How many languages do you speak? If you are Dutch, you were required to learn at least two foreign languages in high school. For those of you who do not happen to be Dutch, the number is probably a bit lower. In fact, most people in the United States know only English.

Only 30 percent of students graduate from high school in the United States with three or more years of a foreign language. Another 53 percent had had one or two years of a foreign language, and 17 percent of high school graduates have had no foreign language at all. In contrast, 62 percent of graduates from Dutch high schools have learned at least three foreign languages.

Even in other English-speaking countries, foreign languages are studied more frequently than in the United States. For example, 89 percent of British middle- and high-school students are learning a foreign language, more than half of whom are learning French.

Earth’s heterogeneous collection of languages is one of its most obvious examples of cultural diversity. Ethnologue, one of the most authoritative sources of languages (SIL International, www.ethnologue.com), estimates that the world has 7,299 languages. Only ten of these languages, including English, are spoken by at least 100 million people. Several of these are relatively familiar to North Americans (Spanish, Portuguese, Russian, and German), and several are less familiar (Mandarin, Hindi, Bengali, Arabic, and Japanese).

Approximately 100 languages are spoken by at least 5 million people, including the ten largest ones, another 70 by between 2 million and 5 million people. The remaining 6,000 or more languages are spoken by fewer than 2 million people. The distribution of some of these languages is easy for geographers to document, whereas others—especially in Africa and Asia—are difficult, if not impossible.
CASE STUDY

French and Spanish in the United States and Canada

The Tremblay family lives in a suburb of Montréal, Québec. The parents and two young children speak French at home, work, school, and shops. The Lopez family—also two parents and two children—lives in San Antonio, Texas, and speaks Spanish in their household.

The Tremblay and Lopez families share a common condition: they live in countries with an English-speaking majority, but English is not their native language. The French-speaking inhabitants of Canada and the Spanish-speaking residents of the United States continue to speak their languages, although English dominates the political, economic, and cultural life of their countries. The two families use languages other than English because they believe that language is important in retaining and enhancing their cultural heritage. At the same time, both families recognize that knowledge of English is essential for career advancement and economic success.

French is one of Canada's two official languages, along with English. French speakers comprise one-fourth of the country's population. Most French-speaking Canadians are clustered in Québec, where they comprise more than three-fourths of the province's speakers. Colonized by the French in the seventeenth century, Québec was captured by the British in 1763 and in 1867 became one of the provinces in the Confederation of Canada.

In the United States, Spanish has become an increasingly important language in recent years because of large-scale immigration from Latin America. In some communities, public notices, government documents, and advertisements are printed in Spanish. Several hundred Spanish-language newspapers and radio and television stations operate in the United States, especially in southern Florida, the Southwest, and large northern cities, where most of the 28 million Spanish-speaking people live.

These examples—French-speaking residents of Canada and Spanish-speaking residents of the United States—illustrate the "where" and "why" questions that concern geographers who study languages. Where are different languages spoken? English, French, Spanish, and other languages are spoken in distinct locations around the world, and geographers can document the distribution of this important element of cultural identity. Why in some cases are two different languages spoken in two locations, whereas in other cases the same language is spoken in two locations? The geography of language displays especially clearly this book's overall theme of interplay between forces of globalization and local diversity.

Language is a system of communication through speech, a collection of sounds that a group of people understands to have the same meaning. Many languages also have a literary tradition, or a system of written communication. However, hundreds of spoken languages lack a literary tradition. The lack of written record makes it difficult to document the distribution of many languages.

Countries designate at least one language as their official language, which is the one used by the government for laws, reports, and public objects, such as road signs, money, and stamps. A country with more than one official language may require all public documents to be in all languages. Logically, an official language would be understood by most if not all of the country's citizens, but some countries that were once British colonies designate English as an official language, even though few of their citizens can speak it.

Language is part of culture, which, as shown in Chapter 1, has two main meanings—people's values and their tangible artifacts. This chapter and the next two discuss the three traits that best distinguish cultural values—language, religion, and ethnicity. Chapter 4 looked at the material objects of culture. We start our study of the geographic elements of cultural values with language in part because it is the means through which other cultural values, such as religion and ethnicity, are communicated.

Consistent with this book's where and why approach, this chapter first looks at where different languages are used, and how these languages can be logically grouped in space. The second and third key issues examine why languages have distinctive distributions. The study of language follows logically from migration, because the contemporary distribution of languages around the world results largely from past migrations of peoples.
Language is like luggage: people carry it with them when they move from place to place. They incorporate new words into their own language when they reach new places, and they contribute words brought with them to the existing language at the new location. Geographers look at the similarities among languages to understand the diffusion and interaction of people around the world.

The final section of the chapter discusses contradictory trends of scale in language. On the one hand, English has achieved an unprecedented globalization, because people around the world are learning it to participate in a global economy and culture. On the other hand, people are trying to preserve local diversity in language, because language is one of the basic elements of cultural identity and a major feature of a region's uniqueness. Language is a source of pride to a people, a symbol of cultural unity. As a culture develops, language is both a cause of that development and a consequence.

The global distribution of languages results from a combination of two geographic processes—interaction and isolation. People in two locations speak the same language because of migration from one of the locations to another. If the two groups have few connections with each other after the migration, the language spoken by each will begin to differ. After a long period without contact, the two groups will speak languages that are so different they are classified as separate languages.

The interplay between interaction and isolation helps to explain regions of individual languages and entire language families. The difference is that individual languages emerged in the recent past as a result of historically documented events, whereas language families emerged several thousand years before recorded history.

For example, English developed as a distinct language in England as a result of migration and subsequent isolation of Germans 1,500 years ago and Normans 1,000 years ago. Similarly, individual Romance languages developed 2,000 years ago as a result of migration and isolation of Romans to other parts of Europe. On the other hand, the Indo-European language family developed as a result of migration and subsequent isolation of people and can only be reconstructed through linguistic and archaeological theories.

