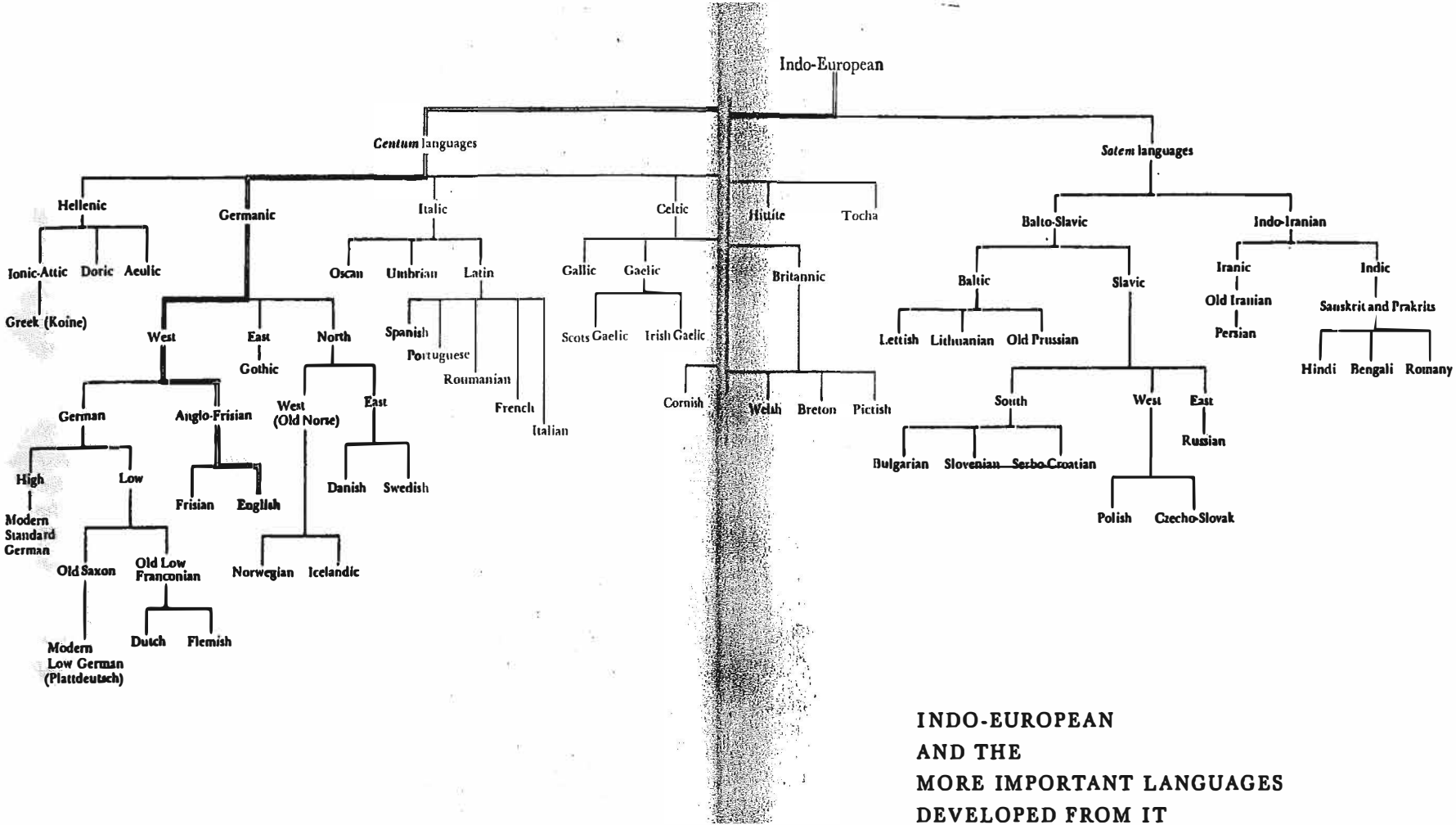
and this was more than the American Indians had done by the nineteenth century. They were small farmers, not nomads, who worked their fields with plows, and they had domesticated animals and fowls. They had religious feeling of a sort, with a conception, not of God, but of gods. This much we can say on the basis of forms which were not actually recorded until long after Indo-European had ceased to be a more or less unified language.

THE INDO-EUROPEAN HOMELAND

Conjectures differ as to the original Indo-European homeland—or at least the earliest for which we have any evidence. Plant and animal names are the principal clues, and the flora and fauna which these denote are northern European. The existence of cognates denoting trees which grow in northern Europe (*oak, birch, willow*), though they may grow elsewhere as well, coupled with the absence of such related words for Mediterranean or Asiatic trees (*olive, cypress, palm*); the similar occurrence of cognate words for *wolf, bear, turtle,* and *salmon,* but none for creatures indigenous to Asia—all this points to northern Europe as the predispersion home, just as the absence of a common word for *ocean* indicates, though it does not in itself prove, that this homeland was inland. Paul Thieme in his cogently reasoned *Die Heimat der indogermanischen Gemeinsprache* (Wiesbaden, 1954) and in the article cited above (n. 7) localizes the Indo-European homeland in the northern part of Central Europe, between the Vistula and the Elbe, on the basis of evidence adduced from the prehistoric geographical distribution of the beech, the turtle, and the salmon. Other Indo-Europeanists have argued from similar evidence for southern Russia, the Carpathians, Scandinavia, and southwestern Asia. The preponderance of scholarly opinion nowadays is in favor of a European center of dispersion—an opinion which implies that the earliest migrations were in a southeasterly direction.

