An Introduction to Language

SEVENTH EDITION

Victoria Fromkin
University of California, Los Angeles (Deceased)

Robert Rodman
North Carolina State University, Raleigh

Nina Hyams
University of California, Los Angeles

THOMSON
HEINLE

Australia Canada Mexico Singapore Spain United Kingdom United States

A group of young linguists known as the **Neo-Grammarians** went beyond the idea that such sound shifts represented only a tendency, and claimed that sound laws have no exception. They viewed linguistics as a natural science and therefore believed that laws of sound change were unexceptionable natural laws. The "laws" they put forth often had exceptions, however, which could not always be explained as dramatically as Verner's Law explained the exceptions to Grimm's Law. Still, the work of these linguists provides important data and insights into language change and why such changes occur.

The linguistic work of the early nineteenth century had some influence on Charles Darwin, and in turn, Darwin's theory of evolution had a profound influence on linguistics and on all science. Some linguists thought that languages had a "life cycle" and developed according to evolutionary laws. In addition, it was believed that every language could be traced to a common ancestor. This theory of biological naturalism has an element of truth to it, but it is an oversimplification of how languages change and evolve into other languages.

Comparative Reconstruction

> . . . Philologists who chase
> A panting syllable through time and space
> Start it at home, and hunt it in the dark,
> To Gaul, to Greece, and into Noah's Ark.
>
> Cowper, "Retirement"

When languages resemble one another in ways not attributable to chance or borrowing, we may conclude they are related. That is, they evolved via linguistic change from an ancestral protolanguage.

The similarity of the basic vocabulary of languages such as English, German, Danish, Dutch, Norwegian, and Swedish is too pervasive for chance or borrowing. We therefore conclude that these languages have a common parent, Proto-Germanic. There are no written records of Proto-Germanic, and certainly no native speakers alive today. Proto-Germanic is a hypothetical language whose properties have been deduced based on its descendants.

In addition to similar vocabulary, the Germanic languages share grammatical properties such as irregularity in the verb *to be*, and similar irregular past-tense forms of verbs, further supporting their relatedness.

Once we know or suspect that several languages are related, their protolanguage may be partially determined by **comparative reconstruction.** One proceeds by applying the **comparative method,** which we illustrate with the following brief example.

Restricting ourselves to English, German, and Swedish, we find the word for "man" is *man*, *Mann*, and *man*, respectively. This is one of many word sets in which we can observe the regular sound correspondence [m]-[m]-[m] and [n]-[n]-[n] in the three languages. Based on this evidence the comparative method has us reconstruct **mVn* as the word for "man" in Proto-Germanic. The *V* indicates a vowel whose quality we are unsure of since, despite the similar spelling, the vowel is phonetically different in the various Germanic languages, and it is unclear how to reconstruct it without further evidence.

Although we are confident that we can reconstruct much of Proto-Germanic with relative accuracy, we can never be sure, and many details remain obscure. To build

confidence in the comparative method, we can apply it to Romance languages such as French, Italian, Spanish, and Portuguese. Their protolanguage is the well-known Latin, so we can verify the method. Consider the following data, focusing on the initial consonant of each word. In these data, *ch* in French is [š] and *c* in the other languages is [k].[7]

French	Italian	Spanish	Portuguese	English
cher	**c**aro	**c**aro	**c**aro	"dear"
champ	**c**ampo	**c**ampo	**c**ampo	"field"
chandelle	**c**andela	**c**andela	**c**andeia	"candle"

The French [š] corresponds to [k] in the three other languages. This regular sound correspondence, [š]-[k]-[k]-[k], supports the view that French, Italian, Spanish, and Portuguese descended from a common language. The comparative method leads to the reconstruction of [k] in "dear," "field," and "candle" of the parent language, and shows that [k] underwent a change to [š] in French, but not in Italian, Spanish, or Portuguese, which retained the original [k] of the parent language, Latin.

To use the comparative method, analysts identify regular sound correspondences in the cognates of potentially related languages. For each correspondence, they deduce the most likely sound in the parent language. In this way, much of the sound system of the parent may be reconstructed. The various phonological changes in the development of each daughter language as it descended and changed from the parent are then identified. Sometimes the sound that analysts choose in their reconstruction of the parent language will be the sound that appears most frequently in the correspondence. This approach was just illustrated with the four Romance languages.

Other considerations may outweigh the "majority rules" principle. The likelihood of certain phonological changes may persuade the analyst to reconstruct a less frequently occurring sound, or even a sound that does not occur in the correspondence. Consider the data in these four hypothetical languages:

Language A	Language B	Language C	Language D
hono	hono	fono	vono
hari	hari	fari	veli
rahima	rahima	rafima	levima
hor	hor	for	vol

Wherever Languages A and B have an *h*, Language C has an *f* and Language D has a *v*. Therefore we have the sound correspondence *h-h-f-v*. Using the comparative method, we might first consider reconstructing the sound *h* in the parent language; but from other data on historical change, and from phonetic research, we know that *h* seldom becomes *v*. The reverse, /f/ and /v/ becoming [h], occurs both historically and as a phonological rule and has an acoustic explanation. Therefore linguists reconstruct an **f* in the parent, and posit the sound change "*f* becomes *h*" in Languages A and B, and "*f* becomes *v*" in Language D. One obviously needs experience and knowledge to conclude this.

[7] Data are taken from Lehmann, 1973.

The other correspondences are not problematic insofar as these data are concerned. They are:

o-o-o-o n-n-n-n a-a-a-e r-r-r-l m-m-m-m

They lead to the reconstructed forms **o, *n, *a, *r,* and **m* for the parent language, and the sound changes "*a* becomes *e*" and "*r* becomes *l*" in Language D. These are natural sound changes found in many of the world's languages.