**KEY ISSUE 1**

Where Are English-Language Speakers Distributed?

- Origin and diffusion of English
- Dialects of English

The location of English-language speakers serves as a case study for understanding the process by which any language is distributed around the world. A language originates at a particular place and diffuses to other locations through the migration of its speakers.

**Origin and Diffusion of English**

English is spoken fluently by one-half billion people, more than any language except for Mandarin. Whereas nearly all Mandarin speakers are clustered in one country—China—English speakers are distributed around the world. English is an official language in 50 countries, more than any other language, and is spoken by a significant percentage of people in a number of other countries (Figure 5-1). Two billion people—three out of the world—live in a country where English is an official language, even if they cannot speak it.

**English Colonies**

The contemporary distribution of English speakers around the world exists because the people of England migrated with their language when they established colonies during the past four centuries. Compare Figure 5-1 to Figure 8-4, which shows the location of former British colonies. English is an official language in most of the former British colonies.

English first diffused west from England to North America in the seventeenth century. The first English colonies were built in North America, beginning with Jamestown, Virginia, in 1607, and Plymouth, Massachusetts, in 1620. After England defeated France in a battle to dominate the North American colonies during the eighteenth century, the position of English as the principal language of North America was assured, even after the United States and Canada became independent countries.

Similarly, the British took control of Ireland in the seventeenth century, South Asia in the mid-eighteenth century, the South Pacific in the late eighteenth and early nineteenth centuries, and southern Africa in the late nineteenth century. In each case, English became an official language, even if only the colonial rulers and a handful of elite local residents could speak it.

More recently, the United States has been responsible for diffusing English to several places, most notably the Philippines, which Spain ceded to the United States in 1899, a year after losing the Spanish-American War. After gaining full independence in 1946, the Philippines retained English as one of its official languages, along with Filipino (Tagalog).

**Origin of English in England**

The global distribution of English may be a function primarily of migration from England since the seventeenth century, but that does not explain how English came to be the principal language of the British Isles in the first place, or why English is classified as a Germanic language.

The British Isles had been inhabited for thousands of years, but we know nothing of their early languages, until tribes called the Celts arrived around 2000 B.C., speaking languages we call Celtic. Then, around A.D. 450, tribes from mainland Europe invaded, pushing the Celts into the remote northern and western parts of Britain, including Cornwall and the highlands of Scotland and Wales.
GERMAN INVASION. The invading tribes were the Angles, Jutes, and Saxons. All three were Germanic tribes—the Jutes from northern Denmark, the Angles from southern Denmark, and the Saxons from northwestern Germany (Figure 5-2). Today, English people and others who trace their cultural heritage back to England are often called Anglo-Saxons, after the two larger tribes.

The name England comes from Angles' land. In Old English, Angles was spelled Angles, and the Angles' language was known as englisc. They came from a corner, or angle, of Germany known as Schleswig-Holstein. Modern English has evolved primarily from the language spoken by the Angles, Jutes, and Saxons when they migrated to England 1,500 years ago. The three tribes who brought the beginnings of English to the British Isles came from present-day Denmark and Germany, where they shared a language similar to that of other peoples in the region.

At some time in history, all Germanic people spoke a common language, but that time predates written records. The common origin of English with other Germanic languages can be reconstructed by analyzing language differences that emerged after Germanic groups migrated to separate territories and lived in isolation from each other, allowing their languages to continue evolving independently.

Other peoples subsequently invaded England and added their languages to the basic English. Vikings from present-day Norway landed on the northeast coast of England in the ninth century. Although defeated in their effort to conquer the islands, many Vikings remained in the country to enrich the language with new words.

NORMAN INVASION. English is a good bit different from German today primarily because England was conquered by the Normans in 1066. The Normans, who came from present-day Normandy in France, spoke French, which they established as England's official language for the next 300 years. The leaders of England, including the royal family, nobles, judges, and clergy, therefore spoke French. However, the majority of the people, who had little education, did not know French, so they continued to speak English to each other.

England lost control of Normandy in 1204, during the reign of King John, and entered a long period of conflict with France. As a result, fewer people in England wished to speak French, and English again became the country's unchallenged dominant language. Recognizing that nearly everyone in England was speaking English, Parliament enacted the Statute of Pleading in 1362 to change the official language of court business from French to English. However, Parliament continued to conduct business in French until 1489.

During the 300-year period that French was the official language of England, the Germanic language used by the common people and the French used by the leaders mingled to form a new language. Modern English owes its simpler, straightforward
words, such as sky, horse, man, and woman, to its Germanic roots, and fancy, more elegant words, such as celestial, equestrian, masculine, and feminine to its French invaders.

**Dialects of English**

A dialect is a regional variation of a language distinguished by distinctive vocabulary, spelling, and pronunciation. Generally, speakers of one dialect can understand speakers of another dialect. Geographers are especially interested in differences in dialects, because they reflect distinctive features of the environments in which groups live.

When speakers of a language migrate to other locations, various dialects of that language may develop. This was the case with the migration of English speakers to North America several hundred years ago. Because of its large number of speakers and widespread distribution, English has an especially large number of dialects.

North Americans are well aware that they speak English differently than the British, not to mention people living in India, Pakistan, Australia, and other English-speaking countries. Further, English varies by regions within individual countries. In both the United States and England, northerners sound different from southerners.

In a language with multiple dialects, one dialect may be recognized as the standard language, which is a dialect that is well established and widely recognized as the most acceptable for government, business, education, and mass communication. One particular dialect of English, the one associated with upper-class Britons living in the London area, is recognized in much of the English-speaking world as the standard form of British speech. This speech, known as British Received Pronunciation (BRP), is well known, because it is commonly used by politicians, broadcasters, and actors. Why don’t Americans or, for that matter, other British people speak that way?

