Phonetics: The Sounds of Language

I gradually came to see that Phonetics had an important bearing on human relations — that when people of different nations pronounce each other’s languages really well (even if vocabulary & grammar not perfect), it has an astonishing effect of bringing them together, it puts people on terms of equality, a good understanding between them immediately springs up.

Daniel Jones

Phonetics is concerned with describing the speech sounds that occur in the languages of the world. We want to know what these sounds are, how they fall into patterns, and how they change in different circumstances. . . . The first job of a phonetician is . . . to try to find out what people are doing when they are talking and when they are listening to speech.

Peter Ladefoged, A Course in Phonetics, 1982, 2nd Edition

Knowledge of a language includes knowledge of the morphemes, words, phrases, and sentences. It also includes the sounds of the language and how they may be “strung” together to form meaningful units. Although there may be some sounds in one language that are not in another, the sounds of all the languages of the world together constitute a limited set of the sounds that the human vocal tract can produce. This chapter will discuss these speech sounds, how they are produced, and how they may be characterized.
The study of speech sounds is called **phonetics**. To describe speech sounds it is necessary to know what an individual sound is, and how each sound differs from all others. This is not as easy as it may seem, and the fact that we generally avoid the confusion of the sign painter in the cartoon is actually remarkable.

A speaker of English knows that there are three sounds in the word *cat*, the initial sound represented by the letter *c*, the second by *a*, and the final sound by *t*. Yet, physically the word is just one continuous sound. You can segment the one sound into parts because you know English. The ability to analyze a word into its individual sounds does not depend on knowledge of spelling. *Not* and *knot* have three sounds even though the first sound in *knot* is represented by the two letters *kn*. The printed word *psycho* has six letters that represent only four sounds — *ps, y, ch, o*.

It is difficult if not impossible to segment the sounds of someone clearing their throat into a sequence of discrete units. This is because these sounds are not the sounds of any morphemes in any human language; it is not because they form a single continuous sound. You do not produce one sound, then another, then another when you say the word *cat*. You move your organs of speech continuously and produce a continuous signal.

Although the sounds we produce and hear are continuous, everyone throughout history who has attempted to analyze language has recognized that speech is divisible into units. According to an ancient Hindu myth, the god *Indra* was the first to segment speech into its separate elements. After he succeeded, the sounds could be perceived as language. Indra may thus be the first phonetician.

Speakers of English can separate *keepout* into the two words *keep* and *out* because they know the language. However, we do not pause between words even though we
sometimes have that illusion. Children learning a language reveal this problem. A two-
year-old child going down stairs heard his mother say, “hold on.” He replied, “I’m hol-
ing don, I’m holing don,” not knowing where the break between words occurred. In fact,
the errors in deciding where a boundary falls between two words has changed the form
of words. At an earlier stage of English, the word apron was napron. However, the
phrase a napron was so often misperceived as an apron that the word lost its initial n.

Some phrases and sentences that are clearly distinct when printed may be ambigu-
ous when read aloud as in the children’s jingle: *I scream, you scream, we all scream for
ice cream*. Read the following pairs aloud and see why we often misinterpret what we
hear:

<table>
<thead>
<tr>
<th>grade A</th>
<th>gray day</th>
</tr>
</thead>
<tbody>
<tr>
<td>It’s hard to recognize speech.</td>
<td>It’s hard to wreck a nice beach.</td>
</tr>
<tr>
<td>The sun’s rays meet.</td>
<td>The sons raise meat.</td>
</tr>
</tbody>
</table>

The lack of breaks between spoken words and individual sounds often makes us
think that speakers of foreign languages run their words together, not realizing that we
also do so. X-ray motion pictures of someone speaking make this lack of breaks in
speech very clear. One can see the tongue, jaw, and lips in continuous motion as the in-
dividual sounds are produced.

Yet, if you know a language you have no difficulty segmenting the continuous
sounds. It doesn’t matter if the language is written or not, or if the person perceiving
the individual sounds can read or not. People who cannot read or write are nevertheless
aware of the individual sounds, and some writing systems do not use spaces between
words on a page, yet everyone who knows the language knows how to segment sen-
tences into words, and words into sounds.

**Identity of Speech Sounds**

It is quite amazing, given the continuity of the speech signal, that we are able to under-
stand the individual words in an utterance. This ability is more surprising because no
two speakers ever say the “same thing” identically. The speech signal produced when
one speaker says *cat* is not the same as that of another speaker’s *cat*. Even two utterances
of *cat* by the same speaker will differ to some degree. George Bernard Shaw pointed to
the impossibility of constructing any set of symbols that will specify all the minute dif-
fferences between sounds in his statement:

> By infinitesimal movements of the tongue countless different vowels can be produced, all of them in use among speakers of English who utter the same vow-
els no oftener than they make the same fingerprints.

Nevertheless, speakers understand each other because they know the same language.

Our knowledge of a language determines when we judge physically different sounds to be the same; we know which aspects or properties of the signal are linguisti-
cally important and which are not. For example, if someone coughs in the middle of say-
ing “How (cough) are you?” a listener will ignore the cough and interpret this simply as
“How are you?” Men’s voices are usually lower in overall pitch than women’s; some
speakers speak slowly, some quickly; others have a "nasal twang." Such pitch or tempo differences, or personal styles of speaking, are not linguistically significant.

Our linguistic knowledge, our mental grammar, makes it possible to ignore nonlinguistic differences in speech. Furthermore, we are capable of making many sounds that we know intuitively are not speech sounds in our language. Many English speakers can make a clicking sound that writers sometimes represent as *tsk tsk tsk*. These sounds are not part of the English sound system. They never occur as part of the words of the utterances we produce. It is even difficult for many English speakers to combine this clicking sound with other sounds. Yet clicks are speech sounds in Xhosa, Zulu, Sosotho, and Khoikhoi — languages spoken in southern Africa — just like the *k* or *t* in English. Speakers of those languages have no difficulty producing them as parts of words. *Xhosa*, the name of a language spoken in South Africa, begins with one of these clicks. Thus, *tsk* is a speech sound in Xhosa but not in English. The sound represented by the letters *th* in the word *think* is a speech sound in English but not in French. The sound produced with a closed mouth when we are trying to clear a tickle in our throats is not a speech sound in any language, nor is the sound produced when we sneeze.

The science of phonetics attempts to describe all the sounds used in human language — sounds that constitute a subset of the totality of sounds that humans can produce.

The way we use our linguistic knowledge to produce a meaningful utterance is complicated. It is a chain of events that starts with an idea or message in the speaker’s brain or mind and ends with a similar message in the hearer’s brain. The language fac-
ulty forms the message in words and transmits it by nerve signals to the organs of speech, which produce the physical sounds.

We can describe the speech sounds at any stage in this chain of events. The study of the physical properties of the sounds themselves is acoustic phonetics. The study of the way listeners perceive these sounds is auditory phonetics. Our primary concern in this chapter is articulatory phonetics, the study of how the vocal tract produces the sounds of language.

Spelling and Speech

The one-I llama,
He's a priest.
The two-I llama,
He's a beast.

And I will bet
A silk pajama
There isn't any
Three-I llama.

Ogden Nash¹

Drawing by Leo Cullum, copyright © 1988 The New Yorker Collection. All rights reserved.

¹ Copyright © 1931 by Ogden Nash, renewed. Reprinted by permission of Curtis Brown, Ltd.
Alphabetic spelling represents the pronunciations of words. However, orthography does not represent the sounds of the words in a language systematically. It is confusing to discuss the production of the different sounds as they are spelled in English words.

