Chapter 3

Phonetics and Phonemic Transcription

When we speak, we are not conscious of the sounds that make up our language(s), but in this chapter and the next you will learn that the sound systems of human language are well organized by subconscious principles that shape the content of speech sounds and their patterns of occurrence. So that you can understand the science that underlies speech sounds, we will introduce some technical concepts from the fields of anatomy and linguistics. We begin by considering the fact that all human languages can be represented in a writing system.

We take it for granted that we can write a language with discrete symbols (e.g., an alphabet). However, speech is for the most part continuous; neither the acoustic signal (the sound wave) nor the movements of the speech articulators (e.g., the tongue and lips) can be broken down into the kind of discrete units that correspond to the units represented by written symbols. For example, look at the waveform of the word *learn* in figure 3.1. (A waveform graphs changes in the amplitude of the sound wave (vertical axis) against time (horizontal axis).) Like this one, the waveforms of most speech samples have continuous patterns; clearly, the discrete symbols of written speech are not reflected in these acoustic representations.

You can observe an overlap in articulation by comparing the pronunciation of the syllables *bee*, *bah*, *boo*. You will find that when you pronounce the *b*, your tongue is already in position to pronounce the "following" vowel. Moreover, you will find that your lips are already pursed when you pronounce the *b* in *boo*, even though the pursing is part of the following vowel.

A writing system, with its set of linearly ordered discrete symbols, turns out to be an idealization of the physical instantiations of speech. So, as we begin our study of the properties of the speech sounds of language,

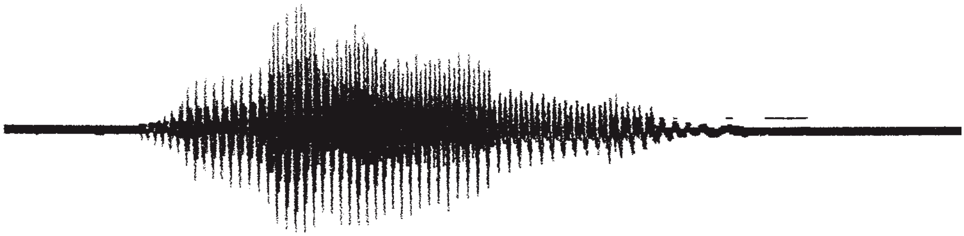


Figure 3.1
Waveform of the English word *learn*. The vertical axis displays the changes in the amplitude of the sound wave and the horizontal axis measures time.

Table 3.1
Different pronunciations of the plural morpheme

Example word	cats	dogs	bushes
Pronunciation of plural morpheme for that word	s-sound	z-sound	vowel + z

we see that what appears to be the most concrete aspect of speech—alphabetic representation—is actually highly abstract in nature.

3.1 SOME BACKGROUND CONCEPTS

Phonetics is concerned with how speech sounds are produced (articulated) in the vocal tract (a field of study known as *articulatory phonetics*), as well as with the physical properties of the speech sound waves generated by the larynx and vocal tract (a field known as *acoustic phonetics*). Whereas the term *phonetics* usually refers to the study of the articulatory and acoustic properties of sounds, the term *phonology*, the subject of chapter 4, is often used to refer to the abstract principles that govern the form and distribution of sounds in a language. In this chapter we will examine the ways in which speech sounds are produced, discussing the articulation of English speech sounds in particular. We will focus on articulation rather than on the acoustic properties of speech sounds; for further information on acoustic phonetics, see Ladefoged 2001 and Johnson 2003.

In chapter 2 we discussed the English plural morpheme -*s*. It turns out that plural nouns formed by attaching the plural morpheme, which is a suffix, do not all end with the same sound (see table 3.1). In chapter 4 we will explore a principled account of the difference, but first we must study the nature of these sounds in order to be equipped with the relevant notions and vocabulary.

Physiology of Speech Production

At its fundamental level the speech signal is a rapidly flowing series of noises that are produced inside the throat, mouth, and nasal passages and that radiate out from the mouth and sometimes the nose. One commonsense view is that learning to speak a language requires only the control of a few muscles that move the lips, jaw, and tongue. These anatomical structures are the most easily observed in any case. In reality the situation is much more complex, for over 100 muscles exercise direct and continuous control during the production of the sound waves that carry speech (Lenneberg 1967). These sound waves are produced by a complex interaction of (1) an outward flow of air from the lungs, (2) modifications of the airflow at the larynx (the Adam's apple or "voice box" in the throat), and (3) additional modifications of the airflow by the position and movement of the tongue and other anatomical structures of the vocal tract. We will consider each of these components in turn.

Airflow from the Lungs during Speech

The flow of air from the lungs during speech differs in several important respects from the airflow during quiet breathing. First, during speech, three to four times as much air is exhaled as during quiet breathing. Second, in speech the normal breathing rhythm is changed radically: inhalation is more rapid and exhalation is much more drawn out. Third, the number of breaths per unit of time decreases during speech. Fourth, the flow of air is unimpeded during quiet breathing, whereas in speech the airflow encounters resistance from the obstructions and closures that occur in the throat and mouth. While these alterations in the normal breathing pattern are occurring during speech, the function of breathing (exchange of oxygen and carbon dioxide) continues with no discomfort to the speaker.

One of the primary mechanisms for expanding the lungs during both quiet breathing and speech is the contraction of the *diaphragm* (see figure 3.2), a sheet of muscular tissue that separates the chest cavity from the abdominal region. This contraction causes the diaphragm to lower and flatten out, leading to an increase in the size of the chest cavity. The other primary mechanism for the expansion of the chest cavity is the set of muscles between the ribs in the rib cage (the *external intercostals*).

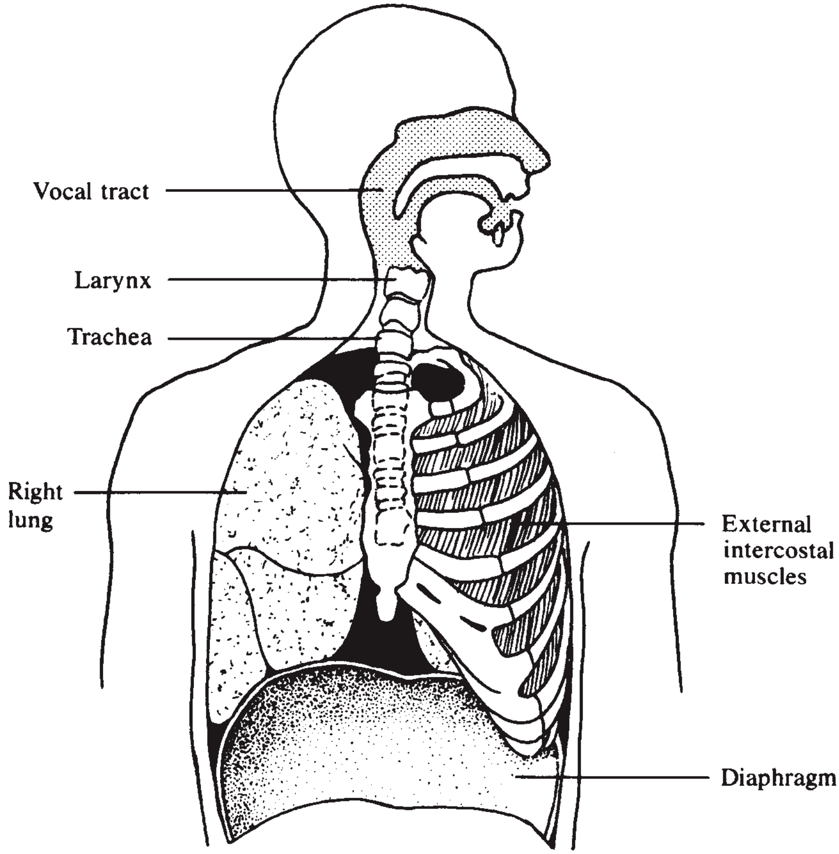


Figure 3.2
Major anatomical structures involved in the production of speech. Air driven from the lungs through the trachea and the larynx into the vocal tract is the primary source of the acoustic energy in speech. The lungs are attached to the chest wall and diaphragm, and when the diaphragm lowers, the size of the chest cavity is increased, the elastic lungs expand, and air flows inward. Similarly, air also flows inward when the muscles between the ribs (the *external intercostals*) contract and the rib cage expands outward, thus increasing the size of the chest cavity. The muscles of the diaphragm and rib cage remain active during speech, acting as a check on the outward flow of air.

Contraction of these muscles causes the ribs to lift up, and because of the way that the ribs are hinged, they swing out, increasing the volume of the chest cavity. Since the lungs are attached to the walls of the chest cavity, when the chest cavity expands, either from diaphragm contraction or from rib movement, the lungs, being elastic, also expand. As the lungs expand, air flows in, up to the point when inhalation is completed. During quiet breathing the diaphragm relaxes at this point, and the stretched lungs begin to shrink, allowing air to flow out quite rapidly at the beginning, as with air escaping from a filled balloon. During speech, however, the muscles of the diaphragm and the rib cage continue to be active, restraining the lungs from emptying too rapidly. Without this checking force, speech would be loud at first and then become quieter as the lungs emptied. Thus, humans have developed special adaptations for breathing during speech: speech is not merely ''added'' to the breathing cycle; rather, the breathing cycle is adapted to the needs of speech.

The Role of the Larynx in Speech

The first point where the airflow from the lungs encounters a controlled resistance is at the *larynx*, a structure of muscle and cartilage located at the upper end of the *trachea* (or *windpipe*) (see figure 3.2). The resistance can be controlled by the different positions and tensions in the *vocal cords* (or *vocal folds*), two muscular bands of tissue that stretch from front to back within the larynx (see figure 3.3). During quiet breathing the cords are relaxed and spread apart to allow the free flow of air to and from the lungs. During swallowing, however, the cords are drawn tightly together to keep foreign material from entering the lungs. For speech the most important feature of the vocal cords is that they can be made to vibrate if the airflow between them is sufficiently rapid and if they have the proper tension and proximity to each other. This rapid vibration is called *voicing* (or *phonation*). The *frequency* of vibration determines the perceived *pitch*. Because the vocal cords of adult males are larger in size, their frequency of vibration is relatively lower than the frequency of vibration in females and children. The pitch of adult males' voices is thus lower than that of females and children.

Voicing is the ''extra noise,'' the ''buzz'' that accompanies the production of the *z*-sound version of the plural morpheme shown in table 3.1. We say that the *z*-sound is *voiced*, whereas the *s*-sound is *voiceless*. The lack of voicing in *s* is due to the fact that the vocal cords are more spread apart and tenser than during the production of *z*, thus creating conditions that inhibit vocal cord vibration.