It is now possible to reconstruct the words of the protolanguage. They are **fono, *fari, *rafima,* and **for.* Language D, in this example, is the most innovative of the three languages, because it has undergone three sound changes. Language C is the most conservative, being identical to the protolanguage insofar as these data are concerned.

The sound changes seen in the previous illustrations are examples of **unconditioned sound change.** The changes occurred irrespective of phonetic context. Below is an example of **conditioned sound change,** taken from three dialects of Italian:

Standard	**Northern**	**Lombard**	
fisso	fiso	fis	"fixed"
kassa	kasa	kasə	"cabinet"

The correspondence sets are:

f-f-f i-i-i o-o-<>[8] k-k-k a-a-a a-a-ə s:-s-s

It is straightforward to reconstruct *f, *i, and *k. Knowing that a geminate like *s:* commonly becomes *s* (recall Old English *f:* became *f*), we reconstruct **s:* for the s:-s-s correspondence. A shortening change took place in the Northern and Lombard dialects.

There is evidence in these (very limited) data for a weakening of word-final vowels, again a change we discussed earlier for English. We reconstruct *o for o-o-<> and *a for a-a-ə. In Lombard, conditioned sound changes took place. The sound *o* was deleted in *word-final position,* but remained *o* elsewhere. The sound *a* became ə in word-final position and remained *a* elsewhere. The conditioning factor is word-final position as far as we can tell from the data presented. Vowels in other position do not undergo change.

We reconstruct the protodialect as having had the words **fisso* meaning "fixed" and **kassa* meaning "cabinet."

It is by means of the comparative method that nineteenth-century linguists were able to initiate the reconstruction of the long-lost ancestral language so aptly conceived by Jones, Bopp, Rask, and Grimm, a language that flourished about 6,000 years ago, the language that we have been calling Indo-European.

Historical Evidence

You know my method. It is founded upon the observance of trifles.

Sir Arthur Conan Doyle, "The Boscombe Valley Mystery," *The Memoirs of Sherlock Holmes*

[8] The empty angled brackets indicate a loss of the sound.

How do we discover phonological changes? How do we know how Shakespeare or Chaucer or the author of *Beowulf* pronounced their versions of English? We have no recordings that give us direct knowledge.

For many languages, written records go back more than a thousand years. Linguists study these records to find out how languages were once pronounced. The spelling in early manuscripts tells us a great deal about the sound systems of older forms of modern languages. Two words spelled differently were probably pronounced differently. Once a number of orthographic contrasts are identified, good guesses can be made as to actual pronunciation. These guesses are supplemented by common words that show up in all stages of the language, allowing their pronunciation to be traced from the present stepwise into the past.

Another clue to earlier pronunciation is provided by non-English words that appear in English manuscripts. Suppose a French word known to contain the vowel [o:] is borrowed into English. The way the borrowed word is spelled reveals a particular letter-sound correspondence.

Other documents can be examined for evidence. Private letters are an excellent source of data. Linguists prefer letters written by naive spellers, who will misspell words according to the way they pronounce them. For instance, at one point in English history all words spelled with *er* in their stems were pronounced as if they were spelled with *ar*, just as in modern British English *clerk* and *derby* are pronounced "clark" and "darby." Some poor speller kept writing *parfet* for *perfect*, which helped linguists to discover the older pronunciation.

Clues are also provided by the writings of the prescriptive grammarians of the period. Between 1550 and 1750 a group of prescriptivists in England known as orthoepists attempted to preserve the "purity" of English. In prescribing how people should speak, they told us how people actually spoke. An orthoepist alive in the United States today might write in a manual: "It is incorrect to pronounce *Cuba* with a final *r*." Future scholars would know that there were speakers of English who pronounced it that way.

Some of the best clues to earlier pronunciation are provided by puns and rhymes in literature. Two words rhyme if the vowels and final consonants are the same. When a poet rhymes the verb *found* with the noun *wound*, it strongly suggests that the vowels of these two words were identical:

BENVOLIO: . . . 'tis in vain to seek him here that means not to be found.
ROMEO: He jests at scars that never felt a wound.

Shakespeare's rhymes are helpful in reconstructing the sound system of Elizabethan English. The rhyming of *convert* with *depart* in Sonnet XI strengthens the conclusion that *er* was pronounced as *ar*.

Dialect differences may provide clues as to what earlier stages of a language were like. Many dialects of English are spoken throughout the world. By comparing the pronunciation of various words in several dialects, we can draw conclusions about earlier forms and see what changes took place in the inventory of sounds and in the phonological rules.

For example, since some speakers of English pronounce *Mary*, *merry*, and *marry* with three different vowels ([meri], [mɛri], and [mæri]), we may conclude that at one

time all speakers of English did so. (The different spellings are also a clue.) For some dialects, however, only one of these sounds can occur before /r/, namely the sound [ɛ]. Those dialects underwent a sound shift in which both /e/ and /æ/ shifted to /ɛ/ when followed immediately by /r/. This is another instance of a conditioned sound shift.

The historical-comparativists working on languages with written records have a difficult job, but not nearly as difficult as scholars who are attempting to discover genetic relationships among languages with no written history.

Linguists must first transcribe large amounts of language data from all the languages, analyze them phonologically, morphologically, and syntactically, and establish a basis for relatedness such as similarities in basic vocabulary, and regular sound correspondences not due to chance or borrowing. Only then can the comparative method be applied to reconstruct the extinct protolanguage.

Linguists proceeding in this manner have discovered many relationships among Native American languages and have successfully reconstructed Amerindian protolanguages. Similar achievements have been made with the numerous languages spoken in Africa. Linguists have been able to group the large number of languages of Africa into four overarching families: Afroasiatic, Nilo-Saharan, Niger-Congo, and Khoisan. For example, Somali is in the Afroasiatic family; Zulu is in the Niger-Congo family; Hottentot, spoken in South Africa, is in the Khoisan family. These familial divisions are subject to revision if new discoveries or analyses deem it necessary.