Suppose some horrible catastrophe destroyed all Earthlings, and years later extraterrestrials exploring Earth discover some fragments of English writing that include the following sentence:

Did he believe that Caesar could see the people seize the seas?

How would an ET phonetician decide that e, ie, ae, ee, eo, ei, and ea all represented the same sound? To add to the confusion, this sentence might crop up later:

The silly amoeba stole the key to the machine.

English speakers learn how to pronounce these words when learning to read and write, and therefore they know that y, oe, ey, and i represent the same sound as the boldface letters in the first sentence.

On the other hand, consider:

My father wanted many a village dame badly

Here the letter a represents the several sounds in father, wanted, many, and so on.

In any science, the objects of study, when different, must be given different names or symbols, and the science of phonetics is no exception. Each distinct sound must have a distinct symbol to represent it; and each symbol must represent one and only one distinct sound.

The Phonetic Alphabet

The English have no respect for their language, and will not teach their children to speak it. They cannot spell it because they have nothing to spell it with but an old foreign alphabet of which only the consonants — and not all of them — have any agreed speech value.

G. B. Shaw, Preface to Pygmalion

The discrepancy between spelling and sounds gave rise to a movement of “spelling reformers” called orthoepists. They wanted to revise the alphabet so that one letter would correspond to one sound and one sound to one letter, thus creating a phonetic alphabet to simplify spelling.

George Bernard Shaw followed in the footsteps of three centuries of spelling reformers in England. In typical Shavian manner he pointed out that we could use the English spelling system to spell fish as ghoti — the gh like the sound in enough, the o like the sound in women, and the ti like the sound in nation. Shaw was so concerned about English spelling that he included a provision in his will for a new “Proposed English Alphabet” to be administered by a “Public Trustee” who would have the duty of seeking and publishing a more efficient alphabet. This alphabet was to have at least forty letters to enable “the said language to be written without indicating single sounds by groups of letters or by diacritical marks.” After Shaw’s death in 1950, 450 designs for such an al-
alphabet were submitted from all parts of the globe. Four alphabets were judged equally
good, and the prize money was shared among their designers, who collaborated to pro-
duce the alphabet designated in Shaw’s will. Shaw also stipulated in his will that his play
Androcles and the Lion be published in the new alphabet, with “the original Doctor
Johnson’s lettering opposite the transliteration page by page and a glossary of the two al-
phabets.” This version of the play came out in 1962.

It is easy to understand why spelling reformers believe there is a need for a phonetic
alphabet. Several letters may represent a single sound:

  to too two through threw clue shoe

A single letter may represent different sounds:

dame dad father call village many
Phonetics: The Sounds of Language

A combination of letters may represent a single sound:

\[
\begin{array}{llll}
\text{shoot} & \text{character} & \text{Thomas} & \text{physics} \\
\text{either} & \text{deal} & \text{rough} & \text{nation} \\
\text{coat} & \text{glacial} & \text{theater} & \text{plain} \\
\end{array}
\]

Some letters have no sound at all in certain words:

\[
\begin{array}{llll}
\text{mnemonic} & \text{autumn} & \text{resign} & \text{ghost} \\
\text{pterodactyl} & \text{write} & \text{hole} & \text{corps} \\
\text{psychology} & \text{sword} & \text{debt} & \text{gnaw} \\
\text{bough} & \text{lamb} & \text{island} & \text{knot} \\
\end{array}
\]

The spelling may fail to represent sounds that occur. In many words, the letter \( u \) represents a \( y \) sound followed by a \( u \) sound:

\[
\begin{array}{ll}
\text{cute} & \text{(compare: coot)} \\
\text{futile} & \text{(compare: rule)} \\
\text{utility} & \text{(compare: Uzbek)} \\
\end{array}
\]

One letter may represent two sounds; the final \( x \) in \( Xerox \) represents a \( k \) followed by an \( s \).

Whether we support or oppose spelling reform, it is clear that we cannot depend on the spelling of words to describe the sounds of English. The alphabets designed to fulfill Shaw’s will were not the first phonetic alphabets. In Shaw’s lifetime, the phonetician Henry Sweet, the prototype for Shaw’s character Henry Higgins in the play \textit{Pygmalion} or the musical play or movie \textit{My Fair Lady}, produced a phonetic alphabet.

In 1888 the interest in the scientific description of speech sounds led the \textit{International Phonetic Association (IPA)} to develop a phonetic alphabet to symbolize the sounds of all languages. Since many languages use a Roman alphabet like that used in the English writing system, the IPA utilized many Roman letters as well as invented symbols. These alphabetic characters have a consistent value, unlike ordinary letters that may or may not represent the same sounds in the same or different languages.

A phonetic alphabet should include enough symbols to represent the “crucial” linguistic differences. At the same time it should not, and cannot, include noncrucial differences, since such differences are infinitely varied.

Table 6.1 is a list of the phonetic symbols that we will use to represent English speech sounds. The symbols do not tell us everything about the sounds, which may vary from person to person, and which may depend on their position in a word. These symbols are intended for use by persons knowing English. They are not all the phonetic symbols needed for English sounds. When we discuss the sounds in more detail later in the chapter, we will add appropriate symbols.

The symbol \( [ə] \) is called a schwa. We use it only to represent vowels in unstressed syllables (we will discuss stress below). In the word \textit{phonetic}, for example, the first syllable \( (pho) \) and last syllable \( (tic) \) are unstressed, while the middle syllable \( (NE) \) is stressed. So we spell it phonetically as \( [fənetɪk] \). (There is great variation in the way
Table 6.1 A Phonetic Alphabet for English Pronunciation

<table>
<thead>
<tr>
<th>Consonants</th>
<th>Vowels</th>
</tr>
</thead>
<tbody>
<tr>
<td>p pill</td>
<td>i beet</td>
</tr>
<tr>
<td>b bill</td>
<td>e bait</td>
</tr>
<tr>
<td>m mill</td>
<td>u boot</td>
</tr>
<tr>
<td>f feel</td>
<td>o boat</td>
</tr>
<tr>
<td>v veal</td>
<td>æ bat</td>
</tr>
<tr>
<td>θ thigh</td>
<td>a butt</td>
</tr>
<tr>
<td>ð thy</td>
<td>ð bite</td>
</tr>
<tr>
<td>jʃ shill</td>
<td>aw bout</td>
</tr>
<tr>
<td>3ʃ azure</td>
<td>ʊj boy</td>
</tr>
</tbody>
</table>

Speakers of English produce this unstressed vowel, but it is often phonetically similar to the sound [ə] of cut [kʌt]. All other vowel symbols are used in syllables that are not unstressed.

Speakers of different English dialects pronounce some words differently from those of other speakers. For example, some of you may pronounce the words which and witch identically. If you do, the initial sound of both words is symbolized by w in the chart. Some speakers of English pronounce bought and pot with the same vowel; others pronounce them with the vowel sounds in bore and bar, respectively. We have thus listed both words in the chart of symbols. It is difficult to include all the phonetic symbols needed to represent all English dialect differences. We are sorry if a vowel sound in your dialect is not included in the table.

Some of the symbols in Table 6.1 are those traditionally used by linguists in the United States in place of IPA symbols. Here are some equivalences:

<table>
<thead>
<tr>
<th>U.S.</th>
<th>IPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>ʃ</td>
<td>f</td>
</tr>
<tr>
<td>ʒ</td>
<td>ʒ</td>
</tr>
<tr>
<td>ɛ</td>
<td>tʃ</td>
</tr>
<tr>
<td>j</td>
<td>dʒ</td>
</tr>
<tr>
<td>u</td>
<td>u</td>
</tr>
</tbody>
</table>

We will use [ʃ], [f] and [ʒ], [ʒ] interchangeably to familiarize readers with both notations, since both are common in books on language and linguistics. We will however, use [ɛ] and [tʃ] instead of the IPA symbols for the first and last sounds in church and judge, respectively.