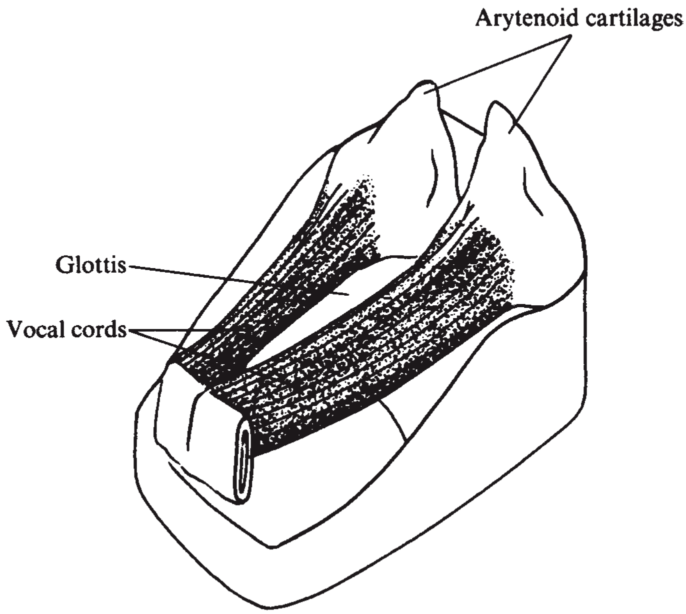


Figure 3.3
View of the vocal cords. The mechanical vibration of these cords during speech is called *voicing* (or *phonation*). The space between the cords is called the *glottis.*

Other speech sounds found in human language also require other types of vocal cord configurations and movements. We will examine some of these later in the chapter.

Speakers have a high degree of control over the sounds the vocal cords can produce. The ability to sing a melody, for example, depends on being able to change the vocal cord positions and tensions rapidly and accurately to hit the right notes. Although the ability to sing well is subject to much individual variation, the ability to control the vocal cord positions and tensions necessary for speech is well within the ability of all normal speakers.

Finally, the space between the vocal cords is called the *glottis* (see figure 3.3), and linguists frequently refer to sounds that involve a constriction or closure of this space between the vocal cords as *glottal sounds*.

The Vocal Tract

The *vocal tract*, the region above the vocal cords that includes the (oral) pharynx, the oral cavity, and the nasal cavity, is the space within which the speech sounds of human language are produced (see figure 3.4). We

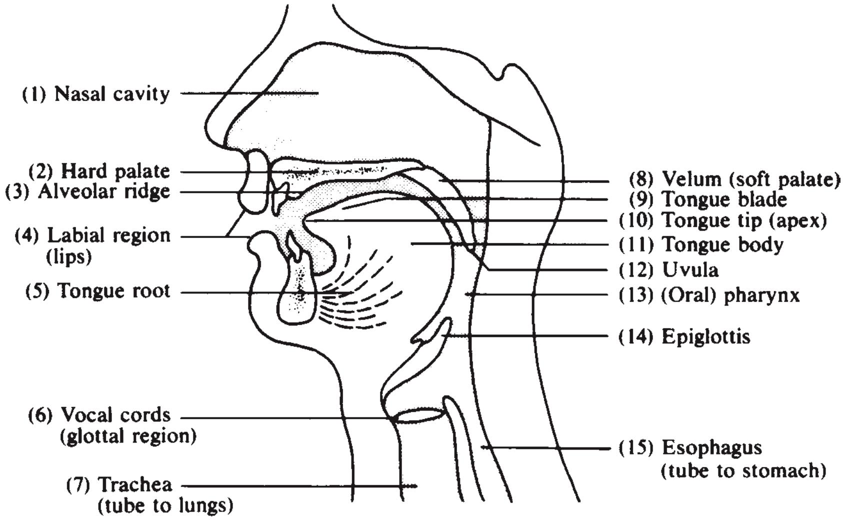


Figure 3.4
Cross section of the human vocal tract

will examine the anatomical features of the vocal tract in the course of discussing how the consonants and vowels of English are formed.

3.2 THE REPRESENTATION OF SPEECH SOUNDS

Phonemic Transcription versus English Orthography

What underlies the continuous flow of human speech is, in fact, a sequence of articulatory configurations that can be represented by a series of discrete units. The basis of the sound component of human language is a *discrete combinatorial* system that is "smeared" together in the overlapping fashion discussed earlier, much like the digital-to-analog conversion that occurs in modern electronic audio devices.

This chapter will introduce you to the discrete units (the *phonemes*) that underlie the articulation of Modern English. In discussing the sounds of English, and the sounds of human language in general, we need a set of symbols to *represent* those sounds. What sort of representational system will be most useful? If we try using the conventional English *orthography* (spelling system) to represent speech sounds, we face problems of two major types: first, a single letter of the alphabet often represents more than one sound; and conversely, a single speech sound is often represented by several different letters (see figure 3.5).

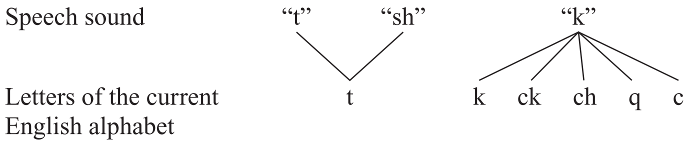


Figure 3.5
Types of inconsistencies in current English orthography. A single letter can stand for more than one sound, or several letters or groups of letters can stand for the same sound. On the left, the letter *t* represents the *t*-sound in *tin* and the *sh*-sound in *nation*. On the right, the *k*-sound is represented by the letters *k* and *ck* as in the word *kick*, *ch* as in *choir*, *q* as in *quick*, and *c* as in *cow*.

As for problems of the first type, we have already seen that the letter *s* represents a *z*-sound in the word *dogs* and an *s*-sound in the word *cats*. To take another case, the letter *t* can represent a *t*-sound, as in the word *tin*; but it can also represent a *sh*-like sound, as in *nation*.

Conversely, consider the *k*-sound in the word *kick*. This sound is orthographically represented in two different ways: the letter *k* at the beginning of the word and the letters *ck* at the end of the word. The word *cow* also begins with a *k*-sound, but here it is represented by the letter *c*. Similar problems arise with the initial sound in *jug*. This initial sound is represented by the letter *j*, but it is sometimes called ‘‘soft *g*’’ (and is spelled *g*) in words such as *giraffe*. Even the sequence of letters *dge* in words such as *ridge* and *edge* represents the *j*-sound.

In sum, English orthography is inadequate for representing the current speech sounds of American English. This lack of consistency in representing sounds is due in part to the fact that the English writing system became fixed several hundred years ago, although the pronunciation of the words has continuously changed since that period. But what system of symbols should we use to represent the speech sounds of English? More importantly, what should the symbols represent? The writing system we will now introduce uses symbols that represent for the most part the *sounds produced by particular configurations of the vocal tract*. A symbol such as *s* therefore represents the vocal tract configuration in which the tongue tip and/or blade are lightly pressed against the roof of the mouth near the teeth ridge so that when air from the lungs passes between the tongue and the teeth ridge and strikes the teeth, a hissing sound is produced.

The first writing system that we will look at is called a *phonemic transcription system*. Later we will have occasion to discuss and distinguish a *phonetic transcription system*. The crucial property of a phonemic system is that each distinctive speech sound of a language is represented with a unique symbol (or unique combination of symbols). This transcription system therefore overcomes the deficiencies of the current English alphabet. Though we will be discussing English almost exclusively, it is important to note that all human languages have a regular and consistent set of distinctive sounds that can be represented phonemically.

The Consonants of American English

Table 3.2 displays the phonemic consonant symbols of English. A *consonant* is a speech sound produced when the speaker either stops or severely constricts the airflow in the vocal tract. In addition to being classified as *voiceless* (like the *s*-sound in *cats*) or *voiced* (like the *z*-sound in *dogs*), consonants are described in terms of (1) the *place* and (2) the *manner* of their articulation. The *places* of articulation (see the top of table 3.2) are labeled in terms of anatomical structures, which (moving from the front of the mouth to the back) include the lips and regions along the roof of the mouth. In the production of most consonants, the lower lip or some part of the tongue approaches or touches the designated places of articulation along the roof of the mouth. The *manners* of articulation (see the left-hand side of table 3.2) refer for the most part to how the articulators (lips or tongue) achieve contact with or proximity to the places of articulation. We will see below that the sounds of English are highly regular in their distribution within and along the vocal tract.

We will now describe the consonants of English in terms of the framework given in table 3.2, making use of the anatomical descriptions shown in figure 3.4.

The phonemic symbols we will use here are those of the *International Phonetic Alphabet (IPA)*. We will also include in parentheses alternative symbols commonly used by many linguists. We enclose the IPA symbols in slant lines, a tradition common in linguistics when discussing phonemic symbols.

Stops

Stops are sounds produced when the airflow is completely obstructed during speech.

/p/ A voiceless bilabial stop. The speech sound symbolized by /p/ does not have accompanying vocal cord vibration and is therefore voiceless. The airflow is stopped by the complete closure of the two lips, which gives rise to the term *bilabial* (see 4, figure 3.4). The symbol /p/ represents the first sound in the word *pin*.

Table 3.2
The consonants of English

		Place of articulation						
		Bilabial	Labiodental	Interdental	Alveolar	Alveopalatal	Velar	Glottal
Manner of articulation								
Stops	voiceless	p			t		k	
	voiced	b			d		g	
Fricatives	voiceless		f	θ	s	ʃ		h
	voiced		v	ð	z	ʒ		
Affricates	voiceless					tʃ		
	voiced					dʒ		
Nasals		m			n		ŋ	
Liquids					l			
Glides		w (ʍ)			ɹ	j		

/b/ A voiced bilabial stop. The sound represented by /b/ has the same place of articulation as /p/ but is accompanied by voicing. The symbol /b/ represents the first and last sounds in the name *Bob*.

/t/ A voiceless alveolar stop. The *alveolar* consonants of English are produced when the tongue tip (or apex; see 10, figure 3.4) or blade approaches or—in the case of /t/ and /d/—touches the roof of the mouth at or near the alveolar ridge *behind* the upper teeth (see 3, figure 3.4). The English sound represented by the symbol /t/ thus differs from the *t*'s of many European languages in which the tongue tip touches the upper teeth. A Spanish /t/, for example, is a voiceless *dental* stop. The symbol /t/ represents the first sound in the English word *tin*.

/d/ A voiced alveolar stop. The sound represented by the symbol /d/ has the same place of articulation as /t/ but is accompanied by voicing. The symbol /d/ represents the first and last sounds in the word *Dad*.

/k/ A voiceless velar stop. *Velar* consonants are formed when the body of the tongue approaches or—in the case of /k/ and /g/—touches the roof of the mouth on the *palate* (the soft palate is called the *velum*; see 8, figure 3.4). The symbol /k/ represents the first sound in the word *kite*.

/g/ A voiced velar stop. The sound represented by the symbol /g/ has the same place of articulation as /k/ but is accompanied by voicing. The symbol /g/ represents the first and last sounds in the word *gag*.

Fricatives

Fricatives are sounds produced when the airflow is forced through a narrow opening in the vocal tract so that noise produced by friction is created.

/f/ A voiceless labiodental fricative. The term *labiodental* indicates that the point of contact involves the (lower) lip and the (upper) teeth. The symbol /f/ represents the first sound in the word *fish*.

/v/ A voiced labiodental fricative. The sounds represented by the symbols /f/ and /v/ differ only in voicing, /v/ being voiced. The symbol /v/ represents the first sound in the word *vine*.