**Dialects in England**

“If you use proper English, you’re regarded as a freak; why can’t the English learn to speak?” asked Professor Henry Higgins in the Broadway musical *My Fair Lady*. He was referring to the Cockney-speaking Eliza Doolittle, who pronounced “rain” like “rine” and dropped the /h/ sound from the beginning of words like “happy.” Eliza Doolittle’s speech illustrates that English, like other languages, has a wide variety of dialects that use different pronunciations, spellings, and meanings for particular words.
As already discussed, English originated with three invading groups from Northern Europe who settled in different parts of Britain—the Angles in the north, the Jutes in the southeast, and the Saxons in the south and west. The language each spoke was the basis of distinct regional dialects of Old English. Kentish dialect in the southeast, West Saxon in the southwest, Mercian in the center of the island, and Northumbrian dialect in the north (Figure 5–3, left).

French replaced English as the language of the government and aristocracy following the Norman invasion of 1066. By the time English again became the country’s dominant language, five major regional dialects had emerged—Northeastern, East Midlands, West Midlands, Southwestern, and Southeastern or Kentish. The boundaries of these five regional dialects roughly paralleled the pattern before the Norman invasion (compare Figure 5–3, left and right). However, after several hundred years of living in isolation in rural settlements under the control of a French-speaking government, people spoke English differently in virtually every county of England.

From this large collection of local dialects, one eventually emerged as the standard language for writing and speech throughout England—the dialect used by upper-class residents in the capital city of London and the two important university cities of Cambridge and Oxford. The diffusion of the dialect spoken in London and the university cities was first encouraged by the introduction of the printing press to England in 1476. Grammar books and dictionaries printed in the eighteenth century established rules for spelling and grammar that were based on the London dialect. These frequently arbitrary rules were then taught in schools throughout the country.

Despite the current dominance of BRP, strong regional differences persist in English dialects spoken in the United Kingdom, especially in rural areas. Although several dozen dialects are identifiable, they can be grouped into three main ones—Northern, Midland, and Southern. People in the south of England pronounce words like grass and path with an /æ/ sound, whereas people in the Midlands and North use a short /æ/ as do most people in the United States. People in the Midlands and North pronounce butter and Sunday with the /oʊ/ sound of words like boot. Northerners pronounce ground and pound like grund and ground, with the /ʌ/ sound similar to the word punt in U.S. football.

Further, distinctive southwestern and southeastern accents occur within the Southern dialect. People in the southwest, for example, pronounce that and thing with the /æt/ sound of then, rather than thin. Fresh and egg have an /æt/ sound. Southeasterners pronounce the /æ/ in apple and cat like the short /æ/ in bet. Local dialects can be further distinguished, and some words have distinctive pronunciations and meanings in each county of the United Kingdom.

Differences Between British and American English

The English language was brought to the North American continent by colonists from England who settled along the Atlantic Coast beginning in the seventeenth century. The early colonists naturally spoke the language used in England at the time and
established seventeenth-century English as the dominant form of European speech in colonial America. Later immigrants from other countries found English already implanted here. Although they made significant contributions to American English, they became acculturated into a society that already spoke English. Therefore, the earliest colonists were most responsible for the dominant language patterns that exist today in the English-speaking part of the Western Hemisphere.

DIFFERENCES IN VOCABULARY AND SPELLING. Why is the English language in the United States so different from that in England? As is so often the case with languages, the answer is isolation. Separated by the Atlantic Ocean, English in the United States and England evolved independently during the eighteenth and nineteenth centuries, with little influence on one another. Few residents of one country could visit the other, and the means to transmit the human voice over long distances would not become available until the twentieth century.

U.S. English differs from that of England in three significant ways—vocabulary, spelling, and pronunciation. The vocabulary is different largely because settlers in America encountered many new objects and experiences. The new continent contained physical features, such as large forests and mountains, which had to be given new names. New animals were encountered, including the moose, raccoon, and chipmunk, all of which were given names borrowed from Native Americans. Indigenous American “Indians” also enriched American English with names for objects such as canoe, moccasin, and squash.

As new inventions appeared, they acquired different names on either side of the Atlantic. For example, the elevator is called a lift in England, and the flashlight is known as a torch. The British call the hood of a car the bonnet and the trunk the boot.

Spelling diverged from the British standard because of a strong national feeling in the United States for an independent identity. Noah Webster, the creator of the first comprehensive American dictionary and grammar books, was not just a documenter of usage, he had an agenda. Webster was determined to develop a uniquely American dialect of English. He either ignored or was unaware of recently created rules of grammar and spelling developed in England.

Webster argued that spelling and grammar reforms would help establish a national language, reduce cultural dependence on England, and inspire national pride. The spelling differences between British and American English, such as the elimination of the “u” from the British spelling of words like “honour” and “colour” and the substitution of “s” for “c” in “defence” are due primarily to the diffusion of Webster’s ideas inside the United States.

DIFFERENCES IN PRONUNCIATION. Differences in pronunciation between British and U.S. speakers are immediately recognizable. Again, geographic concepts help explain the reason for the differences. From the time of their arrival in North America, colonists began to pronounce words differently from the British. Such divergence is normal, for interaction between the two groups was largely confined to exchange of letters and other printed matter rather than direct speech.

One prominent difference between British and U.S. English is the pronunciation of the letters a and r. Such words as fast, path, and half are pronounced in England like the /æ/ in father rather than the /ɑː/ in man. The British also eliminate the letter r from pronunciation except before vowels. Thus lord in British pronunciation sounds like lord. Further, Americans pronounce unaccented syllables with more clarity. The words secret and necessary have four syllables in American English but only three in British (secret’ry and neces’sry).

Surprisingly, pronunciation has changed more in England than in the United States. The letters s and r are pronounced in the United States the way they used to be pronounced in Britain, specifically in the seventeenth century when the first colonists arrived.