Using these symbols, we can now unambiguously represent the pronunciation of words. For example, in the six words below, ou represents six distinct vowel sounds; the gh is silent in all but rough, where it is pronounced [f]; the th represents two sounds, and the l in would is also silent.
We will always use square brackets around the phonetic transcription to distinguish it from ordinary spelling.

Articulatory Phonetics

The voice is articulated by the lips and the tongue... Man speaks by means of the air which he inhales into his entire body and particularly into the body cavities. When the air is expelled through the empty space it produces a sound, because of the resonances in the skull. The tongue articulates by its strokes; it gathers the air in the throat and pushes it against the palate and the teeth, thereby giving the sound a definite shape. If the tongue would not articulate each time, by means of its strokes, man would not speak clearly and would only be able to produce a few simple sounds.

Hippocrates (460–377 B.C.E)

The production of any sound involves the movement of air. Most speech sounds are produced by pushing lung air through the opening between the vocal cords, up the throat, and into the mouth or nose, and finally out of the body. Some technical jargon—see chapter 10 for a discussion of “jargon” in general—is required. The opening between the vocal cords is the glottis and is located in the larynx (often referred to as the “voice box,” and pronounced “lair rinks.”). The tubular part of the throat above the larynx is the pharynx (rhymes with larynx, and produces a very high score at Scrabble). What sensible people call “the mouth,” we linguists call the oral cavity to distinguish it from the nasal cavity, which is the nose and the plumbing that connects it to the throat, plus your sinuses. All of it together is the vocal tract. Figure 6-1 should make these descriptions more clear.

What distinguishes one sound from the other? If you bang a large round drum you will get one sound; if you bang a small round drum you will get a different sound; if you bang a small oblong drum you will get still another sound. The size and shape of the vessel containing the air that is moving makes a difference. This is also true in the production of speech sounds. The vocal tract acts as the vessel of air. When it changes shape, different sounds are produced.

Airstream Mechanisms

Most of the sounds of the world’s languages are produced by pushing air out of the lungs through the vocal tract. Since the air comes from the lungs, it is pulmonic; and since it
Places of articulation: 1. bilabial; 2. labiodental; 3. interdental; 4. alveolar; 5. (alveo)palatal; 6. velar; 7. uvular; 8. glottal.

Figure 6.1 The vocal tract.

is pushed out, it is egressive. All of the speech sounds of English have a pulmonic egressive airstream mechanism.

Other airstream mechanisms are rare in the world’s languages, but are still part of phonetics. Another kind of egressive sound is made when air in the mouth is pressurized by an upward movement of the closed glottis, and then released suddenly, producing a sharp sound called an ejective. An ejective “p” sound makes a distinctive pop.

Sounds may also be ingressive. In one kind of ingressive sound, the air is sucked into the mouth to make clicks (such as tsk). In a different kind, air is drawn from the mouth into the throat to make implosives. There are four airstream mechanisms in all, then. Pulmonic egressive is the most common. The other three are the mechanisms that produce ejectives, clicks, and implosives.

Ejectives are found in many American Indian and African languages, as well as languages spoken in the Caucasus, a region between the Black and Caspian Seas. Implosives also occur in the languages of the American Indians and throughout Africa, India, and Pakistan. Clicks occur in the Southern Bantu languages such as Xhosa and Zulu, and in the languages spoken by the Bushmen and Khoikhoi. A detailed description of these
different airstream mechanisms goes beyond the requirements of an introductory text. We mention them to show that sounds can be classified according to the airstream mechanism used to produce them. In the rest of this chapter, we will discuss only sounds produced by a pulmonic egressive airstream mechanism.

**Consonants**

The sounds of all languages fall into two classes: consonants and vowels. Consonants are produced with some restriction or closure in the vocal tract that impedes the flow of air from the lungs. In phonetics, the terms *consonant* and *vowel* refer to types of sounds, not to the letters that represent them. In speaking of the alphabet, we may call "a" a vowel and "b" a consonant, but that means only that we use the letter "a" to represent vowel sounds, and the letter "b" to represent consonant sounds.

**Places of Articulation**

Different consonantal sounds result according to the **place of articulation**, which is where in the vocal tract the airflow restriction occurs. Movement of the tongue and lips, called the *articulators*, cause the restriction, reshaping the oral cavity in various ways to produce the various consonants. In this section, we discuss the major consonantal place features. As you read the description of each class of sounds, pronounce them and try to feel which articulators are moving and to where. Refer to Figure 6.1 to remind yourself of the terminology.

**Bilabials** [p] [b] [m]  When we produce a [p], [b], or [m] we articulate by bringing both lips together. These sounds are therefore called **bilabials**.

**Labiodentals** [f] [v]  We also use our lips to form [f] and [v] as in *fine* [fain] and *vine* [vain]. We articulate these sounds by touching the bottom lip to the upper teeth, which is why these sounds are called **labiodental**, labio- referring to lips and *dental* to teeth.

**Interdentals** [θ] [ð]  In ordinary spelling, the sounds [θ] and [ð] are both represented by *th*, for example, *thin* [θin] and *then* [ðen]. To articulate these **interdental** ("between the teeth") sounds, one inserts the tip of the tongue between the upper and lower teeth. However, for some speakers the tongue merely touches the teeth, making a sound more correctly called **dental**. We will nevertheless continue to use **interdental** since it describes the most common articulation.

**Alveolars** [t] [d] [n] [s] [z] [l] [r]  Alveolar sounds are articulated by raising the front part of the tongue to the **alveolar ridge** (see Figure 6.1). Pronounce the words *do* [du], *new* [nu], *two* [tu], *sue* [su], *zoo* [zu]. You should feel your tongue touch or almost touch the bony tooth ridge as you produce the first sounds in these words.

For the **lateral** [l], the tip of the tongue rises to the alveolar ridge leaving the rest of the tongue down, permitting the air to escape laterally over its sides. You can feel it in the "la" of "tra la la."

The sound [r] is produced in a variety of ways. Many English speakers produce [r] by curling the tip of the tongue back behind the alveolar ridge. In that case the [r] is a **retroflex** sound. In some languages, the [r] may be an alveolar **trill**, produced by the tip
of the tongue vibrating against the roof of the mouth. Other symbols can be used for these different r sounds, and in a very detailed phonetic description, we would include some of them. For the purposes of this book, however, we will use the symbol [r] for all the varieties produced by speakers of English.

**Palatals [ʃ] [s] [ʒ] [ç] [ʝ]** To produce the sounds in the middle of the words *mission* [mɪˈʃən] and *measure* [meˈzər], the front part of the tongue is raised to a point on the hard palate just behind the alveolar ridge. These are palatal sounds. (The term *alveopalatal* is also used.)

The palatal region is also the place of articulation of [ç] and [ʝ], the sounds that begin and end the words *church* and *judge*.

**Velars [k] [g] [ŋ]** Another class of sounds is produced by raising the back of the tongue to the soft palate or velum. The initial and final sounds of the words *kick* [kɪk], *gig* [gɪg], and the final sounds of the words *back* [bæk], *bag* [bæɡ], and *bang* [bæŋ] — [k], [g], and [ŋ] — are all velar sounds.

**Uvulars [ʁ] [q] [ɢ]** Uvular sounds are produced by raising the back of the tongue to the uvula, the fleshy appendage that hangs down in the back of the throat. The r in French is often a uvular trill and is symbolized by [ʁ]. Uvular sounds occur in other languages. Arabic, for example, has two uvular sounds symbolized as [q] and [ɢ].