/θ/ A voiceless (inter)dental fricative. Both the sound symbolized as /θ/ and its voiced counterpart /ð/ are spelled with *th* in the current English writing system. The *interdental* sounds are produced when the tongue tip is placed against the upper teeth, friction being created by air forced between the upper teeth and the tongue. For most American English speakers, the tongue tip is projected slightly when it rests between the upper and lower teeth. The symbol /θ/ represents the first sound in its own name, the Greek letter *theta*, and in the word *thin*.

/ð/ A voiced interdental fricative. The symbol /ð/ is called *eth* (or *crossed d*). You can hear the difference between the sounds symbolized by /ð/ and /θ/ if you say *then* and *thin* very slowly. You will hear (and feel) the voicing that accompanies the /ð/ at the beginning of *then*, and you will note that the initial consonant of *thin* is not voiced. The symbol /ð/ also represents the first sound in the words *this* and *that*.

/s/ A voiceless alveolar fricative. Note that the fricative sound represented by the symbol /s/ is much harsher than the fricative sound represented by the symbol /θ/. The turbulence for /s/ is created by air passing between either the tongue tip or blade (for some English speakers) and the alveolar ridge, which then strikes the teeth at a high velocity. The symbol /s/ represents the first sound in the word *sit*.

/z/ A voiced alveolar fricative. The sounds represented by /s/ and /z/ differ only in voicing, /z/ being voiced. The symbol /z/ represents the first sound in the name *Zeke*.

/ʃ/ (/š/) A voiceless alveopalatal fricative. The symbol /ʃ/, usually spelled *sh* in English orthography, represents a fricative similar to /s/, but the region of turbulent airflow lies along the sides of the tongue and just behind the alveolar ridge on the hard palate (hence the term *alveopalatal*; see 2 and 3, figure 3.4). During the articulation of /ʃ/ the tongue tip can be positioned either near the alveolar ridge itself (with the tongue blade arched) or just behind the alveolar ridge (in which case the tongue blade does not need to be arched). The symbol /ʃ/ represents the first sound in the word *ship*.

/ʒ/ (/ž/) A voiced alveopalatal fricative. Unlike /ʃ/, the voiced counterpart /ʒ/ is rare. The symbol /ʒ/ represents the first sound in foreign names such as *Zsa-Zsa* or *Jacques*, but no native English words begin with /ʒ/. More commonly, /ʒ/ occurs in the middle of English words. For example, the letter *s* in *decision* and *measure* is pronounced as the sound represented by /ʒ/.

/h/ A voiceless "glottal" fricative. The /h/ sound is often called a *glottal* fricative because the vocal cords are positioned so that a small amount of turbulent airflow is produced across the glottis. However, the primary noise source for this speech sound is turbulence created at different points along the vocal tract where the tongue body (or blade) approaches the roof of the mouth. The point where the friction is created is determined by the vowel that follows the /h/. In the articulation of the English word *heap*, for example, the tongue body is positioned high and forward, and the fricative noise is produced in the palatal region. The symbol /h/ represents the first sound in the words *how* and *here*.

Affricates

An *affricate* is a single but complex sound, beginning as a stop but releasing secondarily into a fricative.

/tʃ/ (/č/) A voiceless alveopalatal affricate. The symbol /tʃ/ represents the first sound in the word *chip* (/tʃ/ is usually spelled as *ch*). In articulating this sound, the tongue makes contact at the same point on the roof of the mouth as in the articulation of the sound represented by /ʃ/. Unlike /ʃ/, though, /tʃ/ begins with a complete blockage of the vocal tract (a stop), but then is immediately released into a fricative sound like /ʃ/.

/dʒ/ (/ǰ/) A voiced alveopalatal affricate. The sounds represented by the symbols /tʃ/ and /dʒ/ differ only in voicing, /dʒ/ being voiced. The symbol /dʒ/ represents the first and last sounds of the word *judge* (/dʒ/ being spelled as both *j* and *dge*, in this case).

Nasals

In English the *nasals* are voiced oral stops, similar to the voiced stops discussed above in that they are voiced and are produced with a complete obstruction in the oral cavity. With nasals, however, the airflow and sound energy are channeled into the nasal passages (see 1, figure 3.4), due to the lowering of the velum (see 8, figure 3.4).

/m/ A bilabial nasal. The sounds represented by the symbols /m/ and /b/ are articulated in the same manner, except that for /m/ the velum is lowered to allow airflow and sound energy into the nasal passages. The symbol /m/ represents the first sound in the word *mice*.

/n/ An alveolar nasal. The sound represented by the symbol /n/ is articulated in the same position as /d/, with the velum lowered. The symbol /n/ represents the first sound in the word *nice*.

/ŋ/ A velar nasal. The symbol /ŋ/ is called *eng* (or even *engma* or *engwa*) and represents the final sound in the word *sing*. The normal English spelling for this single sound is *ng*. In order to hear the sound—and to hear that it *is* only one sound—compare the words *finger* and *singer*. For most speakers of American English the middle consonants of the word *finger* consist of a sequence of the velar nasal /ŋ/ followed by the velar stop /g/. In *singer*, however, only the velar nasal /ŋ/ occurs as the middle consonant, with no following /g/. Similarly, the word *long* ends only in a single consonant, the velar nasal. Note, however, the existence of a dialectal pronunciation of the word *long* in the expression *Long Island*. Certain speakers from the New York City area actually pronounce the final /g/ (*Long Island* = *LonGisland*).

The "*g*-like" quality of /ŋ/ is due to its being articulated in the same way as /g/, except that the velum is lowered. Thus, just as /m/ and /n/ are the nasal counterparts of /b/ and /d/, so /ŋ/ is the nasal counterpart of /g/. The sound represented by the symbol /ŋ/ does not occur in initial position in English words, but only in medial and final positions, as our examples show. A single velar nasal /ŋ/, spelled *Ng* in the United States, is a common surname in Cantonese.

Finally, although English orthography sometimes uses a *digraph* (a combination of two letters) to represent /ŋ/ (namely, *ng*), it should be stressed once again that the velar nasal is a *single* speech sound. Similarly, recall that other consonant sounds of English are represented by two-letter sequences in the current spelling system: *th* for /θ/ and /ð/, *sh* for /ʃ/, and *ch* for /tʃ/. Yet each of these consonants—/ŋ/, /θ/, /ð/, /ʃ/, and /tʃ/—is a single speech sound.

Liquids

Liquid sounds are found in the overwhelming majority of the world's languages, and English has one: /l/. The term *liquid* is a nontechnical, impressionistic expression indicating that the sound is "smooth" and "flows easily." Liquids share properties of both consonants and vowels: as in the articulation of certain consonants, the tongue blade is raised toward the alveolar ridge; as in the articulation of vowels, air is allowed to pass through the oral cavity without friction.

/l/ An alveolar liquid. In the articulation of English /l/, the tongue blade is raised and the apex makes contact with the alveolar ridge. The sides of the tongue are lowered, permitting the air and sound energy to flow outward. The symbol /l/ represents the first sound in the word *life*.

Glides

Glides are vowel-like articulations that precede and follow true vowels. The term *glide* is based on the observation that the sequence of a glide and a vowel is a smooth, continuous gesture. Because the tongue position in articulating the glides /j/ and /w/ is similar to the tongue position of the vowels in *beet* and *boot*, respectively, these glides are sometimes referred to as *semivowels*.

/w/ A bilabial (velarized) glide. The sound represented by the symbol /w/ is formed with the body of the tongue arched in a high, back position, toward the soft palate (velum). Lip rounding also accompanies the production of this sound. The symbol /w/ represents the first sound in the word *wood*.

/ʍ/ A bilabial (velarized) glide (with a voiceless beginning). Some speakers of English have different initial sounds in the words *which* and *witch*. For these speakers the initial sound in *which* begins as a voiceless sound, followed immediately by the glide /w/. Some linguists write this initial sound as the digraph /hw/.

/ɹ/ An alveolar glide. American English /ɹ/ is produced with a tongue blade that is raised toward the alveolar ridge. Many speakers also curl the apex into a *retroflexed* position (curled upward and backward). Others press the tongue tip against the lower gum (below the teeth) and raise the blade of the tongue toward the roof of the mouth. This sound is also produced with lip rounding (a pursing of the lips) and a retraction of the tongue root (see 5, figure 3.4). The symbol /ɹ/ represents the first sound in the word *red*.

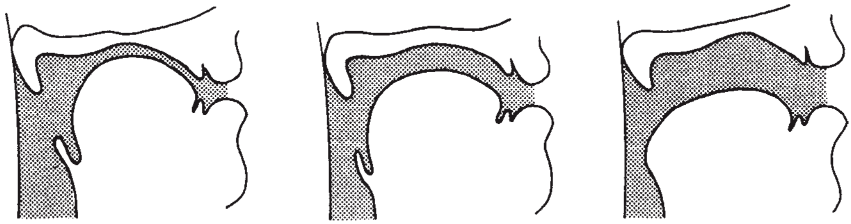
We are following IPA conventions in using the "upside-down *r*" symbol for this English phoneme. The "right-side-up *r*" symbol is reserved for trilled *r*, a sound found in dialects of Scottish English.

Arguments supporting the glide status of /ɹ/ are found in Kahn 1976.

/j/ (/y/) An alveopalatal glide. The sound represented by the symbol /j/ is formed with the body and the blade of the tongue arched in a high, front position, toward the hard palate. The symbol /j/ represents the first sound in the word *yes*.

The Vowels of American English

Whereas consonants are formed by obstructions—either partial or total—in the vocal tract, vowels are produced with a relatively open vocal tract, which functions as a resonating chamber. The different vowels are formed by the different *shapes* of the open, resonating vocal tract, and the variety of shapes is determined by the position of several anatomical structures: the position of the tongue body and blade, the relative opening of the lips, the relative opening of the oral pharynx (see 13, figure 3.4), and the position of the jaw (see figure 3.6). Although these articulators are, to some extent, anatomically connected, they can be independently controlled to produce the different vowels.



/ɪ/ "bit" /ɛ/ "get" /ɑ/ "hot"

Figure 3.6
Vocal tract shapes for given English vowels

	Front	Central	Back
High	ɪ	(ɨ)	ʊ
Mid		ʌ (ə)	
	ɛ	(ɝ, ɚ)*	ɔ
Low	æ		ɑ

Figure 3.7
Lax (short) vowels and reduced vowels of American English. *(These *r*-colored vowels are discussed in "Special Topics: Vowels before /ɹ/.")

There are three major types of vowels in American English: *lax* (or *short*), *tense* (or *long*), and *reduced*. As the labels suggest, the lax vowels are produced with somewhat less muscular tension than the tense ones and are also somewhat shorter in duration. The reduced vowels could equally well be called the *unstressed* vowels, a point we return to below.

Lax (Short) Vowels

The symbols for the English lax vowels are displayed in figure 3.7. If we imagine this figure superimposed on a cross section of the vocal tract (such as that depicted in figure 3.4), then the positions of the vowels in the chart represent the relative positions of the part of the tongue closest to the roof of the mouth (assume the mouth opening to be on the left, as in figure 3.4). We can simplify our description of the articulation of vowels by limiting our discussion to this relative position of the highest part of the tongue during vowel production.