THE MAIN DIVISIONS OF THE INDO-EUROPEAN GROUP

Of some Indo-European languages—for example Phrygian, Scythian, Macedonian, and Illyrian—we possess only the scantiest remains. We may be certain that others have disappeared without leaving a trace. Members of the following subgroups survive as living tongues: Indo-Iranian, Balto-Slavic, Hellenic, Italic, Celtic, and Germanic. Albanian and Armenian are also Indo-European, but do not fit into any of these subgroups.



INDO-EUROPEAN AND THE MORE IMPORTANT LANGUAGES DEVELOPED FROM IT

The Indo-European languages have been conveniently classified into *satem* languages and *centum* languages, *satem* and *centum* being respectively the Avestan (a form of Old Persian) and Latin words corresponding to *hundred*. The classification is based on the development, in very ancient times, of Indo-European palatal *k*. In the *satem* languages—Indo-Iranian, Balto-Slavic, Armenian, and Albanian—this *k* sound became some sort of sibilant: for example Sanskrit (Indic) *śatam*, Lithuanian (Baltic) *šimtas*, Old Slavic *sŭto*.[23] In the other Indo-European languages the earlier *k* of Indo-European **kmtóm* either remained or, in the Germanic group, shifted to *h* in the First Consonant Shift (Grimm's Law), as in Greek (Hellenic) (*he*)*katon*, Welsh (Celtic) *cant*, and Old English (Germanic) *hund*.[24]

The discovery of Tocharian, a *centum* language, early in our century was somewhat disturbing to the general supposition that this division according to the development of Indo-European palatal *k* represented a dialectal split in Indo-European, with those who migrated eastward coming to assibilate the sound. But the assumption of an earlier migration of Tocharians and Hittites, who also spoke a *centum* language, from Central Europe would account for the presence of *centum* languages in what was thought of as *satem* territory. It must be remembered, however, that this is only an inference.

THE INDO-IRANIAN LANGUAGES

The Indo-Iranian (*Iranian* is another form of *Aryan*) group is the oldest for which we have historical records. It is the usual opinion that the Vedic hymns, written in an early form of Sanskrit, date from about the middle of the second millennium B.C. Classical Sanskrit appears more than a thousand years later. It is much more systematized than

[23] Linguistic history often repeats itself: the prehistoric treatment of palatal *k* in the *satem* languages resulting ultimately in its becoming a sibilant was precisely the same thing which took place much later—perhaps about the third century of the Christian Era—in Latin *centum* (and in all other words in which the sound *k*, spelled with *c* in Latin, occurred before the palatal, or front, vowels *e* and *i*). This change is responsible for the occurrence of a sibilant in all the languages derived from Latin; for example the [č] in Italian *cento* and the [s] in French *cent*, Portuguese *cento*, and non-Castilian Spanish *ciento*.

[24] Modern English *hundred* is a compound, first occurring late in the Old English period. The *-red* is a development of what was once an independent word meaning 'number.'

Vedic Sanskrit, for it had been seized upon by a tribe of grammarians who formulated rules for writing it; even so, this was probably not until Sanskrit was ceasing to be widely spoken. The most remarkable of the Indian grammarians was Panini, who, at about the same time (fourth century B.C.) that the Greeks were indulging in more or less reckless speculations about language and in fantastic etymologizing,[25] wrote a grammar of Sanskrit which to this day holds the admiration of linguistic scholars. But there were yet others whose work, motivated as was Panini's by the importance of preserving unchanged the language of the old sacred literature, puts much of the grammatical writing of the Greeks and Romans to shame.

The written language was fixed by these grammarians, and Sanskrit is still written by Indian scholars according to their rules. It is in no sense dead as a written language; its status is roughly comparable to that of Latin in medieval and Renaissance Europe.

Indic dialects had developed, as we might expect, long before Sanskrit became an inflexible and learned language. These are known as Prakrits, and some of them—notably Pali, the religious language of Buddhism—achieved high literary status. From these Prakrits are indirectly derived the various non-Dravidian languages of India, the most widely known of which are Bengali, Hindi, Hindustani (a variety of Hindi, with mixed word stock), and Urdu, derived from Hindustani. Gypsy, or Romany,[26] is also an Indic dialect, with many loan-words from other languages acquired in the course of the Gypsies' wanderings. When they first appeared in Europe in the late Middle Ages, many people supposed them to be Egyptians—whence the name given them in English and some other languages. A long time passed before the study of their language was to indicate unmistakably that they had come originally from northwestern India.

Those Indo-Europeans who settled permanently in the Iranian Plateau developed a sacred language, Avestan, sometimes incorrectly

[25] The Romans later did no better, even deriving names of things from what they were *not*. Thus they fancied *bellum* 'war' was so named because it was not *bellus* 'beautiful.' The Middle Ages and the Renaissance failed to improve much on the Romans.