**Glottal [ʔ] [h]** The [h] sound that starts words such as *house* [haʊs], *who* [hu], and *hair* [heər] is a glottal sound. Although classified as a consonant, there is no airflow restriction in pronouncing [h]. Its sound is from the flow of air through the open glottis. The tongue and lips are usually in the position for the production of the following vowel.

If the air is stopped completely at the glottis by tightly closed vocal cords, the sound produced is a glottal stop. This is the sound sometimes used instead of [t] in *button* and *Latin*. It also may occur in colloquial speech at the end of words like *don’t*, *won’t*, and *can’t*. In some American dialects (we’ll discuss dialects in chapter 10), it regularly replaces the tt sound in words like *bottle* or *glottal*. If you say “ah-ah-ah-ah” with one “ah” right after another and do not sustain the vowel sound, you will be producing glottal stops between the vowels. In some languages, the glottal stop functions like the stops [p], [t], [k] in English. The IPA symbol for a glottal stop looks something like a question mark without the dot on the bottom [ʔ].

Table 6.2 summarizes the classification of English consonants by their place of articulation. The glottal stop is not included in this table since only some speakers use it in some words. The uvular sounds do not occur regularly in English.

<table>
<thead>
<tr>
<th>Place of Articulation of English Consonants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bilabial</td>
</tr>
<tr>
<td>Labiodental</td>
</tr>
<tr>
<td>Interdental</td>
</tr>
<tr>
<td>Alveolar</td>
</tr>
<tr>
<td>Palatal</td>
</tr>
<tr>
<td>Velar</td>
</tr>
<tr>
<td>Glottal</td>
</tr>
</tbody>
</table>
MANNERS OF ARTICULATION

We have described a number of classes of consonants according to their place of articulation, yet we are still unable to distinguish the sounds in each class from each other. What distinguishes [p] from [b], or [b] from [m]? All are bilabial sounds. What is the difference between [t], [d], and [n], which are all alveolar sounds?

Speech sounds also vary in the way the airstream is affected as it flows from the lungs up and out of the mouth and nose. It may be blocked or partially blocked; the vocal cords may vibrate or not vibrate. We refer to this as the manner of articulation.

Voiced and Voiceless Sounds If the vocal cords are apart during airflow, the air flows freely through the glottis and supraglottal cavities (the parts of the vocal tract above the glottis; see Figure 6.1). The sounds produced in this way are voiceless sounds: [p], [t], [k], and [s] in the English words seep [sip], seat [sit], and seek [sik] are voiceless sounds.

If the vocal cords are together, the airstream forces its way through and causes them to vibrate. Such sounds are voiced and are illustrated by the sounds [b], [d], [g], and [z] in the words bate [bet], date [det], gate [get], cob [kab], cod [kad], cog [kag], and daze [dez]. If you put a finger in each ear and say the voiced “z-z-z-z-z,” you can feel the vibrations of the vocal cords. If you now say the voiceless “s-s-s-s-s,” you will not feel these vibrations (although you might hear a hissing sound in your mouth). When you whisper, you are making all the speech sounds voiceless.

The voiced/voiceless distinction is very important in English. This phonetic feature, or property, distinguishes the words in word pairs like the following:

<table>
<thead>
<tr>
<th>rope/robe</th>
<th>fate/fade</th>
<th>rack/rag</th>
<th>wreath/wreathe</th>
</tr>
</thead>
<tbody>
<tr>
<td>[rop]/[rob]</td>
<td>[fet]/[fed]</td>
<td>[rækl]/[ræg]</td>
<td>[riθ]/[rið]</td>
</tr>
</tbody>
</table>

The first word of each pair ends with a voiceless sound and the second word with a voiced sound. All other aspects of the sounds in each word pair are identical; the position of the lips and tongue is the same.

The voiced/voiceless distinction also occurs in the following pairs, where the first word begins with a voiceless sound and the second with a voiced sound:

<table>
<thead>
<tr>
<th>fine/vine</th>
<th>seal/zeal</th>
<th>choke/joke</th>
</tr>
</thead>
<tbody>
<tr>
<td>[fajn]/[vain]</td>
<td>[sil/zil]</td>
<td>[čok]/[jok]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>peat/beat</th>
<th>tote/dote</th>
<th>kale/gale</th>
</tr>
</thead>
<tbody>
<tr>
<td>[pit]/[bit]</td>
<td>[tot]/[dot]</td>
<td>[kel]/[gel]</td>
</tr>
</tbody>
</table>

In our discussion of the voiceless bilabial stop [p], we did not distinguish the initial sound in the word pit from the second sound in the word spit. There is, however, a phonetic difference in these two voiceless stops. During the production of voiceless sounds, the glottis is open and the air passes freely through the opening between the vocal cords. When a voiceless sound is followed by a voiced sound such as a vowel, the vocal cords must close so that they can vibrate.

Voiceless sounds fall into two classes depending on the timing of the vocal cord closure. In English, when we pronounce the word pit, there is a brief period of voicelessness immediately after the p sound is released. That is, after the lips come apart, the
vocal cords remain open for a very short time. Such sounds are called **aspirated** because an extra puff of air escapes through the open glottis.

When we pronounce the *p* in *spit*, however, the vocal cords start vibrating as soon as the lips open. Such sounds are **unaspirated**. The *t* in *tick* and the *k* in *kin* are also aspirated voiceless stops, while the *t* in *stick* and the *k* in *skin* are unaspirated. Hold your palm about 2 inches in front of your lips and say *pit*. You will feel a puff of air, which you will not feel when you say *spit*.

Figure 6.2 shows in diagrammatic form the timing of the articulators (in this case the lips) in relation to the state of the vocal cords.

In the production of the voiced [b], the vocal cords are vibrating throughout the closure of the lips, and continue to vibrate for the vowel production after the lips part. Most English speakers do not voice initial [b] fully. That is, the vocal cords remain open for a fraction of a second after the lips seal. Voiced consonants in other languages such as French are typically fully voiced throughout. These are idiosyncratic differences between languages, or even dialects of languages and individual speakers.

We indicate aspirated sounds by writing the phonetic symbol with a raised *h*, as in the following examples:

<table>
<thead>
<tr>
<th>English</th>
<th>IPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>pate</td>
<td>[pʰæt]</td>
</tr>
<tr>
<td>spate</td>
<td>[spet]</td>
</tr>
<tr>
<td>tale</td>
<td>[tʰel]</td>
</tr>
<tr>
<td>stale</td>
<td>[stel]</td>
</tr>
<tr>
<td>kale</td>
<td>[kʰel]</td>
</tr>
<tr>
<td>scale</td>
<td>[skel]</td>
</tr>
</tbody>
</table>

---

**Figure 6.2** Timing of articulation and vocal-cord vibrations for voiced, voiceless, unaspirated, and voiceless aspirated stops.
Nasal and Oral Sounds  The voiced/voiceless distinction differentiates the bilabials [b] and [p]. The sound [m] is also a bilabial, and it is voiced. What distinguishes it from [b]?

Figure 6.1 shows the roof of the mouth divided into the (hard) palate and the soft palate (or velum). The palate is a hard bony structure at the front of the mouth. You can feel it with your thumb. As you slide your thumb along the hard palate back toward the throat, you will feel the velum, which is where the flesh becomes soft and pliable. The velum terminates in the uvula, which you can see in a mirror if you open your mouth wide and say “aaah.” The velum is movable, and when it is raised all the way to touch the back of the throat, the passage through the nose is cut off and air can escape only through the mouth.