Extinct and Endangered Languages

> Any language is the supreme achievement of a uniquely human collective genius, as divine and unfathomable a mystery as a living organism.
>
> Michael Krauss

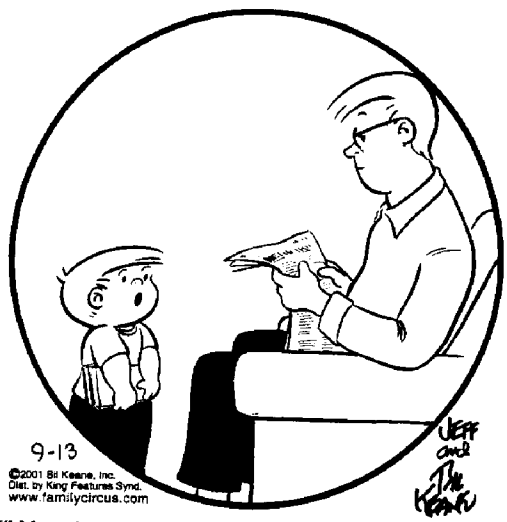
> I am always sorry when any language is lost, because languages are the pedigree of nations.
>
> Samuel Johnson

A language dies and becomes extinct when no children learn it. Linguists have identified four primary types of language death.

Sudden language death occurs when all of the speakers of the language die or are killed. Such was the case with Tasmanian and Nicoleño, a Native American Indian language once spoken in California.

Radical language death is similar to sudden language death in its abruptness. Rather than the speakers dying, however, they all stop speaking the language. Often, the reason for this is survival under the threat of political repression or even genocide. Indigenous languages embedded in other cultures suffer death this way. Speakers, to avoid being identified as "natives," simply stop speaking their native language. Children are unable to learn a language not spoken in their environment, and when the last speaker dies, the language dies.

Gradual language death is the most common way for a language to become extinct. It happens to minority languages that are in contact with a dominant language.



"Was Latin a dead language when you were little, Daddy, or was it still alive?"

much as American Indian languages are in contact with English. In each generation, fewer and fewer children learn the language until there are no new learners. The language is said to be dead when the last generation of speakers dies out. Cornish suffered this fate in Britain in the eighteenth century, as have many Native American languages in both the North and South continents.

Bottom-to-top language death is the term that describes a language that survives only in specific contexts, such as a liturgical language. Latin, and at one time, Hebrew, are such languages. It contrasts with gradual language death, which in its dying throes is spoken casually and informally in homes and villages. People stopped speaking Latin in daily situations centuries ago, and its usage is confined to scholarly and religious contexts.

Language death has befallen, and is befalling, many Native American languages. According to the linguist Michael Krauss, children are learning only 20 percent of the remaining native languages in the United States. Already, hundreds have been lost. Once widely spoken American Indian languages such as Comanche, Apache, and Cherokee have fewer native speakers every generation.

Doomed languages have existed throughout time. The Indo-European languages Hittite and Tocharian no longer exist. Hittite passed away 3,500 years ago, and both dialects of Tocharian gave up the ghost in the first century of the last millennium.

Linguists have placed many languages on an endangered list. They attempt to preserve these languages by studying and documenting their grammars — the phonetics, phonology, and so on — and by recording for posterity the speech of the last few speakers. Through its grammar, each language provides new evidence on the nature of human cognition. In its literature, poetry, ritual speech, and word structure, each language

stores the collective intellectual achievements of a culture, offering unique perspectives on the human condition. The disappearance of a language is tragic; not only are these insights lost, but the major medium through which a culture maintains and renews itself is gone as well.

Dialects, too, may become extinct. Many dialects spoken in the United States are considered endangered by linguists. For example, the sociolinguist Walt Wolfram is studying the dialect spoken on Ocracoke Island off the coast of North Carolina. One reason for the study is to preserve the dialect, which is in danger of extinction because so many young Ocracokers leave the island and raise their children elsewhere, a case of gradual *dialect* death. Vacationers and retirees are diluting the dialect-speaking population, attracted to the island by its unique character, including, ironically, the quaint speech of the islanders.

Linguists are not alone in their preservation efforts. Under the sponsorship of language clubs, and occasionally even governments, adults and children learn an endangered language as a symbol of the culture. Gael Linn is a private organization in Ireland that runs language classes in Irish (Gaelic) for adults. Hundreds of public schools in Ireland and Northern Ireland are conducted entirely in Gaelic. In the state of Hawaii a movement is underway to preserve and teach Hawaiian, the native language of the island.

The United Nations, too, is concerned. In 1991, UNESCO (United Nations Educational, Scientific, and Cultural Organization) passed a resolution that states:

> As the disappearance of any one language constitutes an irretrievable loss to mankind, it is for UNESCO a task of great urgency to respond to this situation by promoting . . . the description — in the form of grammars, dictionaries, and texts — of endangered and dying languages.

Occasionally a language is resurrected from written records. For centuries classical Hebrew was used only in religious ceremonies, but today, with some modernization, and through a great desire among Jews to speak the language of their forefathers, it has become the national language of Israel.

The preservation of dying languages and dialects is essential to the study of Universal Grammar, an attempt to define linguistic properties shared by all languages. This in turn will help linguists develop a comprehensive theory of language that will include a specific description of the innate human capacity for language.