[26] *Romany* has nothing to do with *Rome, Romance, Romaic* (Modern Greek), or *Roumanian*, but is derived from Gypsy *rom* 'man,' ultimately Sanskrit. The *rye* of *Romany rye* (that is, 'Gypsy gentlemen') likewise has nothing to do with the cereal crop, but is a Gypsy word akin to Sanskrit *rājan* 'king,' as well as to Latin *rex* and German *Reich*.

called Zend,[27] preserved in the religious book the Avesta, after which the language is named. There are no modern descendants of Avestan, which is by some believed to be the language of the Medes, whose name is frequently coupled with that of the Persians, most notably in the phrase "the law of the Medes and Persians, which altereth not" (Daniel vi.8). Avestan was the language of the sage Zarathushtra—Zoroaster to the Greeks—many of whose followers fled to India at the time of the Mohammedan conquest of their country in the eighth century. These are the ancestors of the modern Parsees (that is, Persians) of Bombay. Persian is a different dialect from Avestan; it is the dialect of the district known to the Greeks as Persis, whose inhabitants under the leadership of the great Cyrus in the sixth century B.C. became the predominant tribe.

ARMENIAN AND ALBANIAN

Armenian and Albanian, as we have seen, do not fit into any subgroups. The first has in its word stock so many Persian loan-words that it was once supposed to belong to the Indo-Iranian group; there are also many borrowings from Greek and from Arabic and Syrian. Albanian also has a mixed vocabulary, with words from Italian, Slavic, Turkish, and Greek.

THE BALTO-SLAVIC LANGUAGES

The relationship between the Baltic and the Slavic languages is quite remote, yet so unmistakable that we must assume a common ancestor closer than Indo-European, called Balto-Slavic. The chief Baltic language is Lithuanian; the closely related Lettish is spoken in Latvia, to the north of Lithuania and like it now a part of the Soviet Union. Lithuanian is a notable example of a language which has changed little for thousands of years and hence, like Sanskrit and Greek, has preserved old forms lost in related languages. Still another Baltic language, Prussian, was spoken as late as the seventeenth century in what is now called East Prussia, which was considered outside of Germany until the early years of the nineteenth century. Prussia in time became the predominant state of the new German Empire. The Prussians, like the Lithuanians and the Letts, were heathens until the end of the Middle Ages, when they were converted at the point of the

[27] Actually the Middle Persian language of later commentaries on the Avesta.

sword by the Knights of the Teutonic Order—a military order which was an outcome of the Crusades. The aristocracy of the region (their descendants are the Prussian *Junkers*) came to be made up of members of this order, who, having saved the souls of the heathen Balts, proceeded to take over their lands.

Slavic falls into three main subdivisions: East Slavic includes Great Russian (or just Russian), the common and literary language of Russia; Ruthenian (or Ukrainian), spoken in the Ukraine; and White Russian (or Byelorussian), spoken to the north of the Ukraine. West Slavic includes Polish, Czech, the highly similar Slovak, and Sorbian (or Wendish), a language spoken by a small group of people in East Germany; these languages have lost many of the early forms preserved in East Slavic. The South Slavic languages are Bulgarian, Serbo-Croatian, and Slovenian. The oldest Slavic writing which we know is in Old Bulgarian, sometimes called Old Church Slavic (or Old Church Slavonic), which remained a liturgical language long after it ceased to be generally spoken.

THE DEVELOPMENT OF MODERN GREEK

In ancient times there were many Hellenic dialects, among them Aeolic, Doric, and Ionic, which included Attic. As in the course of history Athens came to assume tremendous prestige, its dialect, Attic—that of all the giants of the Age of Pericles—became the basis of a standard for the entire Greek world, a koine (that is, *koinē* [*dialektos*] 'common [dialect]') which was ultimately to drive out the other Hellenic dialects. The various local dialects spoken in Greece today, as well as the standard language, are thus all derived from Attic. With all their glorious ancient literature, the Greeks have not had a modern literary language until comparatively recently. This "purified" literary language makes considerable use of words revived from ancient Greek, as well as a number of ancient inflectional forms; it has become the ordinary language of the upper classes. A more natural development of the Attic koine is spoken by the masses and hence called *demotike*.

THE ITALIC LANGUAGES

As in the ancient Hellenic world, so also in ancient Italia there were a number of dialects, among them Oscan, Umbrian, and Latin, the speech of Latium, whose chief city was Rome. As Rome came to dominate the Italic world, spreading its influence into Gaul, Spain, and the

Illyrian and Danubian countries (and even into Britain, where it failed to displace Celtic), it became a koine as the dialect of Athens had done.

Spoken Latin, as has been noted, survives in the Romance languages. It was quite a different thing from the more or less artificial literary language of Cicero. All the Romance languages—such as Italian, Spanish, Catalan, Galician, Portuguese, French, Provençal, and Roumanian—are developments of the Vulgar Latin (so called because it was the speech of the *vulgus* 'common people') spoken in various parts of the Roman Empire in the early Middle Ages. Rhaeto-Romanic comprises a number of dialects spoken in the most easterly Swiss canton called the Grisons and in the Tyrol. In southern Belgium a dialect of French, called Walloon, is spoken. Other French dialects have included Norman—the source of the Anglo-Norman dialect spoken in England after the Norman Conquest—Picard, and the dialect of Paris and the surrounding regions (the Île-de-France), which for obvious reasons became Standard French.[28] For similar reasons the speech of the old kingdom of Castile, the largest and most important part of Spain, became Standard Spanish.[29] Because of the cultural preeminence of Tuscany during the Italian Renaissance, the speech of that region—and specifically of the city of Florence—became the standard of Italian speech. Both Dante and Petrarch wrote in this form of Italian; their use of it doubtless had much to do with its victory over other forms.