Sounds produced with the velum up, blocking the air from escaping through the nose, are oral sounds, since the air can escape only through the oral cavity. Most sounds in all languages are oral sounds.

When the velum is not in its raised position, air escapes through both the nose and the mouth. Sounds produced this way are nasal sounds. The sound [m] is a nasal consonant. Thus [m] is distinguished from [b] because it is a nasal sound, while [b] is an oral sound.

The diagrams in Figure 6.3 show the position of the lips and the velum when [m], [b], and [p] are articulated. The sounds [p], [b], and [m] are produced by stopping the airflow at the lips; [m] and [b] differ from [p] by being voiced; [m] differs from [b] by being nasal.

The same oral/nasal difference occurs in dear [dir] and near [nir], rug [rəg] and rung [rʌŋ]. The velum is raised in the production of [d] and [g], preventing the air from flowing through the nose, whereas in [n] and [ŋ] the velum is down, letting the air go through both the nose and the mouth when the closure is released. The sounds [m], [n], and [ŋ] are therefore nasal sounds, and [b], [d], and [g] are oral sounds.

These phonetic features permit the classification of all speech sounds into four classes: voiced, voiceless, nasal, and oral. One sound may belong to more than one class, as shown in Table 6.3.
Table 6.3 Four Classes of Speech Sounds

<table>
<thead>
<tr>
<th></th>
<th>Oral</th>
<th>Nasal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voiced</td>
<td>b d g</td>
<td>m n η</td>
</tr>
<tr>
<td>Voiceless</td>
<td>p t k</td>
<td>*</td>
</tr>
</tbody>
</table>

*Nasal consonants in English are usually voiced. Both voiced and voiceless nasal sounds occur in other languages.

We now have three ways of classifying consonants: by voicing, by place of articulation, and oral vs. nasal. For example, [p] is a voiceless, bilabial, oral sound; [n] is a voiced, alveolar, nasal sound, and so on.

**Stops** [p] [b] [m] [t] [d] [n] [k] [g] [ŋ] [ɾ] [j] [ʔ]** We are seeing finer and finer distinctions of speech sounds. However, both [t] and [s] are voiceless, alveolar, oral sounds. What distinguishes them? After all, *sack* and *tack* are different words.

In producing consonants, the airstream may be completely stopped or just partially obstructed. Sounds that are stopped completely in the oral cavity for a brief period are, not surprisingly, called **stops**. The sound [t] is a stop, but the sound [s] is not, and that is what makes them different speech sounds.

The final sounds in the words *top* [tap], *bomb* [bam], *duke* [dud], *dune* [dun], *root* [rut], *rack* [rák], *rag* [rág], and *rang* [ráŋ] are stops that occur in English.

In the production of the nasal stops [m], [n], [ŋ], although the air flows freely through the nose, the airflow is blocked completely in the mouth; therefore, nasal consonants are stops.

Sounds in which there is no stoppage in the oral tract are **continuants**. All the sounds of a language are either stops or continuants (nonstop).

Nonnasal, or oral, stops are also called **plosives** because the air that is blocked in the mouth “explodes” when the closure is released. This explosion does not occur during the production of nasal stops because the air escapes continuously through the nose.

[p], [b], and [m] are **bilabial stops**, with the airstream stopped at the mouth by the complete closure of the lips.

[t], [d], and [n] are **alveolar stops**; the airstream is stopped by the tongue making a complete closure at the alveolar ridge.

[k], [g], and [ŋ] are **velar stops** with the complete closure at the velum.

[ʔ] is a **glottal stop**. Although there is no stoppage of air in the oral cavity, the air is completely stopped at the glottis.

[ɾ] and [ʃ] are **palatal affricates** with complete stop closures. We will describe them below.

We have been discussing the sounds that occur in English. Some sounds, including stops, occur in other languages but not in English. For example, in Quechua, spoken in Bolivia and Peru, uvular stops occur, where the back of the tongue is raised and moved rearward to form a complete closure with the uvula. The phonetic symbol [q] denotes the voiceless version of this stop, which is the initial sound in the name of the language “Quechua.” The voiced uvular stop [g] also occurs in Quechua. We find glottal stops in a number of languages such as Arabic.
Fricatives [f] [v] [θ] [ð] [s] [z] [ʃ] [ʒ] In the production of some continuants, the airflow is so severely obstructed that it causes friction, and the sounds are therefore called fricatives. The sounds [f], [v], [θ], [ð], [s], [z], [ʃ], and [ʒ] are pronounced in this manner.

[f] and [v] are labiodental fricatives; the friction is created at the lips and teeth, where a narrow passage permits the air to escape. The [f] is voiceless and the [v] is voiced. The following pairs follow this pattern.

[θ] and [ð] are the interdental fricatives, represented by th in thin and then. The friction occurs at the opening between the tongue and teeth.

[s] and [z] are alveolar fricatives, with the friction created at the alveolar ridge.

[ʃ] and [ʒ] are the palatal fricatives, and contrast in such pairs as mission [miʃən] and measure [mezər]. They are produced with friction created as the air passes between the tongue and the palate behind the alveolar ridge. In English, the voiced palatal fricative never begins words (except in words borrowed from French like genre, which some English speakers produce with a French pronunciation). The voiceless palatal sound begins the words shoe [ʃu] and sure [ʃər] and ends the words rush [rʌʃ] and push [pʊʃ].

Most modern English dialects do not include velar fricatives, although they occurred in an earlier stage of English in such words as right, knight, enough, and through, where the gh occurs in the spelling. If you raise the back of the tongue as if you were about to produce a [k] or [g], but stop just short of touching the velum, you will produce a velar fricative. The ch ending in the German pronunciation of the composer’s name Bach is a velar fricative. Some speakers of modern English substitute a voiceless velar fricative in words like bucket and a voiced velar fricative in such words as wagon for the velar stops, especially in rapid, informal speech. The IPA symbol for the voiceless velar fricative is [x], and for the voiced velar fricative, it is [ɣ].

In languages like French, the uvular trill [R] occurs as the sound represented by r in French words such as rouge, “red.” Voiced glottal fricatives, which do not occur in English, do occur in other languages, such as Czech. Speakers of Arabic produce pharyngeal fricatives by pulling the tongue root toward the back wall of the pharynx. It is difficult to pull the tongue back far enough to make a complete pharyngeal stop closure, but both voiced and voiceless pharyngeal fricatives can be produced, and can be distinguished from velar fricatives.

All fricatives are continuants: Although the airstream is obstructed as it passes through the oral cavity, it is not completely stopped.

Affricates [ç], [ʃ] Some sounds are produced by a stop closure followed immediately by a gradual release of the closure that produces an effect characteristic of a fricative. These sounds are affricates. The palatal sounds that begin the words church and judge are voiceless and voiced affricates, respectively. Phonetically, an affricate is a sequence of a stop plus a fricative. Thus, the ch in church is the same as the sound combination [tʃ] + [ʃ], as you can see by pronouncing white shoes and why choose rapidly. The two expressions are indistinguishable. The voiceless and voiced affricates are symbolized as [tʃ] (IPA [tʃ]) and [dʒ] (IPA [dʒ]), respectively. In the American tradition, [ç], [ʃ] are the more common symbols for these sounds, and are used in this book.

Because the air is stopped completely during the initial articulation of an affricate, these sounds are also classified as stops.
Liquids [l] [r] In the production of the sounds [l] and [r], there is some obstruction of the airstream in the mouth, but not enough to cause any real constriction or friction. These sounds are liquids. The lateral liquid [l] and the retroflex liquid [r] are described in the earlier alveolar section. They are both voiced sounds. In some contexts the liquids sound similar, as they do to Dennis in the cartoon.