A single dialect of Southern English did not emerge as the British national standard until the late eighteenth century, after the American colonies had declared independence and were politically as well as physically isolated from England. Thus people in the United States do not speak “proper” English because when the colonists left England, “proper” English was not what it is today. Furthermore, few colonists were drawn from the English upper classes.

Dialects in the United States

Major differences in U.S. dialects originated because of differences in dialects among the original settlers. The English dialect spoken by the first colonists, who arrived in the seventeenth century, determined the future speech patterns for their communities because later immigrants adopted the language used in their new homes when they arrived. The language may have been modified somewhat by the new arrivals, but the distinctive elements brought over by the original settlers continued to dominate.

SETTLEMENT IN THE EAST. The original American settlements stretched along the Atlantic Coast in 13 separate colonies. The settlements can be grouped into three areas—New England, Middle Atlantic, and Southeastern. Massachusetts and the other New England colonies were established and inhabited almost entirely by settlers from England. Two-thirds of the New England colonists were Puritans from East Anglia in southeastern England, and only a few came from the north of England.

The nucleus of the southeastern colonies was Virginia, where the first permanent settlement by the English in North America was established at Jamestown in 1607. About half of the southeastern settlers came from southeast England, although they represented a diversity of social-class backgrounds, including deported prisoners, indentured servants, and political and religious refugees. The English dialects now spoken in the U.S. Southeast and New England are easily recognizable. Current distinctions result from the establishment of independent and isolated colonies in the seventeenth century.

The immigrants to the Middle Atlantic colonies were more diverse. The early settlers of Pennsylvania were predominantly Quakers from the north of England. Scots and Irish also went to Pennsylvania, as well as to New Jersey and Delaware. In addition, the Middle Atlantic colonies attracted many German, Dutch, and Swedish immigrants who learned their English from the English-speaking settlers in the area. The dialect spoken in the Middle
Atlantic colonies thus differed significantly from those spoken farther north and south, because most of the settlers came from the north rather than the south of England or from other countries.

**CURRENT DIALECT DIFFERENCES IN THE EAST.** Today, major dialect differences within the United States continue to exist, primarily on the East Coast, although some distinctions can be found elsewhere in the country. The different dialects have been documented through the study of particular words. Every word that is not used nationally has some geographic extent within the country and therefore has boundaries. Such a word-usage boundary, known as an *isogloss*, can be constructed for each word. These isoglosses are determined by collecting data directly from people, particularly natives of rural areas. These people are shown pictures to identify or are given sentences to complete with a particular word. Although every word has a unique isogloss, boundary lines of different words coalesce in some locations to form regions.

Two important isoglosses separate the eastern United States into three major dialect regions, known as Northern, Midlands, and Southern. The northern boundary runs across Pennsylvania, whereas the southern one runs along the Appalachian Mountains (Figure 5-4).

Some words are commonly used within one of the three major dialect areas but rarely in the other two. In most instances, these words relate to rural life, food, and objects from daily activities. Language differences tend to be greater in rural areas than in cities, because farmers are relatively isolated from interaction with people from other dialect regions.

For example, a container commonly used on farms is known as a “pail” in the north and a “bucket” in the Midlands and South. A small stream is known as a “brook” in the North, a “run” in the Midlands, and a “branch” in the South. The term “run” was apparently used in the north of England and Scotland, which was the area of origin for many Middle Atlantic settlers but for few New England or Southern settlers.

Phrases for some farm activities, such as calling cows from pasture, show particularly sharp differences among the three regional dialects. New England farmers call cows with “Boss!” or “Bossee!” sometimes preceded by “Co” or “Come.” In the Midlands the preferred call is “Sook!” or sometimes “Sookie!” or “Sook cow!” The choice in the South is “Co-wench!” or its alternative forms, “Co-inch!” and “Co-eel.”

Many words that were once regionally distinctive are now national in distribution. Mass media, especially television and radio, influence the adoption of the same words throughout the country. For example, a “frying pan” was once commonly called a “skillet” in New England and a “skillet” in the Middle Atlantic area.

**PRONUNCIATION DIFFERENCES.** Regional pronunciation differences are more familiar to us than word differences, although it is harder to draw precise isoglosses for them. Pronunciations that distinguish the Southern dialect include making such words as *bail* and *mine* into two syllables (“ha-aff” and “mi-yeen”), pronouncing *poor* as “po-or,” and pronouncing *Tuesday* and *due* with a /r/ sound (“Tyuesday” and “dyuee”).

The New England accent is well known for dropping the /r/ sound, so that *heart* and *lark* are pronounced “hot” and “lock.” Also, *ear* and *are* are pronounced with /ar/ substituted for the /r/ endings. This characteristic dropping of the /r/ sound is shared with speakers from the south of England and reflects the place of origin of most New England colonists. It also reflects the relatively high degree of contact between the two groups. Residents of Boston, New England’s main port city, maintained especially close ties to the important ports of southern England, such as London, Plymouth, and Bristol. Compared to other colonists, New Englanders received more exposure to changes in pronunciation that occurred in Britain during the eighteenth century.

The New England and southern accents sound odd to the majority of Americans because the standard pronunciation throughout the American West comes from the Middle Atlantic states rather than the New England and Southern regions. This pattern occurred because most western settlers came from the Middle Atlantic states.

The diffusion of particular English dialects into the middle and western parts of the United States is a result of the westward movement of colonists from the three dialect regions of
the East. The area of the Midwest south of the Ohio River was settled first by colonists from Virginia and the other southern areas. The Middle Atlantic colonies sent most of the early settlers north of the Ohio River, although some New Englanders moved to the Great Lakes area. The pattern by which dialects diffused westward resembles the diffusion of East Coast house types discussed in Chapter 4 (compare Figure 5–4 with Figure 4–9).