THE CELTIC LANGUAGES

Celtic shows such striking correspondences with Italic in certain parts of its verbal system and in inflectional endings as to indicate a relationship between them which is rather close, though not so close

[28] The highly similar varieties of French spoken in Quebec, Nova Scotia, New Brunswick, and Louisiana are developments of the dialects of northern France, and are no more to be regarded as "corruptions" of modern Standard French than American English is to be regarded as a corruption of the present British Standard. The "Cajuns" (that is, Acadians) of Louisiana are descendants of exiles from Nova Scotia, which was earlier a French colony called Acadia.

[29] The fact that Spanish America was settled in large part from Andalusia rather than from Castile accounts for the most important differences in pronunciation between Latin-American Spanish and the standard language of Spain.

as that between Indic and Iranian or Baltic and Slavic. Some scholars therefore group them together as developments of a language which they call Celto-Italic.

The Celts were spread over a huge territory in Europe long before the emergence in history of the Germanic peoples. Before the beginning of the Christian Era Celtic languages were spoken over the greater part of central and western Europe, and by the latter part of the third century B.C. even in Asia Minor, in the region called for them Galatia, to whose inhabitants Paul later addressed a famous letter. As the fortunes and the warlike vigor of the Celts declined, their languages were supplanted by those of their conquerors. Thus, the Celtic language spoken in Gaul (called Gallic or Gaulish) gave way completely to the Latin spoken by the Roman conquerors, which was to develop into French.

Roman rule seems to have had comparatively little effect on the continued use of their language by the British Celts. But they were ultimately to give it up after the Angles, Saxons, and Jutes arrived—though not altogether, for British (Brythonic, Britannic) Celtic still lives, if somewhat feebly, in Welsh (Cymric) and in Breton. Breton is the language of the descendants of those Britons who, around the time of the Anglo-Saxon invasion of their island and even somewhat before that time, crossed the Channel and settled in the Gaulish province of Armorica, naming their new home for their old one—Brittany. Breton is thus more closely related to Welsh than to long-extinct Gallic. There have been no native speakers of Cornish, another British language, since the early nineteenth century. Still another British language, Pictish, preserved in a few glosses and place-name elements, was spoken by the Picts in the northwestern part of Britain, in which many Gaelic Celts also settled. These Irishmen, or Scots (*Scotti*) as they were then called, named their new home Scotia, or Scotland.

The Celtic language which spread from Ireland, called Gaelic or Goidelic, was of a type somewhat different from that of the Britons. It was ultimately adopted by the Picts and survives in Scots Gaelic, sometimes called Erse, a word which is simply a variant of *Irish*. Scots Gaelic is spoken in the remoter parts of the Scottish highlands and in a somewhat different development on the Isle of Man (where it is called Manx). In Ireland, which was little affected by either the Roman or the later Anglo-Saxon invasions, Irish Gaelic survived until well into the nineteenth century, but ultimately gave way to English. It has fairly recently been revived for nationalistic reasons in Eire, but

this resuscitation, roughly comparable to the official use of Hebrew in modern Israel, cannot be regarded as in any sense a natural development. It is taught, probably somewhat perfunctorily, in the schools.

THE GERMANIC LANGUAGES

The Germanic group merits a somewhat fuller treatment than has been given to any of the other groups because English belongs to it. In the course of many centuries certain radical developments occurred in the more or less unified language spoken by those Indo-European peoples living in Denmark and the regions thereabout—developments which differentiate it from all the other Indo-European languages more markedly than these are differentiated from Proto-Indo-European and hence from one another. The period during which these developments were in course of occurring we may refer to as Pre-Germanic. *Germanic*[80] is the usual term for the relatively unified language—distinctive in many of its sounds, its inflections, its accentual system, and its word stock—which resulted from these developments.

Unfortunately for us, those who spoke this particular development of Indo-European did not write. Germanic is to German, Dutch, the Scandinavian languages, and English as Latin is to Italian, French, and Spanish. But Germanic, which was probably being spoken shortly before the beginning of the Christian Era, must be reconstructed just like Indo-European. Latin, on the other hand, is amply recorded.

Spread over a large area as Germanic in time came to be, it was inevitable that more and more marked dialectal differences should have occurred, leading to a division into East Germanic, West Germanic, and North Germanic. The only East Germanic language of which we have any detailed knowledge is Gothic. The North Germanic languages are Icelandic, Norwegian, Faroese, Swedish, and Danish. The West Germanic languages are High German, Low German, Dutch, Flemish, Frisian, and English. Some scholars prefer to group Gothic with the North Germanic languages because of certain parallels which indicate that it was closer to Old Norse than to West Germanic. But,

[80] This term seems preferable to Teutonic or Gothonic. When the Romans referred to the Germanic nations (or to Germania), they included in the term Goths, Germans, Angles, Saxons, Frisians, and Scandinavians. Unfortunately *Germanic* has acquired a more limited meaning in English, with nationalistic connotations, but there are equally important objections to the other terms used.