/ɪ/ A lax high front vowel. The terms *high* and *front* describe the position of the tongue in the mouth (see figure 3.6). The symbol /ɪ/ represents the vowel sound in the words *bit* /bɪt/ and *wish* /wɪʃ/.

/ɛ/ A lax mid front vowel. The tongue body is relatively forward, as in the production of /ɪ/, but it is slightly lower (see figure 3.6). The symbol /ɛ/ represents the vowel sound in the words *get* /gɛt/ and *mess* /mɛs/.

/æ/ A lax low front vowel. This vowel (and the symbol for it) is called *ash* by many linguists, and the symbol /æ/ represents, in fact, the vowel sound in the word *ash* /æʃ/. It is produced with a front tongue body and with a lowered tongue body and jaw.

/ʊ/ A lax high back vowel. The vowel sound represented by the symbol /ʊ/ is found in words such as *put* /pʊt/ and *foot* /fʊt/. As you start to pronounce the vowel /ʊ/, you can feel your tongue move back and upward toward the velum. You can also feel your lips become rounded (pursed and brought closer together) during the production of this vowel; hence, it is called a *rounded* vowel.

/ʌ/ A lax mid back vowel. The vowel sound represented by the symbol /ʌ/, sometimes called *wedge*, occurs in words such as *putt* /pʌt/ and *luck* /lʌk/. Note that the words *put* and *putt*, which differ in the number of final *t*'s in the English spelling system, actually differ in their vowels, /ʊ/ versus /ʌ/, respectively.

/ɑ/ A lax low back vowel. The position of the tongue is low and retracted in the articulation of the vowel /ɑ/ (see figure 3.6). There are several varieties of /ɑ/-like vowels in English; these vowels constitute one of the most difficult aspects of the study of English vowel sounds. The difficulty is due in part to the fact that there is considerable dialectal variation in the pronunciation of these vowels. We leave it to your instructor to help you assign the appropriate symbols to represent vowels of your own speech or of the English spoken in your area. The vowel sound represented by the symbol /ɑ/ (*script-a*) is the low back vowel shared by most speakers of American English. It is typically found in words such as *hot* /hɑt/ and *pot* /pɑt/.

Notice that the symbol representing this vowel looks more like an italicized *a* than like a roman-style "a."

/ɔ/ A lax low back (rounded) vowel. If you pronounce the words *cot* and *caught* differently, you probably have the vowel /ɔ/ in your pronunciation of *caught*. There is minor lip rounding in the articulation of this vowel.

For many (if not most) speakers of American English the pronunciation of the vowels in the words *father*, *froth*, and *fraught* will be the same. However, you may speak a dialect (e.g., if you are a speaker of some dialects of British English) in which the vowels in the three words may all be different.

/ɝ/. This *r*-colored vowel is discussed in “Special Topics: Vowels before /ɹ/.”

Reduced Vowels

There are three so-called *reduced* vowels in English, shown in parentheses in figure 3.7.

/ə/ The most common reduced vowel is called *schwa*, a mid back vowel whose symbol is an upside-down and reversed *e*. It is the last vowel sound in the word *sofa* and sounds very much like the lax vowel represented by the symbol /ʌ/ (some linguists, in fact, use the same symbol for both of these sounds). Schwa /ə/ is called a *reduced* vowel because it is frequently an unstressed variant of a stressed (accented) vowel. Note how the accented vowel /ɛ/ in the base word *democrat* /dɛ̲́məkɹæt/ “reduces” or “corresponds” to the unaccented vowel /ə/ in the derived word *democracy* /də̲mɑ́kɹəsi/. Likewise, the vowel /æ/ in *democrat* /dɛ́məkɹæ̲t/ “reduces” or “corresponds” to the second schwa in *democracy* /dəmɑ́kɹə̲si/.

/ɨ/ Another reduced vowel of English is a high back vowel referred to as *barred-i*. It is typically the vowel sound in the second syllable of *chicken* /tʃɪkɨn/. Like /ə/, the vowel /ɨ/ occurs only in unstressed (unaccented) syllables in a word.

/ɚ/ The final reduced vowel of English is discussed in “Special Topics: Vowels before /ɹ/.”

There is considerable variation in the pronunciation of the two vowels /ə/ and /ɨ/. Most likely, English has only one basic reduced vowel, and the appearance of one or the other is determined by the surrounding phonetic environment. In chapter 4 we will discuss the reduced vowel and some properties of English words that account for its distribution.

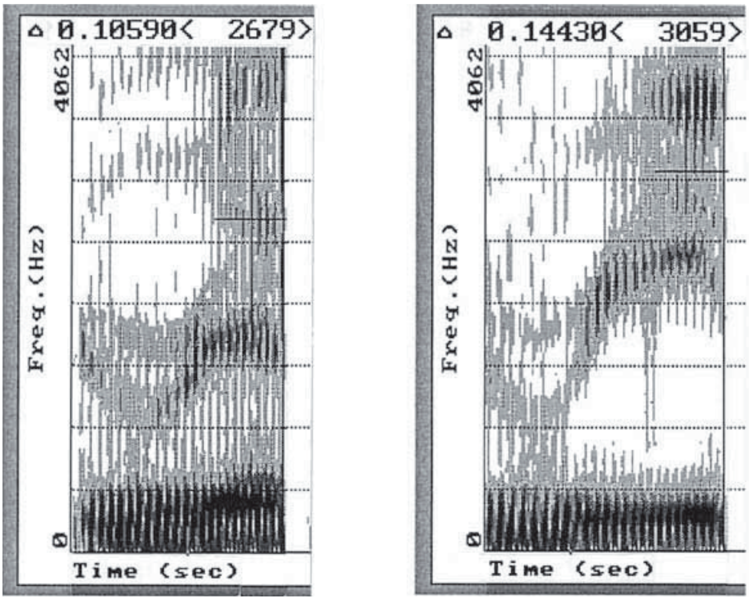
Tense (Long) Vowels and Diphthongs

In addition to its inventory of short and reduced vowels, English has a set of tense vowels (see figure 3.8). The tense vowels are all relatively longer than the lax vowels, and all tense vowels in Standard English end with the tongue body high in the mouth. Tense vowels also sound higher than lax vowels. For example, spectrographic representations (figure 3.9) reveal that the tense vowel /i/ in *reed* is 38 milliseconds longer than the lax vowel /ɪ/ in *rid*; moreover, the second resonant frequency (formant) of

	Front	Back
High	i	u
Mid	eɪ	oʊ, ɔɪ
Low	(æʊ)	aʊ, aɪ

(a)

Figure 3.8
Tense (long) vowels and diphthongs of American English



(a) “rid” (b) “reed”

Figure 3.9
Spectrograms representing the lax vowel /ɪ/ of *rid* (a) and the tense vowel /i/ of *reed* (b). “Δ” marks the vowel’s duration: 106 milliseconds (a) and 144 milliseconds (b). Thus, the tense vowel represented here is 38 milliseconds longer than the lax one, a pattern typical of the length difference between tense and lax vowels. The number in angle brackets is the value of the second formant for these vowels. The higher value for /i/ in *reed* (b) reflects a more advanced tongue position, another characteristic of tense vowels.

/i/ is higher than that of /ɪ/, an acoustic property that corresponds to a more advanced tongue position.

/i/ A tense high front vowel. The symbol /i/ represents the vowel sound in words such as *bead* /bid/ and *three* /θɹi/.

/eɪ/ (alternative IPA transcription /e/; alternative American transcription /ey/) A tense mid front vowel (with an accompanying high front offglide). This high front offglide is represented in the IPA transcription with the symbol /ɪ/. The vowel is found in words such as *clay* /kleɪ/ and *weigh* /weɪ/.

/u/ A tense high back (rounded) vowel. This transcription represents the vowel sound in words such as *crude* /kɹud/ and *shoe* /ʃu/.

/oʊ/ (alternative IPA transcription /o/; alternative American transcription /ow/) A tense mid back (rounded) vowel (with an accompanying high back offglide). This high back offglide is represented in the IPA transcription with the symbol /ʊ/. This transcription represents the vowel sound in the words *boat* /boʊt/ and *toe* /toʊ/.

Diphthongs in English are single vowel sounds that begin in one vowel position and glide into another vowel position. Strictly speaking, the vowels /eɪ/ and /oʊ/ are diphthongs, although they have been traditionally classified with the long vowels /i/ and /u/. The following three vowels are unambiguously diphthongs that have substantial tongue movement in their articulation.

/ɔɪ/ (alternative American transcription /oy/) A tense mid back (rounded) vowel (with an accompanying high front offglide). This transcription represents the vowel sound in words such as *boy* /bɔɪ/ and *Floyd* /flɔɪd/.

/aʊ/ (alternative American transcription /aw/) A tense low back vowel (with an accompanying high back offglide). This transcription represents the vowel sound in the words *cow* /kaʊ/ and *blouse* /blaʊs/. In some dialects of American English this diphthong begins with a low front vowel and should be transcribed as /æʊ/.

/aɪ/ (alternative American transcription /ay/) A tense low back vowel (with an accompanying high front offglide). This transcription represents the vowel sound in words such as *my* /maɪ/ and *thigh* /θaɪ/.

East Coast Dialectal Variant

/a/ A tense low vowel. The vowel sound represented by the symbol /a/ (*printed-a*) is found—among other places—in the speech of New England, especially in Maine and eastern Massachusetts. One characteristic expression of the Boston area, "Park the car," contains two instances of the vowel represented by the symbol /a/.

To conclude our discussion of vowels, we point out that one of the reasons that speakers of English have some difficulty in pronouncing the vowels of languages such as Spanish and Italian is that most of the tense (long) vowels of English are diphthongs, whereas the corresponding vowels in Spanish and Italian are not. For example, a native speaker of American English who is learning Italian is likely to pronounce the word *solo* "alone" with two English *o*'s, as shown most clearly in the IPA transcription /soʊloʊ/. For this reason, teachers of foreign languages often tell American-English-speaking students to use "pure" vowels—that is, ones without velar offglides—in words such as Italian *solo*.

Consonants and Vowels in Other Languages

All spoken human languages have sound systems made up of consonants and vowels. Nevertheless, languages vary greatly in the number of these sound types. Ignoring dialectal differences, American English has 39 phonemes (24 consonants and 15 vowels); Hawaiian has 13 phonemes (8 consonants and 5 vowels); and Khoisan, an African language that has click consonants, has 119 phonemes (95 consonants and 24 vowels). All of these languages function successfully as communication systems in spite of their extremely different numbers of speech sounds.

Also despite numerical differences, the vowels found in the world's languages are often quite similar and are produced in similar portions of the mouth. All languages have an /ɑ/-like vowel, and *i*'s and *u*'s are found in the majority of languages. The vowels *a*, *i*, and *u*, being produced at the periphery of the vocal tract, are the maximally distinct vowels. Consonants are subject to more crosslinguistic variation because languages have more consonants than vowels. Nevertheless, languages share a common core of consonant types. Almost all languages have labial stops (such as *p* and *b*), dental/alveolar stops (such as *t* and *d*), and velar stops (such as *k* and *g*), one or more of the nasals (*m* or *n*), a liquid (*r* or *l*), and some kind of fricative (typically an *s*-like sound).