As mentioned earlier, the r sounds that occur in various dialects of English, and in various languages, differ somewhat from each other. We are using the symbol [r] for this whole class of sounds. An alveolar trilled r occurs in many languages, such as Spanish. In addition, uvular trills occur, produced by vibrating the uvula. Some French speakers use uvular trills in the pronunciation of r; others use uvular fricatives. In yet other languages, the r is produced by a single tap or flap of the tongue against the alveolar ridge, and we sometimes call that sound "a flap." In Spanish both the trilled r and the flap occur, the former in perro, "dog," and the latter in pero, "but."

Some speakers of British English pronounce the r in the word very as a flap. It sounds like a "very fast d." Most American speakers produce a flap instead of a [t] or [d] in words like writer, rider, latter, and ladder. The IPA symbol for the alveolar tap or flap is [r]. American linguists often use the upper case [D] to represent this sound.

In English, [l] and [r] are regularly voiced. When they follow voiceless sounds, as in please and price, they may be partially devoiced, that is, the voicing doesn’t begin until part way through the consonant. Many languages have a voiceless l as an independent
sound, in which case it is actually a fricative. Welsh is such a language; the name Lloyd in Welsh starts with the voiceless fricative l.

Some languages may lack liquids entirely, or may have only a single one. The Cantonese dialect of Chinese has the single liquid [l]. Some English words are difficult for Cantonese speakers to pronounce, and they may substitute an [l] for an [r] when speaking English.

The reason that speakers of languages with only one liquid tend to substitute that sound for the nonoccurring liquid is the acoustic similarity of these sounds. This auditory similarity is the reason that linguists group them in one class, and why they function as a single class of sounds in certain circumstances. For example, in English, the only two consonants that occur after an initial [k], [g], [p], or [b] are the liquids [l] and [r]. Thus we have crate [kret], clock [klak], plate [plet], pray [pre], bleak [blik], break [brek], but no word starting with [ps], [bt], [pk], and so on. (Notice that in words like psychology or pterodactyl the p is not pronounced. Similarly, in knight or knot the k is not pronounced, although at an earlier stage of English it was.)

**Glides [j][w]** The sounds [j] and [w], the initial sounds of you [ju] and woo [wu], are produced with little or no obstruction of the airstream in the mouth. They are always preceded or followed directly by a vowel. In articulating [j] or [w], the tongue moves rapidly in gliding fashion either toward or away from a neighboring vowel, hence the term glide. Glides are transitional sounds that are sometimes called semivowels.

The glide [j] is a palatal sound; the blade of the tongue (the front part minus the tip) is raised toward the hard palate in a position almost identical to that in producing the vowel sound [i] in the word beat [bit]. In pronouncing you [ju], the tongue moves rapidly from the [j] to the [u] vowel.

The glide [w] is produced by both raising the back of the tongue toward the velum and simultaneously rounding the lips. It is thus a labio-velar glide, or a rounded velar glide. In the dialect of English where speakers have different pronunciations for the words which and witch, the velar glide in the first word is voiceless, symbolized as [m] (an “upside-down” w), and in the second word it is voiced, symbolized as [w]. When the pronunciation of the two words is the same, it is the voiced [w]. The position of the tongue and the lips for [w] is similar to that for producing the vowel sound sound in suit [sut]. In pronouncing we [wi], the tongue moves rapidly from the [w] to the [i] vowel.

**Phonetic Symbols for American English Consonants**

We are now capable of distinguishing all of the consonant sounds of English via the properties of voicing, nasality, and place and manner of articulation. For example, [f] is a voiceless, (oral), labiodental fricative; [n] is a (voiced), nasal, alveolar stop. The parenthesized features are usually not mentioned since they are redundant; all sounds are oral unless nasal is specifically mentioned, and all nasals are voiced in English.

Table 6.4 lists the consonants by their phonetic features. The rows stand for manner of articulation and the columns for place of articulation. Symbols for aspirated stops and the glottal stop are not included. The entries are the minimal number of basic sounds needed to distinguish all morphemes and words in English. For example, the one symbol [p] for all voiceless bilabial stops, together with the symbol [b] for the voiced
### Table 6.4 Minimal Set of Phonetic Symbols for American English Consonants

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Interdental</th>
<th>Alveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stop (oral)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless</td>
<td>p</td>
<td></td>
<td></td>
<td>t</td>
<td></td>
<td>k</td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td>b</td>
<td></td>
<td></td>
<td>d</td>
<td></td>
<td>g</td>
<td></td>
</tr>
<tr>
<td><strong>Nasal (stop)</strong></td>
<td>m</td>
<td></td>
<td></td>
<td>n</td>
<td></td>
<td>η</td>
<td></td>
</tr>
<tr>
<td><strong>Fricative</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless</td>
<td>f</td>
<td>θ</td>
<td>s</td>
<td>š</td>
<td></td>
<td>h₁</td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td>v</td>
<td>δ</td>
<td>z</td>
<td>ż</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Affricate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ć</td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>j</td>
<td></td>
</tr>
<tr>
<td><strong>Glide</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless</td>
<td>m</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>w²</td>
<td>h₁</td>
</tr>
<tr>
<td>voiced</td>
<td>w²</td>
<td></td>
<td></td>
<td>j</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Liquid</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>l</td>
<td>r</td>
<td></td>
</tr>
</tbody>
</table>

1. [h] is sometimes classified as a fricative because of the hissing sound produced by air or noise at the glottis. It is also sometimes classified with the glides because in many languages it combines with other sounds the way that glides do.

2. [w] is classified as both a bilabial because it is produced with both lips rounded and as a velar because the back of the tongue is raised toward the velum.

Bilabial stop, are sufficient to differentiate the word *peat* [pit] from *beat*. If a more detailed, or narrow, phonetic transcription of these words is desired, the symbol [pʰ] can be used as in [pʰɪt].

Examples of words in which these sounds occur are given in Table 6.5.

### Table 6.5 Examples of Consonants in English Words

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Interdental</th>
<th>Alveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stop (oral)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless</td>
<td><em>pie</em></td>
<td></td>
<td></td>
<td><em>tie</em></td>
<td></td>
<td><em>kite</em></td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td><em>buy</em></td>
<td></td>
<td></td>
<td><em>die</em></td>
<td></td>
<td><em>guy</em></td>
<td></td>
</tr>
<tr>
<td><strong>Nasal (stop)</strong></td>
<td><em>my</em></td>
<td></td>
<td></td>
<td><em>night</em></td>
<td></td>
<td><em>sing</em></td>
<td></td>
</tr>
<tr>
<td><strong>Fricative</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless</td>
<td><em>fie</em></td>
<td><em>thigh</em></td>
<td><em>sue</em></td>
<td><em>mission</em></td>
<td></td>
<td><em>high</em></td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td><em>vie</em></td>
<td><em>thiy</em></td>
<td><em>zoo</em></td>
<td><em>measure</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Affricate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>chime</em></td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>jive</em></td>
<td></td>
</tr>
<tr>
<td><strong>Glide</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless</td>
<td><em>which</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>which</em></td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td><em>wipe</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>yank</em></td>
<td><em>wipe</em></td>
</tr>
<tr>
<td><strong>Liquid</strong></td>
<td></td>
<td></td>
<td></td>
<td><em>lie, rye</em></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. For speakers with a voiceless *w*. 
Vowels

The quality of a vowel depends on the configuration of the vocal tract during its production. Different parts of the tongue may be high or low in the mouth; the lips may be spread or pursed; the velum may be raised or lowered. The passage through which the air travels, however, is never so narrow as to obstruct the free flow of the airstream.