as E. Prokosch has pointed out, "these are hardly more important than the parallels between Gothic and West Germanic on the one hand, and Norse and West Germanic on the other," concluding that "none of these three groups of parallels is important enough to have any influence on the classification of the Germanic languages, and it seems best to consider the three groups independent branches."[81]

THE MAJOR CHANGES FROM INDO-EUROPEAN TO GERMANIC

Germanic became differentiated from Indo-European principally in the following respects:

1. All Indo-European distinctions of tense and aspect[82] were lost in the verb save for the present and the preterit[83] tenses. This simplification of a more complex Indo-European verbal system (though it was not so complex as what developed in Latin, Greek, and Sanskrit) is reflected in all the languages which have developed out of Germanic—in English *bind–bound*, as well as in German *binden–band*, Old Norse *binda–band*, and all the rest. There is in no Germanic language anything comparable to such forms as those of the Latin future, perfect, pluperfect, and future perfect forms (for instance *laudābō, laudāvī, laudāveram, laudāverō*), which must be rendered in the Germanic languages by verb phrases (for instance English *I shall praise, I have praised, I had praised, I shall have praised*).

2. Germanic developed a preterit tense form with a dental suffix, that is, one containing *d* or *t*. All languages derived from Germanic thus have two types of verbs. Those which are distinctly Germanic—

[81] *A Comparative Germanic Grammar* (Philadelphia, 1939), p. 30.

[82] A modification of the form of verbs to indicate whether an action or state is viewed with regard to its beginning, duration, incompletion, completion, or repetition. Languages which lack inflectional forms to indicate such stages in action or being must employ phrases to express them, for instance *he is walking, he is being* (incompleted action or state), *he has walked, he has been* (completed action or state), and the like.

[83] From a historical and comparative point of view, *preterit* is a better term than *past*, which is good enough if our concern is only with the present state of the Germanic languages. *Preterit* designates the absolute past (*walked, was*) without reference to such aspects of past action or state as are expressed in *was walking, was being* (imperfect, or incompleted, action or state), *had eaten, had been* (perfect, or completed, action or state), or to any other stages of past action or state.

that is, those which employ the dental suffix—were called "weak" by Jacob Grimm because, being incapable of the type of internal change of *rise–rose* and *sing–sang* (which he called "strong"), they had to make do with suffixes (*"mit äusseren Mitteln"*), like *step–stepped* and *talk–talked*. Although Grimm's terminology is not very satisfactory, it has become traditional, and is at least as realistic as "regular" and "irregular." An overwhelming majority of our verbs add the dental suffix in the preterit—it is indeed the only living method of inflection for tense in English as in all the other Germanic languages[34]—and this method has thus been thought of as "regular." Historically speaking, however, the vowel gradation of the strong verbs is quite regular, and some of the weak verbs would probably seem quite irregular to the untrained observer. *Bring, think,* and *buy*, for instance, are weak verbs, as the dental suffix of *brought, thought,* and *bought* indicates; the vowel changes here are due not to Indo-European vowel gradation, but to special factors. The suffix is the real test. No attempt at explaining the origin of this suffix has been wholly satisfactory. Many have thought that it was originally an independent word meaning, and cognate with, *do*; but there are grave objections to this theory.

3. For adjectives, Germanic had a so-called weak declension, to be used chiefly when preceded by a pronoun (that is, a pronominal adjective, including the demonstrative pronoun which developed into the definite article). Thus in Old English, *þā geongan ceorlas* 'the young fellows (churls),' with the weak form of *geong*, but *geonge ceorlas* 'young fellows,' with the strong form; likewise in German, *die jungen Kerle*, but *junge Kerle*. This particular Germanic characteristic cannot be illustrated in Modern English, inasmuch as in the course of its development English has fortunately lost all such declension of the adjective.

4. The "free" accentual system of Indo-European, in which any syllable of a word might be accented, gave way to another type of accentuation in which the first syllable was regularly stressed in all words except compound verbs like modern *believe* and *forget*—that is, verbs in which the initial syllable was a prefix. None of the Germanic languages has anything comparable to the shifting accentuation of Latin *vĭrī* 'men,' *vĭrŏrum* 'of the men' or of *hăbĕō*, 'I have,' *habḗmus* 'we

[34] For example, new verbs form their preterit so: *elbow–elbowed, televise–televised, rev–revved,* and so forth. Furthermore, as we shall see later, many verbs which were once strong have become weak.

have.' Compare the paradigms of the Greek and Old English developments of Indo-European **pəter* 'father.'

GREEK	
Singular nominative	patḗr
Singular genitive	patrós
Singular dative	patrí
Singular accusative	patéra
Singular vocative	páter
Plural nominative	patéres (same for vocative)
Plural genitive	patérōn
Plural dative	patrási
Plural accusative	patéras

OLD ENGLISH	
Singular nom., dat., acc.	fæder
Singular genitive	fæder(es)
Plural nom., acc.	fæderas
Plural genitive	fædera
Plural dative	fæderum

In these paradigms it will be noted that in the Greek forms the accent may occur on the suffix, the ending, or the root, unlike the Old English forms, which are representative of the Germanic accentual system in having their accent fixed on the root. Germanic accent is predominantly a matter of stress (loudness) rather than pitch (tone); Indo-European would seem to have had both types of accent at different stages of its development.