As more of the West was opened to settlement during the nineteenth century, people migrated from all parts of the East Coast. The California gold rush attracted people from throughout the East, many of whom subsequently moved to other parts of the West. The mobility of Americans has been a major reason for the relatively uniform language that exists throughout much of the West.

KEY ISSUE 2

Why Is English Related to Other Languages?

- Indo-European branches
- Origin and diffusion of Indo-European

English is part of the Indo-European language family. A language family is a collection of languages related through a common ancestral language that existed long before recorded history. Indo-European is the world’s most extensively spoken language family by a wide margin. Nearly 3 billion people speak an Indo-European language as their first language.

Indo-European Branches

Within a language family, a language branch is a collection of languages related through a common ancestral language that existed several thousand years ago. Differences are not as extensive as or as old as with language families, and archaeological evidence can confirm that the branches derived from the same family.

Indo-European is divided into eight branches. Four of the branches—Indo-Iranian, Romance, Germanic, and Balto-Slavic—are spoken by large numbers of people. Indo-Iranian languages are clustered in South Asia, Romance languages in southwestern Europe and Latin America, Germanic languages in northwestern Europe and North America, and Balto-Slavic languages in Eastern Europe. The four less extensively used Indo-European language branches are Albanian, Armenian, Greek, and Celtic (Figure 5–5).

Germanic Branch of Indo-European

German may seem a difficult language for many English speakers to learn, but the two languages are actually closely related. Both belong to the Germanic language branch of Indo-European. English is part of the Germanic branch of the Indo-European family because of the language spoken by the Germanic tribes that invaded England 1,500 years ago.

A language group is a collection of languages within a branch that share a common origin in the relatively recent past and display relatively few differences in grammar and vocabulary. West Germanic is the group within the Germanic branch of Indo-European to which English belongs. Although they sound very different, English and German are both languages in the West Germanic group within the Germanic branch because they are structurally similar and have many words in common (Figure 5–6).

West Germanic is further divided into High Germanic and Low Germanic subgroups, so named because they are found in high and low elevations within present-day Germany. High German, spoken in the southern mountains of Germany, is the basis for the modern standard German language. English is classified in the Low Germanic subgroup of the West Germanic group. Other Low Germanic languages include Dutch, which is spoken in the Netherlands, as well as Flemish, which is generally considered a dialect of Dutch spoken in northern Belgium. Afrikaans, a language of South Africa, is similar to Dutch, because Dutch settlers migrated to South Africa 300 years ago. Frisian is spoken by a few residents in northeastern Netherlands. A dialect of German spoken in the northern lowlands of Germany is also classified as Low German.

The Germanic language branch also includes North Germanic languages, spoken in Scandinavia. The four Scandinavian languages—Swedish, Danish, Norwegian, and Icelandic—all derive from Old Norse, which was the principal language spoken throughout Scandinavia before A.D. 1000. Four distinct languages emerged after that time because of migration and the political organization of the region into four independent and isolated countries.

Indo-Iranian Branch of Indo-European

The branch of the Indo-European language family with the most speakers is Indo-Iranian. This branch includes more than 100 individual languages, spoken by 1 billion people. The branch can be divided into an eastern group (Indic) and a western group (Iranian).
INDIC (EASTERN) GROUP OF INDO-IRANIAN LANGUAGE BRANCH. The most widely used languages in India, as well as in the neighboring countries of Pakistan and Bangladesh, belong to the Indic-European language family and, more specifically, to the Indic group of the Indo-Iranian branch of Indo-European.

Approximately one-third of Indians, mostly in the north, use an Indic language called Hindi. Hindi is spoken many different ways—and therefore could be regarded as a collection of many individual languages—but there is only one official way to write the language, using a script called Devanagari, which has been used in India since the seventh century A.D. (For example, the word for "sun" is written in Hindi as सूरा pronounced suraj.) Local differences arose in the spoken forms of Hindi but not in the written form, because until recently few speakers of that language could read or write it.

Pakistan's principal language, Urdu, is spoken very much like Hindi but is written with the Arabic alphabet, a legacy of the fact that most Pakistanis are Muslims, and their holiest book (the Quran) is written in Arabic. The basis of both languages is Hindustani, a form of the language used in communication among different groups of people in much of India for many centuries. Hindi, originally a variety of Hindustani spoken in the area of New Delhi, grew into a national language in the nineteenth century when the British encouraged its use in government. Collectively, Indic languages constitute the world's second-largest language group.

One of the main elements of cultural diversity among the 1 billion plus residents of India is language (Figure 5-7). India has four important language families—Indo-European (predominantly in the north), Dravidian (in the south), Sino-Tibetan (in the northeast), and Austro-Asiatic (in the central and eastern highlands).

After India became an independent state in 1947, Hindi was proposed as the official language, but Dravidian speakers from southern India strongly objected. Therefore, India's constitution as amended recognizes 18 official languages, including 13 Indo-European (Assamese, Bengali, Gujarati, Hindi, Kashmiri, Konkani, Marathi, Nepali, Oriya, Panjabi, Sanskrit, Sindhi, and Urdu), four Dravidian languages (Kannada, Malayalam, Tamil, and Telugu), and one Sino-Tibetan language (Manipur). More than 90 percent of the population speak at least one of these languages, but as many as 10 million Indians use other languages. Bengali is the most important language in Bangladesh.

As the language of India's former colonial ruler, English has an "associate" status, even though only 1 percent of the Indian
Balto-Slavic Branch of Indo-European

The other Indo-European language branch with large numbers of speakers is Balto-Slavic. Slavic was once a single language, but differences developed in the seventh century A.D. when several groups of Slavs migrated from Asia to different areas of Eastern Europe and thereafter lived in isolation from one another. As a result, this branch can be divided into East, West, and South Slavic groups as well as a Baltic group. Figure 7-21 shows the widespread area populated with Balto-Slavic speakers.