A group of sounds that may be unfamiliar to speakers of English and of European and Asian languages are the so-called click sounds found in several African languages. In the production of clicks, the tongue makes a closure with the roof of the mouth not just at one point, but at two points (both at the velum and at one other point farther forward). The primary airflow is created by making the sealed-off space larger, creating a partial vacuum, usually by lowering the tongue and jaw. When the front stoppage is released and air rushes into the partial vacuum, a click sound results. Some click sounds are made by English speakers, and although they are not part of the English language itself, they are still used for communication. The sound that is written *tsk! tsk! tsk!* is not to be pronounced "tisk, tisk, tisk." The *tsk!* is a single click sound made with air rushing in between the tip of the tongue and the alveolar ridge. In the African language Xhosa, spoken by Nelson Mandela, certain "click" phonemes are an integral part of the consonant system. The click consonant that appears at the beginning of the language name *Xhosa*—a click with a lateral release—is the sound that some people use to signal a horse to "giddy-up." Try pronouncing this lateral click and following it immediately with the sequence -*osa*. If you can do this, you will come very close to pronouncing the name of this language correctly. The official IPA representation for this sequence is /ǁosa/.

The Form of the English Plural Rule: Three Hypotheses

Now that we have a set of symbols that permit us to transcribe the consonant and vowel sounds of English in a precise way, we can reformulate table 3.1, more accurately, as table 3.3. Here the plural morpheme can appear as either /s/, /z/, or /ɨz/.

Even though we can now represent the different pronunciations of the plural morpheme, we are still left with accounting for the *distribution* (pattern of occurrence) of the different plural forms. What factors govern, or predict, this distribution? We will pursue this problem by formulating several hypotheses, which we will then test and revise in light of new data.

A given noun can be associated with only one of the three different forms of the plural. Thus, for example, the plural /ɨz/ that is associated with *bush* to make *bushes* cannot be associated with *cat* or *dog*. The result

Table 3.3
Phonemic transcription of different forms of the plural morpheme

Example word	cat<u>s</u>	dog<u>s</u>	bush<u>es</u>
Phonemic transcription of plural morpheme for that word	/s/	/z/	/ɨz/
Phonemic transcription of that word	/kæts/	/dɑgz/	/bʊʃɨz/

of doing so (/kætɨz/, /dɑgɨz/) sounds "foreign" to a native speaker of English. Thus, there must be some principle governing the occurrence of the different plural shapes. One account for the plural distribution would be to say that the form of the plural morpheme to be used with any given noun is unpredictable, and that we must simply list, for each individual noun of the language, which form it takes. This would amount to saying that speakers of English have simply memorized the phonological form of the plural for each individual noun. The distribution of the forms of the plural would then be given by sets of statements such as the following:

(1)
Hypothesis 1 (Listing of words)
{kæt, kæts} "cat"
{mæp, mæps} "map"
{bæk, bæks} "back"
{dɑg, dɑgz} "dog"
{kæn, kænz} "can"
{tæb, tæbz} "tab"
{bʊʃ, bʊʃɨz} "bush"
{dɪʃ, dɪʃɨz} "dish"
{ɹɪdʒ, ɹɪdʒɨz} "ridge"
and so forth

Hypothesis 1 is consistent with the fact that there are nouns such as *child*, *ox*, *sheep*, and *man* for which the shape of the plural ending does seem to be determined by the word itself. However, hypothesis 1 implies that for any new word (not already found in our lists) we will not be able to predict which of the three forms of the plural morpheme it will take. But this is clearly false. Speakers of English can spontaneously and with consensus form the plural for nouns they have never heard before and therefore could not have memorized. We may never have heard the noun *glark* before (since it is a nonsense word), yet we can indeed predict that the form of the plural would be /s/ and not /z/ or /ɨz/; in fact; it seems that every noun that ends in /k/ takes the plural form /s/, whether it is a nonsense word or not. Similarly, every noun that ends in /g/, such as *dog*, takes the plural form /z/; and every noun that ends in /ʃ/, such as *bush*, takes the plural form /ɨz/. It is, in fact, possible to group the nouns that take only /s/ or only /z/ or only /ɨz/ in terms of their last sound. This leads us to a second hypothesis about the distribution of the different forms of the plural morpheme:

(2)
Hypothesis 2 (Listing of final sounds)
The forms of the plural morpheme are distributed according to the following speech sound lists:
a. The plural morpheme takes the form /s/ if the noun ends in /p, t, k, f, or θ/.
b. The plural morpheme takes the form /z/ if the noun ends in /b, m, d, n, g, ŋ, v, ð, l, ɹ, w, j/, or any vowel.
c. The plural morpheme takes the form /ɨz/ if the noun ends in /s, z, ʃ, ʒ, tʃ, or dʒ/.

Notice that hypothesis 2 now reflects a native English speaker's judgments concerning the form that the plural will take for any new word. Accordingly, the task faced by the language learner in learning the distribution of the plural forms is different under hypothesis 2 than under hypothesis 1. That is, language learners do not memorize the particular plural form for every noun; rather, it appears that they acquire a rule to determine what plural form is associated with a particular noun (in terms of its final sound). Of course, there are still nouns whose plural form has to be memorized, as with the exceptional nouns *children*, *oxen*, *sheep*, *men*, and so forth. We can say, then, that there are nouns whose plural follows hypothesis 1 (the exceptional nouns), but the overwhelming majority are subject to hypothesis 2.

To see that hypothesis 2 is still not sufficient to handle all cases of plural formation, we turn to cases in which foreign words are made to undergo English plural formation—in particular, foreign words that contain speech sounds not found in English. Some English speakers, especially announcers on radio stations that play classical music, pronounce the name of the German composer Bach as it is pronounced in German, with a final voiceless velar fricative. This sound, symbolized as /x/, is not part of the English phonemic system. If these English speakers use the name *Bach* (/bɑx/) in the plural, perhaps in referring to two generations of Bachs, it takes /s/ and not /z/ or /ɨz/ (*Bachs* = /bɑxs/). The problem is that the sound /x/ does not appear in the list in hypothesis 2. We therefore need to develop a new hypothesis that reflects the English speaker's ability to assign plurals to words that end in sounds that are foreign to English.

If we compare words that end in, say, /f/ (which take the plural form /s/) and words that end in /v/ (which take the form /z/), we can observe that /f/ and /v/ represent similar sounds that differ only in a single feature—namely, /f/ is *voiceless*, whereas /v/ is *voiced*. Further, words with the final consonant /k/ (which is voiceless) take the plural /s/, whereas words with a final /g/ (which is voiced) take the plural /z/. If we set aside for a moment the nouns that take /ɨz/, we can make the following observation: if a noun ends with a voiceless sound, then it will take the voiceless plural form /s/; but if it ends with a voiced sound, then it will take the voiced plural form /z/. Notice that we now have an account for why hypothesis 2 groups nouns ending in *vowels* with nouns ending in *voiced consonants* such as /b, d, m/ (see hypothesis 2, part (b)): those final sounds are all voiced, and so it follows automatically that all nouns ending in voiced sounds will take the plural form /z/.

Let us now return to the nouns that take the plural form /ɨz/. We note that the final consonants of these nouns (/s, z, ʃ, ʒ, tʃ, or dʒ/) are either *alveolar fricatives*, *alveopalatal fricatives*, or *alveopalatal affricates*.

(3)
Hypothesis 3 (Use of phonetic features)
The forms of the plural morpheme are distributed according to the following conditions:
a. The plural morpheme takes the form /ɨz/ if the last sound in the noun to which it attaches is an alveolar fricative, an alveopalatal fricative, or an alveopalatal affricate.
Otherwise:
b. The plural morpheme takes the voiced form /z/ if the last sound in the noun is voiced.
c. The plural morpheme takes the voiceless form /s/ if the last sound in the noun is voiceless.

English plural formation demonstrates the interaction of two parts of English grammar, where the concept of grammar includes morphology and phonology as well as syntax. English grammar includes a morphological part that specifies that plurals are formed by adding a suffix to nouns, and a phonological part containing rules that determine the actual phonetic shape (or shapes) of that suffix. Linguists hypothesize that grammars of all languages contain a morphological component in which morphemes are combined to form complex or compound words. In the world's languages combinations of morphemes are often subject to phonological rules that determine the ultimate shape of underlying morphemes, both stems, bases, and affixes.

The phonological form of some affixes is invariant. Such a case seems to be the prefix *re*-, which is pronounced /ɹi/ regardless of the phonological shape of the verb to which it is attached. Other affixes may be subject to phonological rules that specify their phonological shape depending on their phonological environment. The English plural morpheme is one of these. Other examples of shape-changing rules are given in the exercises at the end of this chapter and in *A Linguistic Workbook* (Farmer and Demers 2010).

Phonetic Variations on a Phonemic Theme

So far we have assumed that the sounds represented by the phonemic transcription system of English are articulated the same way each time they are produced. This assumption ignores an important aspect of the pronunciation of some phonemes. We discuss below several examples of variation in the pronunciation of certain American English consonants, variations that are common to most speakers of American English.

Types of /t/ in English

Aspirated t. When the sound /t/ occurs at the beginning of a syllable, its pronunciation is accompanied by a puff of air called *aspiration*. You can observe the presence of aspiration if you hold a thin, flexible piece of paper close to the front of your mouth when you say the word *tin*. The paper will flutter immediately after the /t/ is pronounced. You can also place your hand in front of your mouth to feel this puff of air. In contrast, the pronunciation of the /t/'s in the word *stint* is unaspirated; pronouncing these /t/'s will not cause the piece of paper to flutter. Later we will discuss the general conditions under which some English phonemes are aspirated.

In order to represent more detailed aspects of pronunciation (such as aspiration), linguists use a system called (close) *phonetic transcription*. By convention, phonetic symbols are enclosed in square brackets [ ]; the symbols of the more general transcription system we have been using—which, when it satisfies conditions to be discussed below, is called a *phonemic transcription*—are enclosed in slant lines / /. For example, in phonetic transcription *tin* and *stint* are represented as [tʰɪn] and [stɪnt], respectively (where a superscripted *h* indicates an aspirated sound and its absence indicates an unaspirated sound). In phonemic transcription they are represented as /tɪn/ and /stɪnt/. We will discuss the difference between phonetic and phonemic transcriptions after we have discussed some of the finer phonetic details of American English speech.