5. Indo-European vowels underwent Germanic modification. Indo-European *o*, retained in Latin, became *a* (compare Lat. *octo* 'eight,' Gothic *ahtau*); Indo-European *ā* became *ō* (Lat. *māter* 'mother,' OE *mōdor*); and there were other changes as well, which we shall not go into here.

6. The Indo-European stops *bh, dh, gh, p, t, k, b, d,* and *g*—that is, the *sounds* later symbolized by these letters—all underwent modification in what is called the First Sound Shift or, less happily, Grimm's Law. These modifications were gradual, extending over rather long periods of time. Rather than use reconstructed forms preceded by asterisks, we may illustrate this shift—really a series of shifts—by show-

ing correspondences between a non-Germanic language, usually Latin, and English.[85]

a. Indo-European *bh, dh, gh* became respectively Germanic *β̆, ð, ɣ*,[86] and later, in initial position at least, *b, d, g*. Stated in phonetic terms, aspirated voiced stops became voiced fricatives and then unaspirated voiced stops. Sanskrit preserves Indo-European *bh* and *dh*, but not *gh*, which became *h* in that language. *Bh* appears in Latin as *f*, in Greek as *φ* (an aspirated *p*, later becoming [f]), written *ph* in Latin transcriptions of Greek, as we have seen in Chapter II. *Dh* also became *f* in Latin in initial position, but *d* medially; in Greek the Indo-European sound appears as *θ* (an aspirated *t*), the *th* of Latin transcriptions. *Gh* appears in Latin as *h*, in Greek as *χ* (an aspirated *k*), the *ch* of Latin transcriptions. Unless these non-Germanic changes are borne in mind, the examples to be cited below will not make sense. Except for initial *b*, the other stops are preserved in Latin and Greek. Correspondences may be noted in the following pairs of related words, in which the first member of each pair is Latin unless otherwise labeled:

INDO-EUROPEAN *bh* (LATIN *f*, GREEK *ph*)/GERMANIC *b*

frāter / brother	fundus (for **fudnus*) / bottom
fiber / beaver	fāgus / beech
flāre / blow	(Gr.) phōgein 'to roast' / bake
fra(n)go / break	

INDO-EUROPEAN *dh* (LATIN *f*, GREEK *th*)/GERMANIC *d*

fi(n)gere 'to mold' / dough	(Gr.) thĕ- 'to place' / do
foris / door	(Gr.) thygatĕr / daughter

[85] Derivatives of many of the Latin and Greek cognates to be cited below occur in English as loan-words, some having entered by way of French. Compare, for instance, such pairs as *fraternity–brotherhood, fragile–breakable, fundament–bottom, horticullure–gardening, paternal–fatherly, pyrotechnics–fireworks, pedal–foot, tenuous–thin, cornet–horn, cordial–hearty, canine–hound, gelid–cold*, and so forth.

[86] The *β̆* symbolizes a bilabial fricative, the sound symbolized in Spanish by *b* or *v*. The *ð* stands for precisely the sound which it symbolizes in the alphabet of the International Phonetic Association, the initial consonant of *them*. The *ɣ* indicates the velar fricative plus voice, or resonance. Authorities identify it with the medial consonant of North German *sagen*; but, unless one has a North German handy, this is of little help. The symbol was used in the course of the Old English period with this value, as well as for [g] and [j].

INDO-EUROPEAN *gh* (LATIN *h*, GREEK *ch*)/GERMANIC *g*

hortus / garden	(Gr.) cholē (whence *cholera*) / gall
hostis / guest	(pre)he(n)dere 'to take' / get
homo / gome (obsolete)	hædus 'kid' / goat

b. Except when preceded by *s*, the Indo-European voiceless stops *p, t, k* became respectively the voiceless fricatives *f, þ, x*[87] (later *h* in initial position):

INDO-EUROPEAN *p*/ GERMANIC *f*

pater / father	per / for
piscis / fish	(Gr.) plinthos 'tile' / flint
pellis / fell 'animal hide'	ped- / foot
(Gr.) pyr / fire	pecu 'cattle' / fee (cf. Ger. *Vieh* 'cattle')

INDO-EUROPEAN *t*/ GERMANIC *þ*

trēs / three	tenuis / thin
torrēre 'to dry' / thirst	tumēre 'to swell' / thumb (that is, fat finger)
tu / thou	tonāre / thunder

INDO-EUROPEAN *k*/ GERMANIC *h*

cornū / horn	cent- / hund(red)
cord- / heart	celāre 'to hide' / hele (obsolete), hall, hell
quod / what (OE *hwæt*)	capere 'to take' / heave, have
cervus / hart	canis / hound

c. The Indo-European voiced stops *b, d, g* became respectively the voiceless stops *p, t, k*. Initial *b* was very infrequent in Indo-European. Rasmus Rask cites Greek *kannabis* and Old Norse *hampr* (English *hemp*) as showing the correspondence of *b* and *p*. Latin *turba* 'crowd'

[87] That is, the velar fricative and doubtless also its palatal allophone. The *x* is thus used here without brackets and similarly throughout this chapter with the value which it has in the alphabet of the International Phonetic Association. (IPA uses a different symbol, [ç], for the more forward variety.) It should not be confused with the letter *x* as used since Old English times to spell [ks]. By the time of the earliest writings the symbol used for [x], as well as for the aspirate which it had become initially, was *h*.