EAST SLAVIC AND BALTIC GROUPS OF THE BALTO-SLAVIC LANGUAGE BRANCH. The most widely used Slavic languages are the eastern ones, primarily Russian, which is spoken by more than 80 percent of Russian people. Russian is one of the six official languages of the United Nations.

The importance of Russian increased with the Soviet Union's rise to power after the end of World War II in 1945. Soviet officials forced native speakers of other languages to learn Russian as a way of fostering cultural unity among the country's diverse peoples. In Eastern European countries that were dominated politically and economically by the Soviet Union, Russian was taught as the second language. With the demise of the Soviet Union, the newly independent republics adopted official languages other than Russian, although Russian remains the language for communications among officials in the countries that were formerly part of the Soviet Union.

After Russian, Ukrainian and Belarusian are the two most important East Slavic languages, official languages in Ukraine and Belarus. Ukraine is a Slavic word meaning “border,” and Bela- is translated as “white.” The presence of so many non-Russian speakers was a measure of cultural diversity in the Soviet Union, and the desire to use languages other than Russian was a major drive in its breakup a decade ago.

WEST AND SOUTH SLAVIC GROUPS OF THE BALTO-SLAVIC LANGUAGE BRANCH. The most spoken West Slavic language is Polish, followed by Czech and Slovak. The latter two are quite similar, and speakers of one can understand the other. The government of the former state of Czechoslovakia tried to balance the use of the two languages, even though the country contained twice as many Czechs as Slovaks. For example, the announcers on televised sports events used one of the languages during the first half and switched to the other for the second half. These balancing measures were effective in promoting national unity during the Communist era, but in 1993, 4 years after the fall of communism, Slovakia split from the Czech Republic. Slovaks rekindled their long-suppressed resentment of perceived dominance of the national culture by the Czech ethnic group.

The most important South Slavic language is the one spoken in Bosnia & Herzegovina, Croatia, Montenegro, and Serbia. Bosnians and Croats write the language in the Roman alphabet (what you are reading now), whereas Montenegrans and Serbs use the Cyrillic alphabet (for example, Yugoslavia is written ЈУГОСЛАВИЈА).

When Bosnia & Herzegovina, Croatia, Montenegro, and Serbia were all part of Yugoslavia, the language was called Serbo-Croatian. This name now offends Bosnians and Croats,
because it recalls when they were once in a country that was dominated by Serbs. Instead, the names Bosnian, Croatian, and Serbian are preferred by people in these countries, to demonstrate that each language is unique, even though linguists consider them one.

Still, differences have crept into the language of the South Slavs. Bosnian Muslims have introduced Arabic words used in their religion, and Croats have replaced words regarded as having a Serbian origin with words considered to be purely Croatian. For example, the Serbo-Croatian word for martyr or hero—janak—has been changed to heraj by Croats and shahid by Bosnian Muslims. In the future, after a generation of isolation and hostility among Bosnians, Croats, and Serbs, the languages spoken by the three may be sufficiently different to justify their classification as distinct languages.

In general, differences among all of the Slavic languages are relatively small. A Czech, for example, can understand most of what is said or written in Slovak and could become fluent without much difficulty. However, because language is a major element in a people's cultural identity, relatively small differences among Slavic as well as other languages are being preserved and even accentuated in recent independence movements.

**Romance Branch of Indo-European**

The Romance language branch evolved from the Latin language spoken by the Romans 2,000 years ago. The four most widely used contemporary Romance languages are Spanish, Portuguese, French, and Italian. Spanish and French are two of the six official languages of the United Nations.

The European regions in which these four languages are spoken correspond somewhat to the boundaries of the modern states of Spain, Portugal, France, and Italy. Rugged mountains serve as boundaries among these four countries. France is separated from Italy by the Alps and from Spain by the Pyrenees, and several mountain ranges mark the border between Spain and Portugal. Physical boundaries such as mountains are strong intervening obstacles, creating barriers to communication between people living on opposite sides.
The fifth most important Romance language, Romanian, is the principal language of Romania and Moldova. It is separated from the other Romance-speaking European countries by Slavic-speaking peoples.

The distribution of Romance languages shows the difficulty in trying to establish the number of distinct languages in the world. In addition to the five languages already mentioned, two other official Romance languages are Romansh and Catalán. Romansh is one of four official languages of Switzerland, although it is spoken by only 40,000 people. Catalán is the official language of Andorra, a tiny country of 70,000 inhabitants situated in the Pyrenees Mountains between Spain and France. Catalán is also spoken by 6 million people in eastern Spain and is the official language of Spain’s highly autonomous Catalonia province, centered on the city of Barcelona. A third Romance language, Sardinian—a mixture of Italian, Spanish, and Arabic—was once the official language of the Mediterranean island of Sardinia (Figure 5-8).

In addition to these official languages, several other Romance languages have individual literary traditions. In Italy, Ladin (not Latin) is spoken by 30,000 people living in the South Tyrol, and Friulian is spoken by 800,000 people in the northeast. Ladin and Friulian (along with the official Romansh) are dialects of Rhaeto-Romance. A Romance tongue called Ladino—a mixture of Spanish, Greek, Turkish, and Hebrew—is spoken by 100,000 Sephardic Jews, most of whom now live in Israel. None of these languages have an official status in any country, although they are used in literature.

**ORIGIN AND DIFFUSION OF ROMANCE LANGUAGES.** The Romance languages, including Spanish, Portuguese, French, Italian, and Romanian, are part of the same branch, because they all developed from Latin, the “Romans’ language.” The rise in importance of the city of Rome 2,000 years ago brought a diffusion of its Latin language.