Unreleased t. Final /t/ in words such as *kit* is frequently unreleased in the pronunciation of many speakers of American English: the tongue touches the alveolar ridge but does not immediately drop away to "release" the sound. (In contrast, in most American English dialects the pronunciation of the final stop /t/ in words such as *fast* is in fact released.) For most speakers of American English, in the pronunciation of the word *kit*, the voicing ends and the airflow stops *before* the tongue reaches the alveolar ridge in articulating the final /t/. Where and how is the airflow stopped in this case? The primary stop articulation in the pronunciation of final /t/ in words such as *kit* occurs in the larynx, rather than in the region of the alveolar ridge, even though the tongue tip does indeed make contact with the alveolar ridge immediately *after* the closure of the vocal cords. Recall that the *glottis* is the space between the vocal cords, and a stop created by closure at the glottis is called a *glottal stop*, represented as the symbol [ʔ]. A glottal stop appears at the beginning of each of the two *oh*'s of the expression *oh*-*oh!*, which we can phonetically transcribe as [ʔʌʔoʊ] or [ʔoʊʔoʊ]. An unreleased /t/ that is produced with a glottal stop immediately preceding the alveolar articulation is symbolized as [ˀt]. Such sounds are sometimes referred to as *preglottalized*. Thus, the characteristic pronunciation of the word *kit* for most American English dialects is represented phonetically as [kʰɪˀt].

Glottal stop replacement of t. In certain words the tendency to have a glottal closure with the articulation of /t/ in certain environments reaches such an extreme that the glottal stop actually replaces /t/. In many speakers' pronunciation of words such as *button* and *kitten*, the stop articulation is actually carried out at the glottis, and the tongue does not, in fact, move toward the alveolar ridge until the /n/ of the final syllable is articulated. The /t/ is generally replaced by the glottal stop if the following syllable contains a *syllabic* /n/. The term *syllabic* here refers to the fact that nasal consonants (such as /n/) can function as syllables by themselves, without an accompanying vowel. In the word *button*, for example, the only sound in the second syllable is the nasal [n]—there is no true vowel at all in that syllable. A syllabic /n/ is indicated by placing a straight apostrophe (or tick mark) under the symbol: [n̩]. The phonetic transcription of *kitten* would thus be [kʰɪʔn̩].

Flapped t. In words such as *pitted*, /t/ is regularly pronounced as a voiced "*d*-like" sound by most speakers of American (but not British) English. This sound is articulated by making a quick "tap" with the tongue tip on the alveolar ridge. Because of the rapidity of the articulation of this sound, it is referred to as a *flap* (or a *tap*), transcribed phonetically with the symbol [ɾ]. Thus, a word such as *pitted* is phonetically transcribed as [pʰɪɾɨd]. The flap [ɾ] is always voiced and occurs primarily intervocalically (between vowels).

Alveopalatal t. Children who are learning to write English sometimes spell the word *truck* as *chruk* or *chuk*. In doing so, they reveal that they are quite good phoneticians. What they are noticing is that the /t/ in the word *truck* is pronounced much farther back along the roof of the mouth than is the regular /t/. For many speakers, in fact, the tongue tip touches *behind* the alveolar ridge, at exactly the point where the /tʃ/ phoneme is produced. Moreover, the /ɹ/ phoneme in many dialects is voiceless following /t/ and sounds similar to /ʃ/. Since the combination of the alveopalatal stop followed by the alveopalatal "fricative" (the voiceless *r*) sounds like the /tʃ/ phoneme, it is understandable that children might spell initial *tr* sequences as *ch*. Linguists transcribe this phonetic realization of /t/ as [t̠].

Retraction of an alveolar sound under the influence of a following /ɹ/ also accounts for a dialectal difference in the American English pronunciation of the word *groceries*. In many parts of the eastern United States, speakers pronounce this word as three syllables: /gɹoʊsəɹis/. In the western states, many speakers pronounce this word with two syllables with the /ə/ omitted. Under these conditions the word-internal /s/ is adjacent to a following /ɹ/. The /ɹ/ induces retraction of the /s/ and the following pronunciation results: /gɹoʊʃɹis/.

To sum up, there are several phonetic realizations of the phoneme /t/ in American English. These variations and their conditioning environments are shown in table 3.4. These variations are all heard as /t/'s by speakers of English in spite of the wide phonetic variation.

Types of /l/ in English

The English language has two types of /l/, referred to informally as *dark-l* and *light-l*. Dark-*l*, which occurs in words such as *luck* and *bell*, has a lower sound than light-*l*, which occurs in words such as *leek*. In English dark-*l* is basic. Its dark quality is due to a coarticulation effect caused by an accompanying raised and retracted tongue body. (Because of this high and back (velar) tongue body, dark-*l* is sometimes referred to as *velarized-l*.) Light-*l* is a positional variant occurring before front vowels such as /ɪ/ and /i/. Before front vowels /l/ is not produced with a retracted tongue body—the body is more forward—and thus the light variant results. An English speaker learning French, Spanish, or German must learn to pronounce all of the *l*'s in these languages as light since

Table 3.4
Phonetic variants of the phoneme /t/ in American English

Articulatory description	Phonetic symbol	Conditioning environments	Example words
Released, aspirated	[tʰ]	when syllable-initial	tin [tʰɪn]
Unreleased, preglottalized	[ˀt]	word-final, after a vowel	kit [kʰɪˀt]
Glottal stop	[ʔ]	before a syllabic *n*	kitten [kʰɪʔn̩]
Flap	[ɾ]	between vowels, when the first vowel is stressed (approximate environment)	pitted [pʰíɾɨd]
Alveopalatal stop	[t̠]	syllable-initial before *r*	truck [t̠ɹ̥ʌk]
Released, unaspirated	[t]	when the above conditions are not met first	stint [stɪnt]

none of them has dark-*l*. The IPA symbols for light-*l* and dark-*l* are *l* and ɫ (or ʟ), respectively.

The Relationship between Phonetic and Phonemic Representation

We have seen that the phoneme /t/ has a number of phonetic variants depending on its *position* in a word. Keeping this in mind, we can see that the phonemic symbol /t/ is actually a *cover symbol* for a range of different sounds (or *phones*) that occur in actual speech. We can refer to all of the sounds/phones for which /t/ is a cover symbol as its *allophones* (sometimes also called *positional variants*, since they occur in specific environments). The positional variants that we transcribe as [t], [tʰ], [ˀt], [ɾ], [t̠], and [ʔ] are all instances of the same phoneme /t/. It is important to stress that every positional variant is represented by a phone. Indeed, every phone is an allophone of some phoneme. Thus, we can refer to the allophones [kʰ], [tʰ], or [t], but we must keep in mind that [kʰ] is an allophone of the phoneme /k/ whereas [t] and [tʰ] are allophones of the phoneme /t/. Criteria for determining whether two or more phones are members of the same phoneme or different phonemes are discussed below.

It is clear, then, that we are using two distinct systems of representation for the sounds of English (and of human language in general) and that different information is encoded in each system. For example, the phonetic representation system explicitly represents information concerning aspiration, preglottalization, and flapping, using notational devices such as superscripted *h* and other special symbols summarized in table 3.4. In contrast, the phonemic representation system is more abstract in nature; it ignores such features as aspiration, preglottalization, and flapping.

Since we are using two representation systems for sounds, the question immediately arises, Why should this be so? How can we justify two systems for encoding phonological information? Why should one representation system ignore (or leave unrepresented) articulatory information encoded by the other system? Why shouldn't we simplify our phonological theory and use only one representation system for sounds?

There are some fairly intuitive ways to answer these questions, and so we must stress that we will provide informal answers here rather than precise definitions. Furthermore, we must point out that part of our discussion will assume certain traditional (or "classical") views on the distinction between phonemic and phonetic representations, in which, for the sake of exposition, we will gloss over a number of problems that have arisen in recent work.

The basic idea behind the distinction between phonetic and phonemic representation systems can be best illustrated by considering pairs of words that linguists refer to as *minimal pairs*: pairs of words that (1) have the same number of phonemes, (2) differ in a single sound in a corresponding position in the two words, and (3) differ in meaning. An example is the pair of words *fine* and *vine*. They differ in meaning, but phonologically they differ only in the contrast between initial /f/ and initial /v/. Thus, /faɪn/ and /vaɪn/ constitute a minimal pair.

Now let us consider two possible pronunciations of the word *kit*: [kʰɪt] and [kʰɪˀt]. As noted earlier, for some speakers of English, the final consonant of *kit* is sometimes released (= [t]) and sometimes unreleased (= [ˀt]). The important point is that no meaning difference is associated with the different pronunciations [kʰɪt] and [kʰɪˀt]: both versions are perceived by native speakers of American English as instances of the same word *kit*. Thus, the distinction between the allophones [t] and [ˀt] in word-final position is not contrastive, and we can say that, for some speakers, these allophones of /t/ are in *free variation* (or of optional occurrence) in that position.

The substitution of /v/ for /f/ can create a minimal pair, as we saw in the case of the words *fine* and *vine*; the sounds /f/ and /v/ are therefore members of different phonemes. By contrast, the substitution of [t] for

[ʔt] does not create a minimal pair; they are therefore members of the same phoneme.

The allophones of a phoneme can also occur in what is called *complementary distribution*; that is, one allophone can occur in a position where the other allophone(s) can never appear, and vice versa. The term *complementary distribution* is used because the distribution of one allophone is the complement of the distribution of the other(s). For example, in the position following word-initial /s/, the phoneme /t/ has the obligatory positional variant [t], and the allophones [t^h] and [ʔt] never occur in this position. Allophones of a single phoneme, then, are always either in free variation or in complementary distribution, but in either case they are not contrastive with one another. To repeat, it is only when phones function contrastively to create different meanings that they are members of different phonemes.

The phoneme is actually more than just a cover symbol for a collection of sounds (its allophones)—it has a psychological aspect as well. The phoneme can be viewed as the speaker's internalized representation of a single speech sound, which, however, can have different phonetic shapes depending on the environment in which it appears. To speakers of American English, for example, the phones [t^h], [t], [ʔt], and so forth, are all heard as a "single *t*-sound," the phoneme /t/.

Some linguists understand the phoneme somewhat more concretely and view it as a representation of an ideal articulatory target. Because of the effects of the environment in which the phoneme occurs, however, it may be produced in different allophonic versions. In any case, phonemic writing represents the *basic*, *contrasting sound units* of a language, and many languages use the phonemic principle as the basis of their alphabet.

We write phonemically, then, to represent the minimally contrasting speech sounds of a language. Nevertheless, linguists also have occasion to represent the finer phonetic details of a language. For example, there is often a need to specify just what phonetic features speakers of American English may be carrying over to speaking another language—the features that give them their "American accent." The aspiration of syllable-initial voiceless stops is one such regularly observable feature of English pronunciation, and we want to represent it in some way. To fail to do so would be to fail to give a proper characterization of American English pronunciation. For this reason, we require a phonetic representation system as well as a phonemic representation system in order to characterize the sounds of English (and of human language in general). Speakers of

French and Spanish, for example, do not aspirate syllable-initial voiceless stops, and speakers of American English can pronounce these two Romance languages better if they learn to suppress their aspiration rule. Moreover, the fine phonetic details of the pronunciation of /t/ discussed above are typical of American English but not British English. British English does not have the flap rule, nor does it for the most part have the glottal stop reinforcement rule in word-final position. Thus, the word *pity* has the same phonemic representation in both British and American English (/pɪti/), but the phonetic representations differ: [pɪti] in British English, but [pɪɾi] in American English.