The Genetic Classification of Languages

> The Sanskrit language, whatever be its antiquity, is of a wonderful structure, more perfect than the Greek, more copious than the Latin, and more exquisitely refined than either, yet bearing to both of them a stronger affinity, both in the roots of verbs and in the forms of grammar, than could possibly have been produced by accident; so strong, indeed, that no philologer could examine all three, without believing that they have sprung from some common source, which, perhaps, no longer exists. . . .
>
> Sir William Jones (1786)

We have discussed how different languages evolve from one language and how historical and comparative linguists classify languages into families such as Germanic or Romance and reconstruct earlier forms of the ancestral language. When we examine the languages of the world, we perceive similarities and differences among them that provide evidence for degrees of relatedness or for non-relatedness.

Counting to five in English, German, and Vietnamese shows similarities between English and German not shared by Vietnamese.

English	German	Vietnamese[9]
one	eins	mot
two	zwei	hai
three	drei	ba
four	vier	bon
five	fünf	nam

The similarity between English and German is pervasive. Sometimes it is extremely obvious (*man/Mann*), at other times a little less obvious (*child/Kind*). No regular similarities or differences apart from those due to chance are found between them and Vietnamese.

Pursuing the metaphor of human genealogy, we say that English, German, Norwegian, Danish, Swedish, Icelandic, and so on are sisters in that they descended from one parent and are more closely related to one another than any of them are to non-Germanic languages such as French or Russian.

The Romance languages are also sister languages whose parent is Latin. If we carry the family metaphor to an extreme, we might describe the Germanic languages and the Romance languages as cousins, since their respective parents, Proto-Germanic and early forms of Latin were siblings.

As anyone from a large family knows, there are cousins, and then there are "distant" cousins, encompassing nearly anyone with a claim to family bloodlines. This is true of the Indo-European family of languages. If the Germanic and Romance languages are truly cousins, then languages such as Greek, Armenian, Albanian, and even the extinct Hittite and Tocharian are distant cousins. So are Irish, Scots Gaelic, Welsh, and Breton, whose protolanguage, Celtic, was once widespread throughout Europe and the British Isles. Breton is spoken in Brittany in the northwest coastal regions of France. It was brought there by Celts fleeing from Britain in the seventh century.

Russian is also a distant cousin, as are its sisters, Bulgarian, Serbo-Croatian, Polish, Czech, and Slovak. The Baltic language Lithuanian is related to English, as is its sister language, Latvian. A neighboring language, Estonian, however, is not a relative. Sanskrit, as pointed out by Sir William Jones, though far removed geographically, is nonetheless a relative. Its offspring, Hindi and Bengali, spoken primarily in South Asia, are distantly related to English. Persian (or Farsi), spoken in modern Iran, is a distant cousin of English, as is Kurdish, spoken in Iran, Iraq, and Turkey, and Pashto spoken in Afghanistan and Pakistan.

All the languages mentioned in the last paragraph, except for Estonian, are related, more or less distantly, to one another because they all descended from Indo-European.

[9] Tones are omitted for simplicity.

Figure 11.5 is an abbreviated family tree of the Indo-European languages that gives a genealogical and historical classification of the languages shown. This diagram is somewhat simplified. For example, it appears that all the Slavic languages are sisters. This suggests the comical scenario of speakers of Proto-Slavic dividing themselves into nine clans one fine morning, with each going its separate way. In fact the nine languages shown can be organized hierarchically, showing some more closely related than others. In other words, the various separations that resulted in the nine languages we see today occurred several times over a long stretch of time. Similar remarks apply to the other families, including Indo-European.

Another simplification is that the "dead ends" — languages that evolved and died leaving no offspring — are not included. We have already mentioned Hittite and Tocharian as two such Indo-European languages.

The family tree also fails to show a number of intermediate stages that must have existed in the evolution of modern languages. Languages do not evolve abruptly, which is why comparisons with the genealogical trees of biology have limited usefulness.

Finally, the diagram fails to show a number of Indo-European languages because of lack of space.

Languages of the World

> And the whole earth was of one language, and of one speech.
>
> Genesis 11:1

> Let us go down, and there confound their language, that they may not understand one another's speech.
>
> Genesis 11:7

Most of the world's languages do not belong to the Indo-European family. Linguists have also attempted to classify the non-Indo-European languages according to their genetic relationships. The task is to identify the languages that constitute a family and the relationships that exist among them.

The two most common questions asked of linguists are: "How many languages do you speak?" and "How many languages are there in the world?" Both are difficult to answer precisely. Most linguists have varying degrees of familiarity with several languages, and many are **polyglots,** persons who speak and understand several language. Charles V, the Holy Roman Emperor from 1500 to1558 was a polyglot, for he proclaimed: "I speak Spanish to God, Italian to women, French to men, and German to my horse."

As to the second question, it's hard to ascertain the number of languages in the world because of disagreement as to what comprises a language as opposed to a dialect.

A difficulty with both these questions is that the answers rely on a sliding scale. Familiarity with a language is not an all-or-nothing affair, so how much of a language do you have to know before you can be said to "speak and understand" that language? And how different must two dialects be before they become separate languages? One criterion is that of mutual intelligibility. As long as two dialects remain mutually intelligible, it is generally believed that they cannot be considered separate languages. But mutual