At its height in the second century A.D., the Roman Empire extended from the Atlantic Ocean on the west to the Black Sea on the east and encompassed all lands bordering the Mediterranean Sea (the empire’s boundary is shown in Figure 6-5). As the conquering Roman armies occupied the provinces of this vast empire, they brought the Latin language with them. In the process the languages spoken by the natives of the provinces were either extinguished or suppressed in favor of the language of the conquerors.

Even during the period of the Roman Empire, Latin varied to some extent from one province to another. The empire grew over a period of several hundred years, so the Latin used in each province was based on that spoken by the Roman army at the time of occupation. The Latin spoken in each province also integrated words from the language formerly spoken in the area.

The Latin that people in the provinces learned was not the standard literary form but a spoken form, known as Vulgar Latin, from the Latin word referring to “the masses” of the populace. Vulgar Latin was introduced to the provinces by the soldiers stationed throughout the empire. For example, the literary term for “horse” was equus, from which English has derived such words as equine and equestrian. However, the Vulgar term, used by the common people, was caballus, from which are derived the modern terms for “horse” in Italian (cavallo), Spanish (caballo), Portuguese (cavalo), French (cheval), and Romanian (cal)

Following the collapse of the Roman Empire in the fifth century, communication among the former provinces declined, creating still greater regional variation in spoken Latin. By the eighth century, regions of the former empire had been isolated from each other long enough for distinct languages to evolve.
Latin persisted in parts of the former empire. People in some areas reverted to former languages, whereas others adopted the languages of conquering groups of people from the north and east who spoke Germanic and Slavic. In the past, when migrants were unable to communicate with speakers of the same language back home, major differences emerged between the languages spoken in the old and new locations, leading to the emergence of distinct, separate languages. This was the case with the migration of Latin speakers 2,000 years ago.

ROMANCE LANGUAGE DIALECTS. Distinct Romance languages did not suddenly appear in the former Roman Empire. As with other languages, they evolved over time. Numerous dialects existed within each province, many of which are still spoken today. The creation of standard national languages, such as French and Spanish, was relatively recent.

The dialect of the Ile-de-France region, known as Francien, became the standard form of French because the region included Paris, which became the capital and largest city of the country. Francien French became the country's official language in the sixteenth century, and local dialects tended to disappear as a result of the capital's longtime dominance over French political, economic, and social life.

The most important surviving dialect difference within France is between the north and the south (refer to Figure 5–8). The northern dialect is known as langue d'oïl and the southern as langue d'oc. It is worth exploring these names, for they provide insight into how languages evolve. These terms derive from different ways in which the word for “yes” was said.

One Roman term for “yes” was boc illad est, meaning “that is so.” In the south, the phrase was shortened to boc, or oc, because the /b/ sound was generally dropped, just as we drop it on the word honor today. Northerners shortened the phrase to o-il after the first sound in the first two words of the phrase, again with the initial /b/ suppressed. If the two syllables of o-il are spoken very rapidly, they are combined into a sound like the English word “wheel.” Eventually the final consonant was eliminated, as in many French words, giving a sound for “yes” like the English we, spelled in French oui.

A province where the southern dialect is spoken in southwestern France is known as Languedoc. The southern French dialect is now sometimes called Occitan, derived from the French region of Aquitaine, which in French has a similar pronunciation to Occitan. About 2 million people in southern France speak one of a number of Occitan dialects, including Auvergnat, Gascon, and Provençal.

Spain, like France, contained many dialects during the Middle Ages. One dialect, known as Castilian, arose during the ninth century in Old Castile, located in the north-central part of the country. The dialect spread southward over the next several hundred years as independent kingdoms were unified into one large country. Spain grew to its approximate present boundaries in the fifteenth century, when the Kingdom of Castile and Léon merged with the Kingdom of Aragón. At that time, Castilian became the official language for the entire country. Regional dialects, such as Aragón, Navarre, Léon, Asturias, and Santander, survived only in secluded rural areas. The official language of Spain is now called Spanish, although the term Castilian is still used in Latin America.

Spanish and Portuguese have achieved worldwide importance because of the colonial activities of their European speakers. Approximately 90 percent of the speakers of these two languages live outside Europe, mainly in Central and South America. Spanish is the official language of 18 Latin American states, whereas Portuguese is spoken in Brazil, which has as many people as all the other South American countries combined and 18 times more than Portugal itself.

These two Romance languages were diffused to the Americas by Spanish and Portuguese explorers. The division of Central and South America into Portuguese- and Spanish-speaking regions is the result of a 1493 decision by Pope Alexander VI to give the western portion of the New World to Spain and the eastern part to Portugal. The Treaty of Tordesillas, signed a year later, carried out the papal decision.

The Portuguese and Spanish languages spoken in the Western Hemisphere differ somewhat from their European versions, as is the case with English. The 46 members of the Spanish Royal Academy meet every week in a mansion in Madrid to clarify rules for the vocabulary, spelling, and pronunciation of the Spanish language around the world. The Academy’s official dictionary, published in 1992, has added hundreds of “Spanish” words that originated either in the regional dialects of Spain or the Indian languages of Latin America.

Brazil, Portugal, and several Portuguese-speaking countries in Africa agreed in 1994 to standardize the way their common language is written. Many people in Portugal are upset that the new standard language more closely resembles the Brazilian version, which eliminates most of the accent marks—such as tildes (São Paulo), cedillas (Alcobaça), circumflexes (Estância), and hyphens—and the agreement recognizes as standard thousands of words that Brazilians have added to the language.

The standardization of Portuguese is a reflection of the level of interaction that is possible in the modern world.
between groups of people who live tens of thousands of kilometers apart. Books and television programs produced in one country diffuse rapidly to other countries where the same language is used.

Difficulties arise in determining whether two languages are distinct or whether they are merely two dialects of the same language. Galician, spoken in northwestern Spain and northeastern Portugal, is as distinct from Portuguese as, say, Catalan is from Spanish. However, Catalan is generally classified as a distinct language, and Galician is classified as a dialect of Portuguese. Moldovan is the official language of Moldova yet is generally classified as a dialect of Romanian, and Flemish, the official language of northern Belgium, is generally considered a dialect of Dutch.