and English *thorp* 'town' (as in *Halethorp* and, with metathesis,[88] *Winthrop*) have also been cited. Other certain examples are impossible to find. The shifting of *d* and *g* is illustrated by the following cognates:

INDO-EUROPEAN *d*/ GERMANIC *t*

duo / two	(Gr.) drys 'oak' / tree
dentis / tooth	decem / ten (Gothic *taihun*)
domāre / tame	edere / eat

INDO-EUROPEAN *g*/ GERMANIC *k*

genu / knee (loss of [k-] is modern)	(Gr.) gynĕ 'woman' / queen, quean
ager 'field' / acre	grānum / corn
genus / kin	(g)noscere / know, can

Although we cannot be sure of the chronology of these consonant changes, it is certain that they stretched over centuries—perhaps as much as a millennium. Each set of shifts was completed before the next began; the First Sound Shift was no circular process. It is obvious, for instance, that the shift of Indo-European *b*, *d*, and *g* to Germanic *p*, *t*, and *k* must have occurred long after Indo-European *p*, *t*, and *k* had become Germanic *f*, *þ*, and *x*; otherwise, the Germanic *p*, *t*, and *k* from Indo-European *b*, *d*, and *g* would have gone on to become *f*, *þ*, and *x* also, and we should have no native words with *p*, *t*, and *k*.

FIRST SOUND SHIFT (GRIMM'S LAW)

IE bh, dh, gh ⟶ (respectively) Gmc b, ð, ɣ ⟶ b, d, g
IE p, t, k ⟶ (respectively) Gmc f, þ, x (→ h initially)
IE b, d, g ⟶ (respectively) Gmc p, t, k

The First Sound Shift antedated the Stress Shift described above, for, at a time when the stress had not yet settled on the first syllables of all Germanic words, the voiceless fricatives *f*, *þ*, and *x* underwent a further modification in those words in which they were not immediately preceded by the stressed syllable: they were voiced, becoming respectively *b*, *ð*, and *ɣ*.[89] Under the same circumstances, the voiceless fricative *s*, hitherto unchanged from Indo-European, became *z*. The

[88] A change in the position of sounds or syllables in a word, in this instance of [r].

[89] For a description of the sounds indicated by these symbols, see n. 36.

new *þ*, *ð*, and *ȝ* underwent the same later developments as the *þ*, *ð*, and *ȝ* which had resulted from the shifting of Indo-European *bh*, *dh*, and *gh*. In Old English, for instance, the *ð* appears as *d* (as in all other old West Germanic languages) and the *ȝ* as *g* (under certain circumstances as *w*). The *z* appears as *r*[40] in all recorded Germanic languages save Gothic.

These occurrences of voiced fricatives where according to Grimm's Law we should expect to find voiceless ones were explained by a Danish scholar named Karl Verner in 1875 as being due to the post-position of the stress before the Germanic shift to the first syllable. Hence voicing under such circumstances is said to be according to Verner's Law. Grimm, who recognized but failed to comprehend the reason for the voicing, called it "grammatical change" (*grammatischer Wechsel*). The phenomenon is most obvious in the preterit plural and the past participial forms of some Germanic strong verbs—forms in which the stress did not originally fall on the root: thus, in Old English, *sēaþ* '(I, he, she, it) seethed,' but *sudon* '(we, you, they) seethed' and *soden* (past participle). The *d* of the last two forms has developed from Germanic *ð*; in pre-Stress-Shift Germanic these forms did not have initial stress. It is similar with Old English *lēas* '(I, he, she, it) lost' contrasted with *luron* '(we, you, they) lost' and *loren* (past participle), in which the *s*, after being voiced to become *z* in those forms lacking initial stress, has been subsequently rhotacized.

One further illustration must suffice. Germanic developments of the Indo-European word meaning 'male parent' are evidenced in Gothic *faðar*, written *fadar* (but the *d* in this position indicated the voiced fricative), Icelandic *faðir*, and Old English *fæder* (in which the *d* is, as we have seen, a West Germanic development of earlier *ð*).[41] In all these forms we should ordinarily expect to find *þ* medially, since the Indo-European medial consonant was *t*. But examination of early

[40] This shift of *z* to *r*, known as *rhotacism* (that is, *r*-ing, from Gr. *rho*, the name of the letter) is by no means peculiar to Germanic: compare Latin *flōs* '*flower*,' which has *r* in all forms other than the nominative singular—for instance the genitive singular *flōris*, from earlier **flōz-*, the original *s* being here voiced because of its position between vowels.

[41] The fact that Modern English *father* has the same medial consonant as Proto-Germanic is sheer coincidence. The [ð] in this word, like that in *mother*, is comparatively recent—perhaps no older than the sixteenth century. Earlier [d] was "sometimes heard" in northern England and the lowlands of Scotland at the time of the publication of the relevant section of the *OED* (1895), and may still be.

cognate forms in non-Germanic languages reveals that the stressed syllable in this word followed rather than immediately preceded the *t*, as in Greek *patḗr*, Sanskrit *pitā́*. When the Germanic stress shifted to the first syllable, the reason for the voicing of the consonant was completely obscured. It required comparative linguistics (or comparative philology, as it was then called) to solve the puzzle presented by such apparent exceptions to the workings of the First Sound Shift—to demonstate that they were *only* apparent, and not real exceptions.