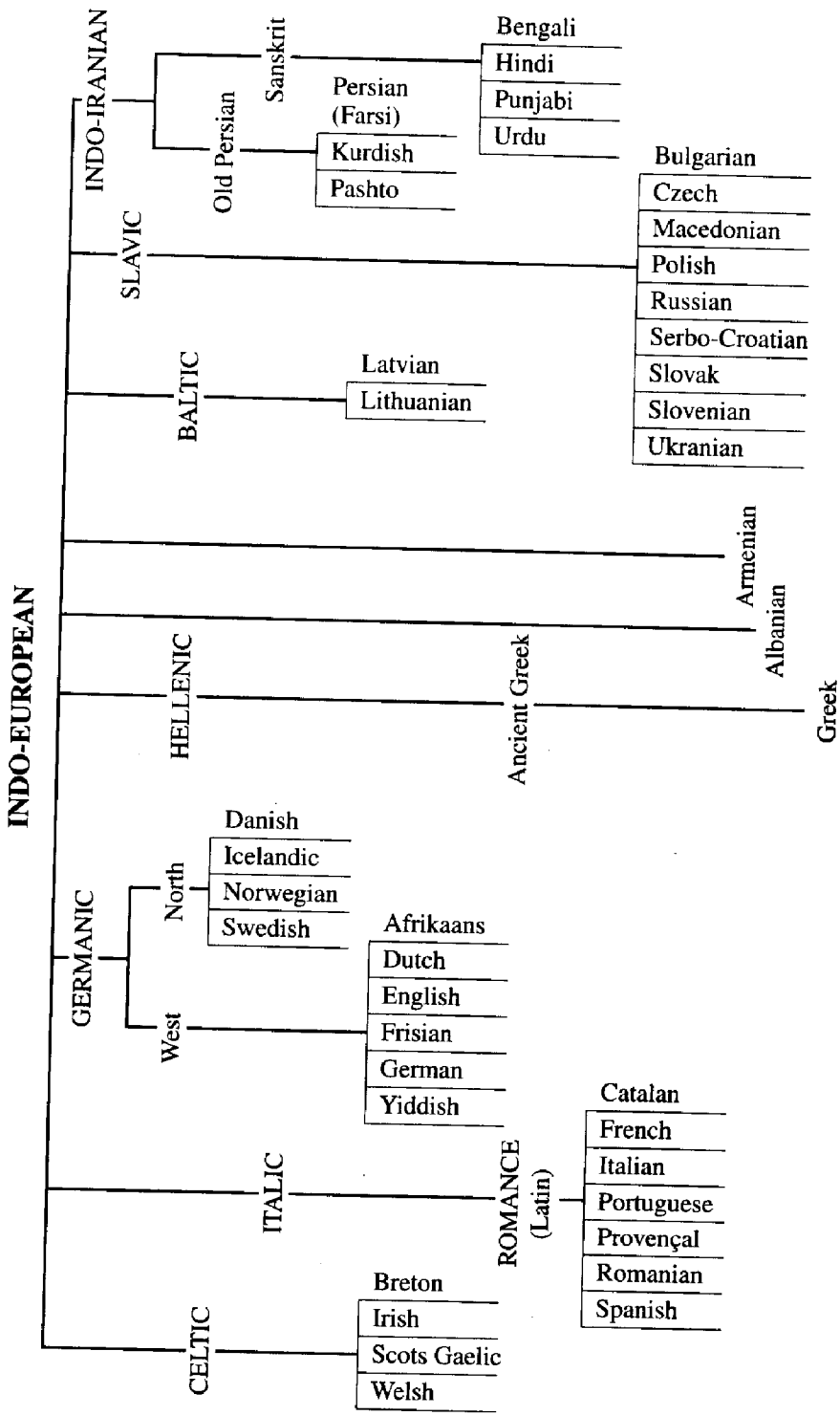
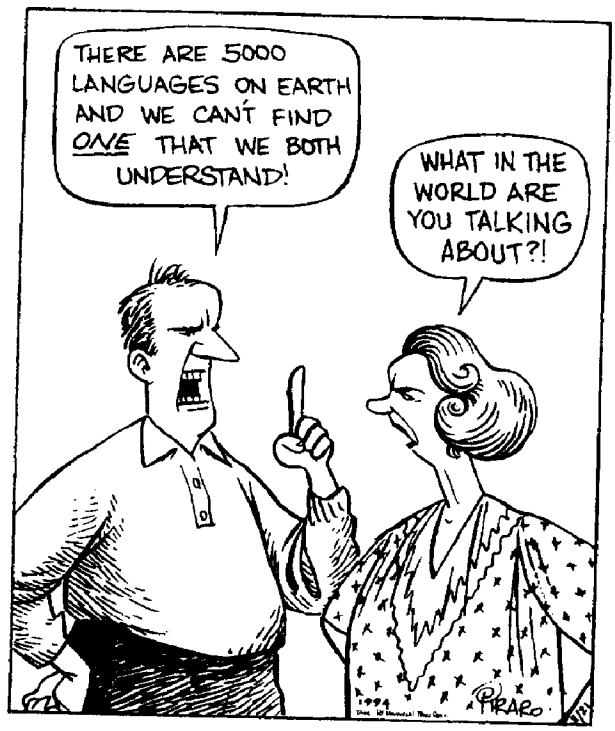


Figure 11.5 The Indo-European family of languages.



intelligibility itself lies on a sliding scale, as all of us know who have conversed with persons speaking dialects of our native language that we do not understand completely.

The Indo-Iranian languages Hindi and Urdu are listed as separate languages in Figure 11.5, yet they are mutually intelligible in their spoken form and are arguably dialects of one language. However, each uses a different writing system and each is spoken in communities of differing religious beliefs and nationalities. (Hindi, for the most part, is spoken in India by Hindus; Urdu is spoken in Pakistan by Muslims.) So what constitutes a separate language is not always determined by linguistic factors alone.

On the other hand, mutually unintelligible languages spoken in China are often thought of as dialects because they have a common writing system and culture, and are spoken within a single political boundary.

Estimates of the number of languages in the world vary widely. The minimum has been set at 4,000 and the maximum at 8,000. In the city of Los Angeles alone, more than 80 languages are spoken. Students at Hollywood High School go home to hear their parents speak Amharic, Armenian, Arabic, Marshallese, Urdu, Sinhalese, Ibo, Gujarati, Hmong, Afrikaans, Khmer, Ukrainian, Cambodian, Spanish, Tagalog, Russian, and more.