Several dialects of Italian are viewed as different enough to merit consideration as separate languages, according to Ethnologue. In southern Italy, the most widespread of those possible distinct languages are Napoletano-Calebrese, spoken by 7 million people, and Sicilian, spoken by 5 million. In the north, the most widespread are Lombard, spoken by 9 million people; Piemontese, spoken by 3 million; and Emiliano-Romagnolo, Liguria, and Venetian, spoken by 2 million each. Distinguishing individual languages from dialects is difficult, because many speakers choose to regard their languages as distinct.

Romance languages spoken in some former colonies can also be classified as separate languages because they differ substantially from the original introduced by European colonizers. Examples include French Créole in Haiti, Papiamentu (Creole Spanish) in Netherland Antilles (West Indies), and Portuguese Creole in Cape Verde Islands off the African coast. A créole or creolized language is defined as a language that results from the mixing of the colonizer's language with the indigenous language of the people being dominated.

A creolized language forms when the colonized group adopts the language of the dominant group but makes some changes, such as simplifying the grammar and adding words from their former language. The word créole derives from a word in several Romance languages for a slave who is born in the master's house.

**Origin and Diffusion of Indo-European**

If Germanic, Romance, Balto-Slavic, and Indo-Iranian languages are all part of the same Indo-European language family, then they must be descended from a single common ancestral language. Unfortunately, the existence of a single ancestor—which can be called Proto-Indo-European—cannot be proved with certainty, because it would have existed thousands of years before the invention of writing or recorded history.

The evidence that Proto-Indo-European once existed is "internal," derived from the physical attributes of words themselves in various Indo-European languages. For example, the words for some animals and trees in modern Indo-European languages have common roots, including *beech, oak, bear,* *deer, pheasant,* and *bee*. Because all Indo-European languages share these similar words, linguists believe the words must represent things experienced in the daily lives of the original Proto-Indo-European speakers.

In contrast, words for other features, such as *elephant, camel, rice,* and *bamboo,* have different roots in the various Indo-European languages. Such words therefore cannot be traced back to a common Proto-Indo-European ancestor and must have been added later, after the root language split into many branches.

Interestingly, individual Indo-European languages share common root words for *winter* and *snow* but not for *ocean*. Therefore, linguists conclude that original Proto-Indo-European speakers probably lived in a cold climate, or one that had a winter season, but did not come in contact with oceans.

Linguists and anthropologists generally accept that Proto-Indo-European must have existed, but they disagree on when and where the language originated and the process and routes by which it diffused. The debate over place of origin and paths of diffusion is significant, because one theory argues that language diffused primarily through warfare and conquest, whereas the other theory argues that the diffusion resulted from peaceful sharing of food.

So where did Indo-European originate? One influential hypothesis, espoused by Marija Gimbutas, is that the first Proto-Indo-European speakers were the Kurgan people, whose homeland was in the steppes near the border between present-day Russia and Kazakhstan. The earliest archaeological evidence of the Kurgans dates to around 4300 B.C.

The Kurgans were nomadic herders. Among the first to domesticate horses and cattle, they migrated in search of grasslands for their animals. This took them westward through Europe, eastward to Siberia, and southeastward to Iran and South Asia. Between 3500 and 2500 B.C., Kurgan warriors, using their domesticated horses as weapons, conquered much of Europe and South Asia (Figure 5–9).
FIGURE 5-9 Origin and diffusion of Indo-European (Kurgan hearth theory). The Kurgan homeland was north of the Caspian Sea, near the present-day border between Russia and Kazakhstan. According to this theory, the Kurgans may have infiltrated into Eastern Europe beginning around 4000 B.C. and into central Europe and southwestern Asia beginning around 2500 B.C.

Not surprisingly, scholars disagree on where and when the first speakers of Proto-Indo-European lived. Archaeologist Colin Renfrew argues that they lived 2,000 years before the Kurgans, in eastern Anatolia, part of present-day Turkey (Figure 5-10). Biologist Russell D. Gray supports the Renfrew position but dates the first speakers even earlier, at around 6700 B.C.

Renfrew believes they diffused from Anatolia westward to Greece (the origin of the Greek language branch) and from Greece westward toward Italy, Sicily, Corsica, the Mediterranean coast of France, Spain, and Portugal (the origin of the Romance language branch). From the Mediterranean coast, the speakers migrated northward toward central and northern France and on to the British Isles (perhaps the origin of the Celtic language branch).

Indo-European also diffused northward from Greece toward the Danube River (Romania) and westward to central Europe, according to Renfrew. From there the language diffused northward toward the Baltic Sea (the origin of the Germanic language branch) and eastward toward the Dniepr River near Ukraine (the origin of the Slavic language branch). From the Dniepr River, speakers migrated eastward to the Dniepr River (the homeland of the Kurgans).

The Indo-Iranian branch of the Indo-European language family originated either directly through migration from Anatolia along the south shores of the Black and Caspian seas by way of Iran and Pakistan, or indirectly by way of Russia north of the Black and Caspian seas.

Renfrew argues that Indo-European diffused into Europe and South Asia along with agricultural practices rather than by military conquest. The language triumphed because its speakers became more numerous and prosperous by growing their own food instead of relying on hunting. Regardless of how Indo-European diffused, communication was poor among different peoples, whether warriors or farmers. After many generations of complete isolation, individual groups evolved increasingly distinct languages.

FIGURE 5-10 Origin and diffusion of Indo-European (Anatolian hearth theory). Indo-European may have originated in present-day Turkey 2,000 years before the Kurgans. According to this theory, the language diffused along with agricultural innovations west into Europe and east into Asia.