Vowel sounds carry pitch and loudness; you can sing vowels. They may be long or short. Vowels can “stand alone”—they can be produced without consonants before or after them. You can say the vowels of beat, [bit], bit [bɪt], or boot [bʊt], for example, without the initial [b] or the final [t], but you cannot say a [b] or a [t] alone without at least a “little bit” of vowel sound.

Linguists and speech scientists can describe vowels acoustically, or electronically. We will discuss that in chapter 9. In this chapter, we describe vowels by their articulatory features, just as we described consonants. Many beginning students of phonetics find this method more difficult to apply to vowel articulations than to consonant articulations. When you articulate a [t], you can feel your tongue touch the alveolar ridge. When you make a [p], you can feel your two lips come together, or you can watch your lips move in a mirror. Because we produce vowels without articulators touching or even coming close, it is often difficult to feel what is happening. You may not understand at first what we mean by “front,” “back,” “high,” and “low” vowels, but we encourage you to persist. It will come.

These terms do have meaning, though. If you watch an X-ray movie of someone talking, you can see why vowels have traditionally been classified according to three questions:

1. How high is the tongue?
2. What part of the tongue is involved, and is that part up, down, or neutral in the mouth?
3. What is the position of the lips?

TONGUE POSITION

HIGGINS: Tired of listening to sounds?
PICKERING: Yes. It’s a fearful strain. I rather fancied myself because I can pronounce twenty-four distinct vowel sounds, but your hundred and thirty beat me. I can’t hear a bit of difference between most of them.
HIGGINS: Oh, that comes with practice. You hear no difference at first, but you keep on listening and presently you find they're all as different as A from B.

G. B. Shaw, *Pygmalion*

The upper two diagrams in Figure 6.4 show that the tongue is high in the mouth in the production of the vowels [i] and [u] in the words *he* [hi] and *who* [hu]. In *he* the front part (but not the tip) of the tongue is raised; in *who* it is the back of the tongue. (Prolong the vowels of these words and try to feel the raised part of your tongue.)

To produce the vowel sound [a] of *hah* [ha], the back of the tongue is low in the mouth, as the lower diagram in Figure 6.4 shows. (The reason a doctor examining your throat may ask you to say “ah” is that the tongue is low and easy to see over.) This vowel is therefore a low, back vowel.

The vowels [i] and [u] in the words *hit* [hit] and *put* [put] are similar to those in *he* [hi] and *who* [hu] with slightly lowered tongue positions.

The vowel [æ] in *hack* [hæk] is produced with the front part of the tongue low in the mouth, similar to the low vowel [a], but with the front rather than the back part of the tongue lowered. Say “hack, hah, hack, hah, hack, hah, . . .” and you should feel your tongue moving forward and back in the low part of your mouth.

The vowels [e] and [o] in *bait* [bet] and *boat* [bot] are mid vowels, produced by raising the tongue to a position midway between the high and low vowels just discussed.

---

**Figure 6.4** Position of the tongue in producing the vowels in he, who, and hah.
[e] and [o] in the words *bet* [bet] and *bore* [bør] are also mid vowels, produced with a slightly lower tongue position than [e] and [o].

To produce the vowel [a] in the word *but* [bat], the tongue is not strictly high nor low, front nor back. It is a lower mid, central vowel. The schwa vowel [ə], which occurs as the first sound in *about* [əbaut], or the final sound of *sofa* [sofə], is also articulated with the tongue in a more or less neutral position between the extremes of high/low, front/back. The schwa is used only to represent unstressed vowels. (We will discuss stress below.)

**Lip Rounding**

Vowels also differ as to whether the lips are rounded or spread. The vowels [u], [ɯ], [o], and [ɔ], in *boot*, *put*, *boat*, and *bore* are *rounded vowels*. They are produced with the lips pursed, or rounded, and the back of the tongue at decreasing heights, as shown in Figure 6.5. You can get a feel for the rounding by prolonging the word *who*, as if you were an owl: *whooooooo*.[1] Now pose for the camera and say *cheese*, only say it with a prolonged vowel: *cheeeeeeeeeeese*. The high front [i] in *cheese* is unrounded, with the lips in the shape of a smile, and you can feel it. The low vowel [a] in the words *bar*, *bah*, and *aha* is the only English back vowel that occurs without lip rounding. All nonback vowels in English are also unrounded.

This is not true of all languages. French and Swedish, for example, have both front and back rounded vowels. In English, a high back unrounded vowel does not occur, but in Mandarin Chinese, Japanese, the Cameroonian language Fe?Fe?, and many other languages, this vowel is part of the phonetic inventory of sounds. There is a Chinese word meaning “four” with an initial [s] followed by a vowel similar to the one in *boot* but with unrounded, spread lips. This Chinese word is distinguished from the word meaning “speed,” pronounced like the English word *sue*, with a high back rounded vowel.

Figure 6.5 shows the vowels based on tongue “geography.” The position of the vowel relative to the horizontal axis is a measure of the vowel’s front/back dimension.

<table>
<thead>
<tr>
<th>Tongue Height</th>
<th>FRONT</th>
<th>CENTRAL</th>
<th>BACK</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIGH</td>
<td>i</td>
<td><em>beet</em></td>
<td>boot</td>
</tr>
<tr>
<td></td>
<td>↑</td>
<td><em>bit</em></td>
<td>put</td>
</tr>
<tr>
<td>MID</td>
<td>e</td>
<td><em>bait</em></td>
<td><em>boat</em></td>
</tr>
<tr>
<td></td>
<td>↓</td>
<td><em>bet</em></td>
<td><em>Rosa</em></td>
</tr>
<tr>
<td>LOW</td>
<td>æ</td>
<td><em>bat</em></td>
<td><em>bore</em></td>
</tr>
</tbody>
</table>

*Figure 6.5* Classification of American English vowels.
Its position relative to the vertical axis is a measure of tongue height. For example, we see that [i] is a high front vowel, [o] is a mid back (rounded) vowel, and [ə] is a lower mid, central vowel, tending toward backness.

**Diphthongs**

A *diphthong* is a sequence of two sounds, vowel + glide. Diphthongs are present in the phonetic inventory of many languages, including English. The vowels we have studied so far are simple vowels, called monophthongs. The vowel sound in the word *bite* [bait], however, is the [a] vowel sound of *father* followed by the [j] glide, resulting in the diphthong [aj]. Similarly, the vowel in *bout* [bawt] is [a] followed by the glide [w], resulting in [aw]. (Some speakers of English pronounce this diphthong as [əw], with the front low unrounded vowel instead of the back vowel.) The third diphthong that occurs in English is the vowel sound in *boy* [bɔj], which is the vowel [ɔ] of *bore* (without the [r]) followed by the palatal glide [j], resulting in [ɔj].

**Nasalization of Vowels**

Vowels, like consonants, can be produced with a raised velum that prevents the air from escaping through the nose, or with a lowered velum that permits air to pass through the nasal passage. When the nasal passage is blocked, oral vowels result; when the nasal passage is open, nasal (or nasalized) vowels result. In English, nasal vowels occur for the most part before nasal consonants in the same syllable, and oral vowels occur in all other places.

The words *bean, bin, bane, been, ban, boon, bun, bone, beam, bam, boom, bing, bang, and bong* are examples of words that contain nasalized vowels. To show the nasalization of a vowel in a phonetic transcription, a *diacritic* mark [˘] (tilde) is placed over the vowel, as in *bean* [bīn] and *bone* [bōn].