It is often surprising to discover what languages are genetically related and which ones aren't. Nepali, the language of remote Nepal is an Indo-European language, whereas Hungarian, surrounded on all sides by Indo-European languages, is not.

It is not possible in an introductory text to give an exhaustive table of families, sub-

families, and individual languages. Besides, a number of genetic relationships have not yet been firmly established. For example, linguists are divided as to whether Japanese and Turkish are related. We'll simply mention several language families with a few of their members. These language families appear not to be related to one another or to Indo-European. This, however, may be an artifact of being unable to delve into the past far enough to see common features that time has erased. We cannot eliminate the possibility that all the world's languages spring ultimately from a single source, an "ur-language" that some have termed **Nostratic,** buried, if not concealed, in the depths of the past. Readers interested in this fascinating topic may wish to read the writings of Professor Johanna Nichols of the University of California at Berkeley.

Uralic is the other major family of languages, besides Indo-European, spoken on the European continent. Hungarian, Finnish, and Estonian are the major representatives of this group.

Afro-Asiatic languages comprise a large family spoken in northern Africa and the Middle East. They include the modern *Semitic* languages of Hebrew and Arabic, as well as languages spoken in biblical times such as Aramaic, Babylonian, Canaanite, and Moabite.

The *Sino-Tibetan* family includes Mandarin, the most populous language in the world, spoken by around one billion Chinese. This family also includes all of the Chinese "dialects," as well as Burmese and Tibetan.

Most of the languages of Africa belong to the *Niger-Congo* family. These include over nine hundred languages grouped into a number of subfamilies such as Kordofanian and Atlantic-Congo. The latter includes individual languages such as Swahili and Zulu.

Equally numerous, the *Austronesian* family contains about nine hundred languages, spoken over a wide expanse of the globe, from Madagascar, off the coast of Africa, to Hawaii. Hawaiian itself, of course, is an Austronesian language, as are Maori, spoken in New Zealand; Tagalog, spoken in the Philippine Islands; and Malay, spoken in Malaysia and Singapore, to mention just a few.

Dozens of families and hundreds of languages are, or were, spoken in North and South America. Knowledge of the genetic relationships among these families of languages is often tenuous, and because so many of the languages are approaching extinction, there may be little hope for as thorough an understanding of the Amerindian language families as linguists have achieved for Indo-European.

Types of Languages

> All the Oriental nations jam tongue and words together in the throat, like the Hebrews and Syrians. All the Mediterranean peoples push their enunciation forward to the palate, like the Greeks and the Asians. All the Occidentals break their words on the teeth, like the Italians and Spaniards. . . .
>
> Isidore of Seville, seventh century C.E.

There are many ways to classify languages. One way already discussed in this chapter is according to the language "family." This method would be like classifying people

"We get a lot of foreign visitors."

"Herman" ® is reprinted with permission from Laughing Stock Licensing Inc., Ottowa, Canada.

according to whether they were related by blood. Another way of classifying languages is by certain linguistic traits, regardless of family. With people, this method would be like classifying them according to height and weight, or hair and eye color.

Every language has sentences that include a subject (S),[10] an object (O), and a verb (V), although some sentences lack all three elements. Languages have been classified according to the basic or most common order in which these occur in sentences.

There are six possible orders — SOV (subject, object, verb), SVO, VSO, VOS, OVS, OSV — permitting six possible language types. Here are examples of some of the languages in these classes.[11]

SVO: English, French, Swahili, Hausa, Thai
VSO: Tagalog, Irish, (Classical) Arabic, (Biblical) Hebrew
SOV: Turkish, Japanese, Persian, Georgian, Eskimo
OVS: Apalai (Brazil), Barasano (Colombia), Panare (Venezuela)
OSV: Apurina and Xavante (Brazil)
VOS: Cakchiquel (Guatemala), Huave (Mexico), Coeur d'Alene (Idaho)

The most frequent word orders in languages of the world are SVO, VSO, and SOV. The basic VSO and SOV sentences may be illustrated as follows:

[10] In this section *only*, S will abbreviate *subject* not *sentence*.

[11] The examples of VOS, OVS, and OSV languages are from Pullum, 1981.

VSO (Tagalog): Sumagot siya sa propesor
answered he the professor
"He answered the professor."

SOV (Turkish): Romalılar barbarları yendiler
Romans barbarians defeated
"The Romans defeated the Barbarians."

Languages with OVS, OSV, and VOS basic word order are much rarer.

The order of other sentence components in a language is most frequently correlated with the language type. If a language is of a type in which the verb precedes the object — a VO language, which includes SVO, VSO, or VOS — then the auxiliary verb tends to precede the verb; adverbs tend to follow the verb; and the language use *pre*positions, which precede the noun, among other such ordering relationships. English exhibits all these tendencies.

In OV languages, most of which are SOV, the opposite tendency occurs: Auxiliary verbs tend to follow the verb; adverbs tend to precede the verb; and there are *post*positions, which function similarly to prepositions but follow the noun. Japanese is an SOV language. It has postpositions, so to say "from Tokyo" in Japanese you say *Tokyo kara*, "Tokyo from." Also in Japanese, the auxiliary verb follows the verb, as illustrated by the following sentence:

Akiko	wa	sakana	o	tabete	iru
Akiko	*topic marker*	fish	*object marker*	eating	is

"Akiko is eating fish."

The correlations between language type and the word order of syntactic categories in sentences are *preferred* word orders, and for the most part are violable tendencies. Different languages follow them to a greater or lesser degree.