So far we have taken care to specify that our phonemic and phonetic generalizations are based on American English. It is important to note that languages can differ with respect to what phonetic features function distinctively. For example, in Hindi, a language spoken in India, the feature of aspiration does in fact function distinctively in voiceless stops. For speakers of Hindi, the consonants /kʰ/ (aspirated) and /k/ (unaspirated) are perceived as two completely different consonant sounds, and indeed we can find minimal pairs in Hindi showing the contrast between the two. For example, /kʰiil/ means "parched grain," whereas /kiil/ means "nail." Speakers of English tend to hear Hindi /kʰ/ and /k/ as free variants of one another, or else they perceive Hindi unaspirated /k/ as English /g/, given that voiced stops in English are unaspirated. But Hindi /kʰ/ and /k/ also contrast with Hindi /g/. This example brings up an important point: whether or not a phonetic feature (or the phoneme that contains it) is contrastive (phonemic) is a language-particular phenomenon. That is, a phonetic distinction that functions phonemically in one language may or may not function phonemically in another language. Aspiration functions phonemically in voiceless stops in Hindi, but it has no such function in English.

To take another example, there is no phonemic distinction between an *r*-sound and an *l*-sound in Japanese and Korean. In Korean these two sounds are in complementary distribution; they are allophones of a single phoneme. In Japanese only a single *r*-like phoneme occurs. Speakers of American English are baffled by the fact that to a native Japanese speaker the English words *red* and *led* sound like the same word. How can sounds that seem so different sound the same? The answer is that differences that function phonemically in a language are easy for a native speaker to distinguish. In contrast, differences that do not function distinctively may be hard to distinguish. Speakers of Japanese have trouble distinguishing English /r/ and /l/ in the same way that speakers of English have trouble distinguishing Hindi /k/ and /kʰ/ as two separate phonemes.

In most cases the distinction between phonemic and phonetic representations will not be crucial for our purposes. Generally speaking, we will use phonetic representations, using square brackets ([ ]), when discussing specific details of the pronunciation of a word or syllable, and phonemic representations, using slant lines (/ /), when discussing individual consonants and vowels at a more abstract level, as part of a phonological system. When neither the phonemic nor the phonetic transcription is relevant, we will italicize the letter representing the sound under discussion.

3.3 SPECIAL TOPICS

Vowels before /ɹ/

American English /ɹ/ is often one of the most difficult features of pronunciation for speakers of other languages to learn. It is even hard for native speakers themselves, being one of the last sounds that children acquire when they learn American English. It is also one of the sources of extreme dialectal variation—for instance, imagine the word *fire* being pronounced by Ted Kennedy (U.S. senator from Massachusetts), a country music singer such as Randy Travis, and Brian Williams (NBC Evening News anchor). In fact, differences in the pronunciation of /ɹ/ are so complex that we leave it to your instructor to explore with you the features of /ɹ/ in your region.

An interesting aspect of the pronunciation of /ɹ/—one that also has a bearing on dialectal variation, as we will see—lies in the relationship between /ɹ/ and the vowel that precedes it in a word. When beginning students of linguistics transcribe the word *fear*, they often use the tense vowel /i/: /fiɹ/. They notice that the vowel in *fear* sounds higher than the lax vowel /ɪ/ in *bid*, even though they admit that it doesn't seem quite as high as the tense vowel /i/ in *bead* (/bid/). In reality, the vowel in *fear* lies between /ɪ/ and /i/. In fact, the vowel before /ɹ/ is a positional variant—namely, a raised variant of the vowel phoneme /ɪ/, the raising of which is due to the anticipated articulation of the /ɹ/. You can hear that /ɪ/ is the correct vowel by pronouncing both high vowels in the context *s—r*. When you use /ɪ/, the word will sound like *sear* /sɪɹ/. When you use /i/, it will sound like *seer*. Listening to these two words, you will hear that *sear* contains one syllable and *seer* two—the second syllable of *seer* being an *r*-colored vowel transcribed as /ɚ/. The word *seer* is thus written phonemically as /siɚ/. /ɚ/ is an unstressed vowel; when the *r*-colored vowel is stressed, it is transcribed /ɝ/. As you work through this paragraph, it will help to utter the pair of words *sear* and *seer* several times. Ultimately you will recognize a rhythmical difference in these words. The word *sear* /sɪɹ/ is monosyllabic and has one "beat." The word *seer* /siɚ/ is bisyllabic and has two beats. In section 4.4 we will discuss a difference in the tonal patterns that also accompanies the pronunciation of these two words.

The term *r-colored vowel* refers to English vocalic sounds that have an *r*-like quality. The *r*-like quality is a consequence of superimposing the articulatory properties of the /ɹ/ glide onto the articulation of a mid central vowel. It is telling that in British English, which does not have *r*-colored vowels, the vowels that correspond to American English *r*-colored vowels are mid central vowels. Thus, the word *brother* is pronounced /bɹʌðə/.

The difference in syllable structure between the two words *sear* and *seer* results from a property of American English that only a lax vowel can appear in the same syllable with a following /ɹ/; if an *r*-sound alone follows a long (or tense) vowel (i.e., an *r*-sound is the only following phoneme), then it must always occur as an *r*-colored vowel in a second, immediately following syllable. The distributional properties of tense and lax vowels and a following *r*-sound can be stated even more strongly: if a single *r*-sound follows a lax vowel, forming a monosyllable, then this *r* must be the phoneme /ɹ/, and not the *r*-colored vowel /ɚ/. Figure 3.10 displays words that contain the sequence "vowel + /ɹ/." The lax vowels that do not appear in figure 3.10 are /æ/ and /ʌ/. For most speakers of American English, /æ/ does not occur before /ɹ/. The vowel /ʌ/ has actually merged with /r/ to form the *r*-colored vowel written as /ɝ/. In chapter 4 we will

(a)

sear	/sɪɹ/			tour	/tʊɹ/
air	/ɛɹ/	fur	/fɝ/	for	/fɔɹ/
				far	/fɑɹ/

(b)

seer	/siɚ/	sewer	/suɚ/
Bayer	/beɪɚ/	lower	/loʊɚ/
tire	/taɪɚ/	tower	/taʊɚ/
lawyer	/lɔɪɚ/		

Figure 3.10
Vowels that can appear before an *r*-sound: (a) lax, (b) tense

see why several symbols—/ɹ/, /ɝ/, and /ɚ/—are used to represent *r*-like sounds.

As an example of dialectal variation involving vowels before /ɹ/, consider the words *marry*, *merry*, and *Mary*. Speakers in most parts of the United States, especially in the West, pronounce these words the same: /mɛɹi/. However, many speakers on the East Coast, especially those in New York City, pronounce them all differently: *marry* /mæɹi/, *merry* /mɛɹi/, *Mary* /maɹi/, where the first vowel in the last word is the tense /a/ discussed earlier. Since the tense /a/ does not occur in most dialects, it is not available before /ɹ/.

One additional point needs to be made about the lax vowels that can appear before /ɹ/. Although not all dialects of American English make the /ɑ/–/ɔ/ distinction in pronouncing *cot* and *caught* (/kɑt/–/kɔt/), most, if not all, dialects have the vowel /ɔ/ in monosyllables before /ɹ/. This is the vowel in a word such as *lore* /lɔɹ/. As you pronounce this word, you will perceive that it is a monosyllable, and this monosyllabic pronunciation is consistent with the "lax vowel + *r*" principle discussion above.

The vowel in *lore* may sound like the tense vowel /oʊ/, but it is not. The vowel in *lore* may sound "higher" and more *o*-like, but this raising is due to the influence of the following /ɹ/. Moreover, the vowel in *lore* is not as long as the vowel /oʊ/. In fact, if you pronounce the sequence *l*, followed by /oʊ/, followed by an *r*-sound, you will pronounce the word *lower* /loʊɚ/. The difference between *lore* and *lower* further underscores the importance of the conditions that govern the occurrence of vowels before *r*-phonemes in English.

Contractions in Casual Spoken English

In discussing the phonetic properties of English, we have so far focused our attention on phonetic details within single words. Now we must note that in casual spoken forms of American English there are a number of phonological contraction processes in which a *sequence* of words is contracted, or reduced, to a shorter sequence. In all cases contracted forms conform to the phonological rules of English. For example, consider the various phonological contractions of forms of the verb *to be*, illustrated in tables 3.5 and 3.6. Taking table 3.5 first, notice that a sequence of words from formal written language such as *she is* will be pronounced in careful, or formal, speech as a sequence of two separate words /ʃi/ /ɪz/, whereas in more casual, rapid speech they are "merged" into a single bisyllabic (two-syllable) form /ʃíɨz/, with stress on the first syllable, indicated by an

Table 3.5
Phonetic form of contractions of the verb *to be* with personal pronouns in American English: Bisyllabic forms

Formal written	Formal spoken	Casual spoken bisyllabic forms
I am	/aɪ æm/	/áɪəm/ (or /aɪm̩/)
you are	/ju ɑɹ/	/júɚ/
she is	/ʃi ɪz/	/ʃíɨz/
he is	/hi ɪz/	/híɨz/
it is	/ɪt ɪz/	/íɾɨz/
we are	/wi ɑɹ/	/wíɚ/
they are	/ðeɪ ɑɹ/	/ðéɪɚ/

Table 3.6
Phonetic form of contractions of the verb *to be* with personal pronouns in American English: Monosyllabic forms

Casual written	Casual spoken monosyllabic forms
I'm	[aɪm] or [ɑm]
you're	[jʊɹ] or [jɝ]
she's	[ʃiz]
he's	[hiz]
it's	[ɪts]
we're	[wɪɹ]
they're	[ðɛɹ]

accent mark, ´, above the first vowel. Notice further that in the bisyllabic form /ʃíɨz/, the vowel /ɪ/ of /ɪz/ is reduced to /ɨ/, a reduction phenomenon that also takes place when the two-word sequence *I am* becomes a single bisyllabic form /áɪəm/, where /æ/ is reduced to /ə/ in the unstressed syllable. Recall that the reduced vowels /ɨ/ and /ə/ occur only in unstressed syllables of a word, as in *sofa* /sóʊfə/ and *chicken* /tʃíkɨn/. In other words, the bisyllabic forms /ʃíɨz/ and /áɪəm/ (or /aɪm/) reflect phonetic patterns characteristic of single words, and indeed we can consider such contractions as single phonological words.

To take a final example from table 3.5, consider the sequences with the verb *are*: *you are*, *we are*, *they are*. Notice that in the bisyllabic contracted forms of casual speech, *are* [ɑɹ] is reduced to [ɚ] alone (the vowel [ɑ] having been reduced and merged with the /ɹ/), and in fact this /ɚ/ functions as the second (unstressed) syllable. In the forms /juɚ/, /wiɚ/, and /ðeɪɚ/, notice that the tense vowels /u/, /i/, and /eɪ/ are in the first (stressed) syllable, and /ɚ/ forms the second syllable. This sequence "tense vowel + /ɚ/" reflects the syllabic pattern discussed earlier, which is found quite generally in single words of American English: the two members of the sequence "tense vowel + *r*-sound" must be in different syllables. Therefore, this syllabic pattern is just what we find in the bisyllabic contractions /juɚ/, /wiɚ/, and /ðeɪɚ/.