VERNER'S LAW

Later developments in the various Germanic languages which have obscured the workings of the shift are not indicated in the table below, where each of the consonants in question plus *a* stands for a syllable headed by that consonant, *o* standing for any preceding syllable; thus *otá* represents Indo-European *pətḗr* 'father,' *kmtóm* 'hundred,' and so forth:

IE opá ——→ Gmc ofá → oƀá → óƀa
IE otá ——→ Gmc oþá → oðá → óða (→ WGmc óda)
IE oká ——→ Gmc oxá → oɣá → óɣa
IE osá (remains) Gmc osá → ozá → óza → óra (except in Gothic)

7. Germanic has a large number of words which have no known cognates in other Indo-European languages. These could of course have existed in Indo-European and have been lost; it is also possible that they were taken from non-Indo-European languages originally spoken in the area occupied by the Germanic peoples. A few words which are apparently distinctively Germanic, given in their Modern English forms, are *rain, drink, drive, broad, hold,*[42] *wife, meat,* and *fowl.*

The earliest records in any Germanic language, aside from a few proper names recorded by classical authors, a few loan-words in Finnish, and some runic inscriptions found in Scandinavia, are those of Gothic.[43] For almost all our knowledge of Gothic we are indebted to

[42] These are all cited in their modern High German forms (*Regen, trinken, treiben, breit, halten*) by Robert Priebsch and W. E. Collinson, *The German Language,* 4th ed., rev. (London, 1958), pp. 281–82.

[43] *Gothic* was in the seventeenth and eighteenth centuries extended to mean 'Germanic,' even in the linguistic sense, but this meaning is now happily obsolete. It also came to mean 'romantically medieval'—a meaning which survives in the name of a fictional genre (*Gothic novel*) and of a style of architecture.

a translation of parts of the New Testament made in the fourth century by Wulfila (*Ulfilas* to the Greeks), bishop of the Visigoths, those Goths who lived north of the Danube. There are also small fragments of two other books of the Bible and of a commentary on the Gospel of John. Late as they are in comparison with the literary records of Sanskrit, Iranian, Greek, and Latin, these remains of Gothic provide us with a clear picture of a Germanic language in an early stage of development and hence are of tremendous importance to the student of Germanic languages. Etymological dictionaries of English cite Gothic cognates of English words (for instance *light–leihts, find–finþan*) when the related Gothic form occurs in the literature cited above. Gothic as a spoken tongue disappeared a long time ago without leaving a trace. No modern Germanic languages are derived from it, nor are there any Gothic loan-words in any of the Germanic languages. Vandalic and Burgundian were apparently also East Germanic in structure, but we know little more of them at first hand than a few proper names.

Certain differences between modern Standard German and the other West Germanic languages are due to a second sound shift—the so-called High German Shift—which occurred comparatively recently as linguistic history goes. It was nearing its completion by the end of the eighth century of our era. This shift began in the southern, mountainous part of Germany and spread northward, stopping short of the low-lying northern section of the country. The *high* in High German (*Hochdeutsch*) and the *low* in Low German (*Plattdeutsch*) refer only to relative distances above sea level. High German became in time Standard German, relegating Low German to the status of a peasant patois in Germany.

FROM GERMANIC TO ENGLISH

The Continental home of the English was north of the area in which the High German Shift occurred. But even if this had not been so, the English language would have been unaffected by changes which had not begun to occur at the time of the Anglo-Saxon migrations to Britain, which began as early as the mid-fifth century. Consequently English has the earlier consonantal characteristics of Germanic, which among the West Germanic languages it shares with Low German, Dutch, Flemish, and Frisian. We may illustrate the High German shift in part by contrasting English and High German forms, as follows: earlier Germanic *p* appears in High German as *pf* or, after vowels, as *ff* (*pepper–Pfeffer*); earlier *t* appears as *ts* (spelled z) or, after vowels,

as *ss* (*tongue–Zunge*; *water–Wasser*); earlier *k* appears after vowels as *ch* (*break–brechen*); earlier *d* appears as *t* (*dance–tanzen*).

The German spoken by more or less simple folk in northern Germany is a development of Old Saxon, and it alone now bears the proper name Low German, though as we have seen it is only one type of low German. Dutch and the practically identical Flemish are the modern forms of Low Franconian, spoken respectively in Holland and, side by side with French, in Belgium. Formerly spoken in a much larger area, including the west coast of Schleswig, Frisian has survived principally in the northern Dutch province of Friesland and in some of the islands off the coast. English and Frisian share certain features not found elsewhere in the Germanic group to such an extent that some scholars regard them as developments of a relatively unified prehistoric language called Anglo-Frisian, a subgroup of West Germanic.

English, then, began its separate existence as a form of Germanic brought by pagan warrior-adventurers from the Continent to the then relatively obscure island which the Romans called Britannia and which had up until a short time before been part of their mighty empire. There, in the next five centuries or so, it was to develop into an independent language quite distinct from any Germanic language spoken on the Continent—a language sufficiently rich in its word stock, thanks largely to the impetus given to learning by the introduction of Christianity, that, as Kemp Malone so well puts it, "by the year 1000, this newcomer could measure swords with Latin in every department of expression, and was incomparably superior to the French speech that came in with William of Normandy."[44]

[44] *A Literary History of England*, ed. Albert C. Baugh (New York, 1948), p. 10.