The knowledge that speakers of various languages have about word order is captured in the phrase structure rules of the language. In English, an SVO language, the V precedes its NP Object, so the grammar contains the rule VP → V NP. In the SOV languages Turkish and Japanese, the NP Object precedes the Verb and the corresponding phrase structure rule is VP → NP V. Similarly, the rule PP → P NP occurs in SVO languages, whereas the rule PP → NP P is the correlate occurring in SOV languages.

That a language is SVO does not mean that SVO is the only possible word order. When a famous comedian said, "Believe you me" on network TV, he was understood and imitated despite the VSO word order. Yoda, the Jedi Master from the motion picture *Return of the Jedi*, speaks a strange but perfectly understandable style of English that achieves its eccentricity by being OSV. (Objects may be complements other than Noun Phrases.) Some of Yoda's utterances are:

Sick I've become.
Strong with the Force you are.
Your father he is.
When nine hundred years you reach, look as good you will not.

For linguists, the many languages and language families provide essential data for the study of universal grammar. Although these languages are diverse in many ways, they are also remarkably similar in many ways. We find that the languages of the "wretched Greenlanders," the Maoris of New Zealand, the Zulus of Africa, and the native peoples of North and South America all have similar sounds, similar phonological and syntactic rules, and similar semantic systems.

Why Do Languages Change?

> Some method should be thought on for ascertaining and fixing our language forever. . . . I see no absolute necessity why any language should be perpetually changing.
>
> Jonathan Swift (1712)

> Stability in language is synonymous with rigor mortis.
>
> Ernest Weekley

No one knows exactly how or why languages change. As we have shown, linguistic changes do not happen suddenly. Speakers of English did not wake up one morning and decide to use the word *beef* for "ox meat," nor do all the children of one particular generation grow up to adopt a new word. Changes are more gradual, particularly changes in the phonological and syntactic system.

Of course, certain changes may occur instantaneously for any one speaker. When someone acquires a new word, it is not acquired gradually, although full appreciation for all of its possible uses may come slowly. When a new rule enters a speaker's grammar, it is either in or not in the grammar. It may at first be an optional rule, so that sometimes it is used and sometimes it is not, possibly determined by social context or other external factors, but the rule is either there and available for use or not. What is gradual about language change is the spread of certain changes through an entire speech community.

A basic cause of change is the way children acquire the language. No one teaches a child the rules of the grammar. Each child constructs a personal grammar alone, generalizing rules from the linguistic input received. As discussed in chapter 8, the child's language develops in stages until it approximates the adult grammar. The child's grammar is never exactly like that of the adult community, because children receive diverse linguistic input. Certain rules may be simplified or overgeneralized, and vocabularies may show small differences that accumulate over several generations.

The older generation may be using certain rules optionally. For example, at certain times they may say "It's I" and at other times "It's me." The less formal style is usually used with children, who as the next generation may use only the "me" form of the pronoun in this construction. In such cases the grammar will have changed.

The reasons for some changes are relatively easy to understand. Before television there was no such word as *television*. It soon became a common lexical item. Borrowed words, too, generally serve a useful purpose, and their entry into the language is not mysterious. Other changes are more difficult to explain, such as the Great Vowel Shift in English.

One plausible source of change is *assimilation*, a kind of ***ease of articulation*** process in which one sound influences the pronunciation of an adjacent or nearby sound. Due to assimilation, vowels are frequently nasalized before nasal consonants because it is easiest to lower the velum to produce nasality in advance of the actual consonant articulation. This results in the preceding vowel being nasalized. Once the vowel is nasalized, the contrast that the nasal consonant provided can be equally well provided by the nasalized vowel alone, and the redundant consonant may be deleted. The contrast between oral and nasal vowels that exists in many languages of the world today resulted from just such a historical sound change.

In reconstructing older versions of French it has been hypothesized that *bol*, "basin," *botte*, "high boot," *bog*, "a card game," *bock*, "Bock beer," and *bon*, "good," were pronounced [bɔl], [bɔt], [bɔg], [bɔk], and [bɔ̃n], respectively. The nasalized vowel in *bon* was due to the final nasal consonant. Owing to a conditioned sound change that deleted nasal consonants in word-final position, *bon* is pronounced [bɔ̃] in modern French. The nasal vowel alone maintains the contrast with the other words.

Another example from English illustrates how such assimilative processes can change a language. In Old English, word initial [kʲ] (like the initial sound of *cute*), when followed by /i/, was further palatalized to become our modern palatal affricate /č/, as illustrated by the following words:

Old English (c = [kʲ])	**Modern English (ch = [č])**
ciese	cheese
cinn	chin
cild	child

The process of palatalization is found in the history of many languages. In Twi, the word meaning "to hate" was once pronounced [ki]. The [k] became first [kʲ] and then finally [č], so that today "to hate" is [či].

Ease of articulation processes, which make sounds more alike, are countered by the need to maintain contrast. Thus sound change also occurs when two sounds are acoustically similar, with risk of confusion. We saw a sound change of /f/ to /h/ in an earlier example that can be explained by the acoustic similarity of [f] to other sounds.

Analogic change is an "economy of memory" change that results in a reduction of the number of exceptional or irregular morphemes that must be individually learned and remembered. It may be by analogy to *foe/foes* and *dog/dogs* that speakers started saying *cows* as the plural of *cow* instead of the earlier plural *kine*. By analogy to *reap/reaped, seem/seemed,* and *ignite/ignited,* children and adults are presently saying *I sweeped the floor* (instead of *swept*), *I waked last night* (instead of *woke*), and *She lighted the bonfire* (instead of *lit*).