Notice that in very casual speech the bisyllabic forms of the contractions in table 3.5 can be realized as monosyllabic forms (table 3.6). In these examples we see that *am*, *are*, and *is* have lost their vowels entirely and have become reduced to /m/, /ɹ/, and /z/, respectively. Thus, *I'm* is pronounced as monosyllabic /aɪm/ or /ɑm/, having lost the schwa (and the glide in the second form) in /áɪəm/. In the forms *you're* (/jʊɹ/), *we're* (/wɪɹ/), and *they're* (/ðɛɹ/), notice that /ɹ/ is now in the same syllable as the preceding vowel; however, the vowel is now a lax vowel (/ʊ, ɪ, ɛ/) and thus /ɹ/ can occur with it as part of the same syllable. There is another variant pronunciation of the contraction *you're*, namely, /jɝ/. In this case the /ʊ/ and the /ɹ/ have merged to create the *r*-colored vowel /ɝ/.

Consonant Clusters

The sequence of English speech sounds in a word is not arbitrary. In fact, there are strict conditions on the order and type of speech sounds that can appear. At the beginning of a word all consonants except /ŋ/ can appear. If two consonants occur at the beginning, however, the possibilities are quite limited. Consider the sequences in (4):

(4)
*bt, *nk, *gb, *pb, *pt, *pk

None of these combinations can begin an English word, even though they can all be found word-internally (e.g., *na<u>pk</u>in*). By contrast, all the combinations in (5) are permissible word-initial sequences of English:

(5)
br, dr, gr, bl, gl, pr, tr, kr, pl, kl

Native speakers of English can instantly tell if a combination of sounds is possible, suggesting that speakers have internalized a set of principles that determine well-formedness. To begin to form an idea of what these principles are, note that the difference between the disallowed sequences

m	ɛ	ɹ	i	k		ɹ	ɪ	s		m	ʌ	s	
m	e	l	i	k	a	l	i	k	a	m	a	k	a

Figure 3.11
How a speaker of Hawaiian pronounces the English expression *Merry Christmas*

in (4) and the allowed sequences in (5) is that the former consist of two stops and the latter consist of a stop followed by /l/ or /ɹ/. In English a word-initial sequence of two stops is not possible, but a sequence of a stop plus /l/ or /ɹ/ is possible (with a couple of exceptions). Conditions of this type are generally referred to as the *phonotactic constraints* (or *phonotactics*) of a language.

Every language has its own set of conditions on consonant sequencing. When a word is borrowed into one language from another, the borrowed word is often restructured to conform to the sequencing conditions in the borrowing language. When English words are borrowed into the Hawaiian language, first, the consonants and vowels in Hawaiian that are closest to the English counterparts are employed, and second, the English words are restructured to conform to Hawaiian phonotactic constraints. The English greeting *Merry Christmas* sounds very different when pronounced by a native speaker of Hawaiian. Figure 3.11 displays the alterations that occur when the English version is converted into Hawaiian.

Earlier we noted that Hawaiian has 8 consonants (/p, m, n, l, k, h, w, ʔ/) and 5 vowels (/a, e, i, o, u/) and that English has 24 consonants and 15 vowels. There are therefore fewer consonants and vowels available in Hawaiian to represent the consonants and vowels of English. The closest sound to English /ɹ/ is Hawaiian /l/. Somewhat surprising is the fact that the closest consonant to English /s/ is Hawaiian /k/. The other big adjustment in this Hawaiian borrowing is a phonotactic one: Hawaiian does not permit consonant clusters or syllable-final obstruents. As a result, the Hawaiian vowel /a/ is inserted after every consonant that is not immediately followed by a vowel in the borrowed word. *Meli Kalikamaka* is thus the Hawaiian version of *Merry Christmas.*

Exercises

1. George Bernard Shaw, in ridiculing the English spelling system, claimed that a possible spelling for *fish* could be *ghoti*. Why did he claim this? (Hint: The *o* in *women* /wɪmɨn/ is pronounced as an /ɪ/.)

2. Give the English speech sound symbol that corresponds to the following articulatory descriptions:

a. voiceless bilabial stop
b. voiced alveolar stop
c. lax high front vowel
d. voiceless alveolar fricative
e. liquid
f. voiced interdental fricative
g. voiceless alveopalatal affricate
h. tense high back vowel
i. lax low front vowel
j. voiceless velar stop

3. Describe each of the following speech sound symbols using articulatory features:

a. /n/
b. /ʊ/
c. /s/
d. /z/
e. /m/
f. /a/
g. /ɛ/
h. /h/
i. /g/
j. /ʌ/

4. Write the speech sound symbol for the *first* sound in each of the following words. Examples: *fish* /f/, *chagrin* /ʃ/.

a. psychology
b. use
c. thought
d. cow
e. knowledge
f. though
g. pneumonia
h. cybernetics
i. physics
j. memory

5. Write the speech sound symbol for the *last* sound in each of the following words. Examples: *bleach* /tʃ/, *sigh* /aɪ/.

a. cats
b. dogs
c. bushes
d. sighed
e. bleached
f. judge
g. rough
h. tongue
i. garage
j. climb

6. Write the speech sound symbol for the *vowel* in each of the following words. Examples: *fish* /ɪ/, *table* /eɪ/.

a. mood
b. caught
c. cot
d. and
e. tree
f. five
g. bait
h. toy
i. said
j. soot

7. Note the following pairs of words:

a. /bæd/ *bad* and /bæg/ *bag*
b. /sɪn/ *sin* and /sɪŋ/ *sing*
c. /bɛd/ *bed* and /bɛg/ *beg*

You may speak a dialect of American English in which the vowels in the words on the right differ from those in the words on the left. Describe the differences and determine why the vowels are different. (Hint: Consider tongue movement.)

8. Write the following words in the transcription system given in this chapter:

a.
1. through
2. rough
3. gouge
4. Knox
5. draft
6. though
7. blink
8. hinge
9. hang
10. try

b.
1. miss
2. his
3. shoe
4. edge
5. foot
6. three
7. paste
8. trash
9. blunt
10. thigh

c.
1. bow (bend at waist)
2. bow (for shooting arrows)
3. hand
4. which
5. witch
6. hands
7. loose
8. lose
9. tasks
10. chat

d.
1. strengths
2. halve
3. salve
4. cloths
5. clothes
6. yeast
7. gym
8. mend
9. sixths
10. boil

9. Write the names of the letters of the alphabet using the phonemic symbols given in this chapter. For example, *a* = /eɪ/, *b* = /bi/, *c* = /si/, and so forth. Can you find any "rhyme or reason" to the vowels that appear with the alphabetic consonants?

10. Write the following words using the *phonetic* symbols discussed in this chapter:

a. water
b. lit
c. eaten
d. pull
e. craft
f. splat
g. tin
h. beading
i. beating
j. beatin' (casual speech)

11. In some of the following words (e.g., *play*) the *l*'s and the *r*'s are voiceless. Identify these words and try to establish the conditions under which *l* and *r* lose their voicing.

a. Alpo
b. archive
c. black
d. play
e. dream
f. try
g. splat
h. spread
i. leap
j. read

12. Transcribe the following words exhibiting vowels before *r*. (See section 3.3; be aware that dialectal variations will abound in these words.)

a. boor
b. bore
c. poor
d. care
e. car
f. dear
g. fir
h. mire
i. sewer
j. mirror

13. Write the following combinations as contractions (monosyllables, if possible), using the phonetic symbols given in this chapter. Example: *she will* = /ʃɪl/.

a. I will
b. you will
c. he will
d. it will
e. we will
f. they will
g. I would
h. you would
i. she would
j. it would
k. we would
l. they would

14. Using phonetic symbols where possible, write a contracted form (there is more than one version for each of these expressions) for the following sequences, as though they were pronounced in the frame "____ want?" Example: In *What do I want?*, *what do I* = [wʌ́ɾəwaɪ].

a. what do I
b. what do you
c. what does she
d. what does it
e. what do we
f. what do they

15. Nicholas, the 6-year-old son of one of the authors, used the creative spelling *thingck* to spell the word *think*. What assumptions on his part produced this spelling?

Further Reading

General

The study of phonetics is typically divided into *articulatory* and *acoustic phonetics*, and more recently *speech perception* is included. Lieberman and Blumstein 1988, Ladefoged 2001, and Reetz and Jongman 2008 are good introductions to articulatory and acoustic phonetics. Johnson 2003 provides a serious introduction to acoustics and perception. Raphael, Borden, and Harris 2007 covers articulation, acoustics, and perception. Denes and Pinson 1993 provides a good overview of the physics underlying the acoustic study of language. For a discussion of the *International Phonetic Alphabet* (IPA) and other symbol systems for transcribing speech sounds, see Pullum and Ladusaw 1996.

Special Topics

Kahn 1976 is still an excellent and current discussion of the */ɹ/ phoneme* and the vowels that co-occur with it. *Consonant clusters in English* are treated in Clements and Keyser 1983.

Journals

Journal of Phonetics, Phonetica, Journal of the Acoustic Society of America, Journal of Speech and Hearing Sciences

Bibliography

Chomsky, N., and M. Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Clements, G. N., and S. J. Keyser. 1983. *CV phonology: A generative theory of the syllable.* Cambridge, Mass.: MIT Press.

Denes, P., and E. Pinson. 1993. *The speech chain: The physics and biology of spoken language*. New York: W. H. Freeman.

Farmer, A. K., and R. A. Demers. 2010. *A linguistics workbook.* 5th ed. Cambridge, Mass.: MIT Press.

Fry, D. 1979. *The physics of speech.* Cambridge: Cambridge University Press.

Johnson, K. 2003. *Acoustic and auditory phonetics*. 2nd ed. Cambridge, Mass.: Blackwell.

Kahn, D. 1976. Syllable-based generalizations in English phonology. Doctoral dissertation, MIT.

Ladefoged, P. 2001. *A course in phonetics*. 4th ed. New York: Harcourt Brace Jovanovich.

Lenneberg, E. 1967. *Biological foundations of language.* New York: Wiley.

Lieberman, P., and S. Blumstein. 1988. *Speech physiology, speech perception, and acoustic phonetics: An introduction.* New York: Cambridge University Press.

MacKay, I. 1991. *Phonetics: The science of speech production.* 2nd ed. Boston: Little Brown.

Pickett, J. M. 1999. *The acoustics of speech communication: Fundamentals, speech perception theory, and technology.* Needham Heights, Mass.: Allyn and Bacon.

Pullum, G. K., and W. A. Ladusaw. 1996. *Phonetic symbol guide.* 2nd ed. Chicago: University of Chicago Press.

Raphael, L., G. Borden, and K. Harris. 2007. *Speech science primer: Physiology, acoustics, and perception of speech*. 5th ed. Hagerstown, Md.: Lippincott Williams and Wilkins.

Reetz, H., and A. Jongman. 2008. *Phonetics: Transcription, production, acoustics, and perception*. Somerset, N.J.: Wiley-Blackwell.

Small, L. 1999. *Fundamentals of phonetics: A practical guide for students*. Needham Heights, Mass.: Allyn and Bacon.