The same kind of analogic change is exemplified by our regularization of exceptional plural forms. We have borrowed words like *datum/data, agendum/agenda, curriculum/curricula, memorandum/memoranda, medium/media, criterion/criteria,* and *virtuoso/virtuosi,* to name just a few. The irregular plurals of these nouns have been replaced by regular plurals among many speakers: *agendas, curriculums, memorandums, criterias,* and *virtuosos.* In some cases the borrowed original plural forms were considered to be the singular (as in *agenda* and *criteria*) and the new plural (e.g., *agendas*) is

therefore a "plural-plural." In addition, many speakers now regard *data* and *media* as nouns that do not have plural forms, like *information.* All these changes lessen the number of irregular forms that must be remembered.

Assimilation and analogic change account for some linguistic changes, but they cannot account for others. Simplification and regularization of grammars occur, but so does elaboration or complication. Old English rules of syntax became more complex, imposing a stricter word order on the language, at the same time that case endings were being simplified. A tendency toward simplification is counteracted by the need to limit potential ambiguity. Much of language change is a balance between the two.

Many factors contribute to linguistic change: simplification of grammars, elaboration to maintain intelligibility, borrowing, and so on. Changes are actualized by children learning the language, who incorporate them into their grammar. The exact reasons for linguistic change are still elusive, though it is clear that the imperfect learning of the adult dialects by children is a contributing factor. Perhaps language changes for the same reason all things change: it is the nature of things to change. As Heraclitus pointed out centuries ago, "All is flux, nothing stays still. Nothing endures but change."

Summary

Languages change. Linguistic change such as **sound shift** is found in the history of all languages, as evidenced by the **regular sound correspondences** that exist between different stages of the same language, different dialects of the same language, and different languages. Languages that evolve from a common source are **genetically related.** Genetically related languages were once dialects of the same language. For example, English, German, and Swedish were dialects of an earlier form of Germanic called **Proto-Germanic,** while earlier forms of Romance languages, such as Spanish, French, and Italian were dialects of Latin. Going back even further in time, earlier forms of Proto-Germanic, Latin, and other languages were dialects of **Indo-European.**

All components of the grammar may change. Phonological, morphological, syntactic, lexical, and semantic changes occur. Words, morphemes, phonemes, and rules of all types may be added, lost, or altered. The meaning of words and morphemes may **broaden, narrow,** or shift. The lexicon may expand by **borrowing,** which results in **loan words** in the vocabulary. It also grows through word **coinage, blends, acronyms** and other processes of word formation. On the other hand, the lexicon may shrink as certain words are no longer used and become obsolete.

No one knows all the causes of linguistic change. Change comes about through the restructuring of the grammar by children learning the language. Grammars may appear to change in the direction of simplicity and regularity, as in the loss of the Indo-European case morphology, but such simplifications may be compensated for by other complexities, such as stricter word order. A balance is always present between simplicity — languages must be learnable — and complexity — languages must be expressive and relatively unambiguous.

Some sound changes result from **assimilation,** a fundamentally physiological

process of **ease of articulation.** Others, like the **Great Vowel Shift,** are more difficult to explain. Some grammatical changes are **analogic changes,** generalizations that lead to more regularity, such as *sweeped* instead of *swept.*

The study of linguistic change is called **historical and comparative linguistics.** Linguists use the **comparative method** to identify regular sound correspondences among the **cognates** of related languages and systematically reconstruct an earlier **protolanguage.** This **comparative reconstruction** allows linguists to peer backward in time and determine the linguistic history of a language family, which may then be represented in a tree diagram similar to Figure 11.5.

Linguists estimate that there are 4,000 to 8,000 languages spoken in the world today (2002). These languages are grouped into families, subfamilies, and so on, based on their genetic relationships. A vast number of these languages are dying out because in each generation fewer children learn them. However, attempts are being made to preserve dying languages and dialects for the knowledge they bring to the study of Universal Grammar and the culture in which they are spoken.

References for Further Reading

Aitchison, J. 1985. *Language Change: Progress or Decay.* New York: Universe Books.

Anttila, R. 1989. *Historical and Comparative Linguistics.* New York: John Benjamins.

Baugh, A. C. 1978. *A History of the English Language,* 3rd ed. Englewood Cliffs, NJ: Prentice-Hall.

Cassidy, F. G., ed. 1986. *Dictionary of American Regional English.* Cambridge, MA: The Belknap Press of Harvard University Press.

Comrie, B., ed. 1990. *The World's Major Languages.* New York: Oxford University Press.

Hock, H. H. 1986. *Principles of Historical Linguistics.* New York: Mouton de Gruyter.

Jeffers, R. J., and I. Lehiste. 1979. *Principles and Methods for Historical Linguistics.* Cambridge, MA: MIT Press.

Lehmann, W. P. 1973. *Historical Linguistics: An Introduction,* 2d ed. New York: Holt, Rinehart and Winston.

Lyovin, A. V. 1997. *An Introduction to Languages of the World.* New York: Oxford University Press.

Nichols, J. 1992. *Linguistic Diversity in Space and Time.* Chicago: University of Chicago Press.

Pullum, G. K. 1981. "Languages with Object before Subject: A Comment and a Catalogue," *Linguistics* 19:147–55.

Pyles, T. 1993. *The Origins and Development of the English Language,* 4th ed. New York: Harcourt Brace.

Renfrew, C. 1989. "The Origins of the Indo-European Languages." *Scientific American* 261.4: 106 – 114.

Ruhlen, M. 1994. *On the Origin of Languages.* Stanford, CA: Stanford University Press.

Traugott, E. C. 1972. *A History of English Syntax.* New York: Holt, Rinehart and Winston.

Voegelin, C. F., and F. M. Voegelin. 1977. *Classification and Index of the World's Languages.* New York: Elsevier.

Wolfram, W. 2001. "Language Death and Dying," In *The Handbook on Language Variation and Change,* Chambers, J. K., Trudgill, P., and Schilling-Estes, N. (eds.) Oxford, UK: Basil Blackwell.