In languages like French, Polish, and Portuguese, nasalized vowels occur without nasal consonants. The French word meaning “sound” is *son* [sɔ̃]. The *n* in the spelling is not pronounced but indicates that the vowel is nasal.

**Tense and Lax Vowels**

Figure 6.5 shows that the vowel [i] has a slightly higher tongue position than [ɪ]. This is also true for [ɛ] and [ɛ], [u] and [ʊ], and [o] and [ɔ]. The first vowel in each pair is generally produced with greater tension of the tongue muscles than its counterpart, and they are often a little longer in duration. These vowels can be distinguished from the shorter and less tense vowels by the phonetic features *tense* and *lax* as shown in the following:

<table>
<thead>
<tr>
<th>Tense (Longer)²</th>
<th>Lax (Shorter)</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>beat</td>
</tr>
<tr>
<td>e</td>
<td>bait</td>
</tr>
<tr>
<td>u</td>
<td>boot</td>
</tr>
<tr>
<td>o</td>
<td>boat</td>
</tr>
<tr>
<td></td>
<td>i</td>
</tr>
<tr>
<td></td>
<td>ɛ</td>
</tr>
<tr>
<td></td>
<td>ʊ</td>
</tr>
<tr>
<td></td>
<td>ɔ</td>
</tr>
</tbody>
</table>

² The term *long* refers to duration, not to the “long vowels” of traditional grammar and dictionaries, where a “long i” is [aj], and the “short i” is [ɪ], and so on.
Some speakers of English may diphthongize the tense vowels somewhat. For these speakers, the tense front vowels are followed by a short [i] glide, so [ij] and [ej] replace [i] and [e]. The tense back vowels are followed by a short [w] glide, so [uw] and [ow] replace [u] and [o]. These are sometimes written as [i], [e], [uw], and [ow]. We will continue to denote these sounds as [i], [e], [u], and [o].

**Dialect Differences**

As already mentioned, but perhaps worth repeating because of the many dialects of English, the vowels in Figure 6.5 do not represent all the vowels of all dialects of English. One dialect spoken in the United Kingdom, called British RP or "Received Pronunciation" (as the dialect spoken by the upper classes and "received" in court), has a low rounded back vowel in the word *hot* that does not occur in American English dialects, and that contrasts with the unrounded low back vowel [a] in *bah*. The long tense vowels in British RP are all diphthongs. Thus the vowel in *bay* is [ej] and the vowel in *bow* is [ow], as is true for some dialects of American English. On the other hand, in some dialects of English spoken in Ireland, these vowels are pure monophthongs. These are just a few examples of dialect differences that occur primarily in the pronunciation of vowels.
Major Classes

Biologists describe classes of life in broader or narrower terms. They may distinguish between animals and plants, or within animals, between vertebrates and invertebrates, and within vertebrates, between mammals and reptiles, and so on.

Linguists describe speech sounds in a similar manner. All sounds are consonants or vowels. Within consonants, all are voiced or unvoiced, and so on. All the classes of sounds described so far in this chapter combine to form larger, more general classes that are important in the patterning of sounds in the world’s languages.

Noncontinuants and Continuants

As we mentioned, stop sounds are **noncontinuants**. There is a total obstruction of the airstream in the oral cavity. They include the nasal stops (despite the fact that air flows continuously out the nose). All other consonants, and all vowels, are continuants, in which the stream of air flows continuously out the mouth.

Obstruents and Sonorants

The nonnasal stops, the fricatives, and the affricates form a major class of sounds called **obstruents**. The airstream may be fully obstructed, as in nonnasal stops and affricates, or partially obstructed, as in the production of fricatives.

Fricatives are continuant obstruents. The air flows continuously out the mouth, though it is obstructed enough to cause the frictional sound that characterizes this class of consonants.

Nonnasal stops and affricates are noncontinuant obstruents; there is a complete blockage of the air during the production of these sounds. The closure of a stop is released abruptly as opposed to the closure of an affricate, which is released gradually, causing friction.

Sounds that are not obstruents are **sonorants**. Sonorants are produced with relatively free airflow through either the mouth or nose. They have greater acoustic energy than obstruents. Nasal stops are sonorants because although the air is blocked in the mouth, it continues to resonate and move through the nose. Vowels, the liquids [l] and [r], and the glides [w] and [j] are also sonorants because the air resonates as it flows relatively undisturbed through the vocal tract.

Consonants and Vowels

As stated, the sounds of all the languages of the world fall into two major natural classes—consonants and vowels. Consonants include a number of subclasses: stops (including affricates and nasals), fricatives, liquids, and glides. The class of vowels include oral, nasal, front, central, back, high, mid, and low vowels.

Nasals and liquids are sonorants because they resonate; yet, they resemble the obstruents in that the oral cavity is constricted or even closed during their articulation. Linguists group nasals, liquids, and obstruents into a larger class of sounds that they (confusingly) call **consonantal**. This is done not out of perversity, but because these sounds form a natural class, as we will see in the next chapter.
While all consonantal sounds are consonants, not all consonants are consonantal. Glides, in particular, are nonconsonantal consonants. They pattern with the vowels to make up the class of nonconsonantal sounds that are sometimes referred to as vocalic sounds.

Here are some other terms used to form classes of consonants. These are not exhaustive (though they may exhaust you while learning them). A full course in phonetics would note further classes that we omit.

Labials [p] [b] [m] [f] [v] The class of labial consonants includes the class of bilabial sounds — [p] [b] [m] — as well as the labiodentals, [f] and [v]. Labial sounds are those articulated with the involvement of the lips.

Coronals [t] [d] [n] [s] [z] [ʃ] [ʒ] [ʃ] [ʒ] Coronals include the alveolars, [t] [d] [n] [s] [z]; the palatals, [ʃ] [ʒ]; the affricates, [ʃʃ] [ʒʒ]; and the liquid [l]. (Some articulations of [r] are also coronal.) These sounds are produced by raising the tongue blade.

Anterior sounds are consonants produced in the front part of the mouth, that is, from the alveolar area forward. They include the labials, the interdentals, and the alveolars.

Sibilants [s] [z] [ʃ] [ʒ] [ʃʃ] [ʒʒ] Another class of consonantal sounds is characterized by an acoustic rather than an articulatory property of its members. The friction created in the production of fricatives and affricates causes a hissing sound, which is a mixture of high-frequency sounds. These sounds form the class of sibilants.

**Syllabic Sounds**

In the following chapter, we will give a precise definition of syllable. Traditionally it has been difficult to provide such a definition, although speakers seem to be able to determine the syllabic structure of a word. From an auditory point of view, syllables have peaks of sonorance (which are also difficult to define). Every vowel is at the center of a single syllable.

Liquids and nasals can also be syllabic — function as a syllable — as shown by the words Rachel [reʃl], faker [fekr], rhythm [riðm], and button [bʌtn]. (The diacritical mark under the [l] [r] [n] and [n] show that these sounds are syllabic.) Placing a schwa [ə] before the syllabic liquid or nasal also shows that these are separate syllables. The four words could be written as [reʃəl], [fekrə], [riðm], and [bʌtn]. We will use this transcription. Similarly, the vowel sound in words like bird and verb are sometimes written as a syllabic r, [brd] and [vrb]. For consistency we shall transcribe these words using the schwa — [bərd] and [vərb] — the only instances where a schwa represents a stressed vowel.

**Prosodic Suprasegmental Features**

Speech sounds that are identical in their place or manner features may differ in length (duration), pitch, or loudness. Tense vowels are generally longer than lax vowels, but only by a small amount, perhaps a few milliseconds. (A millisecond is 1/1000 of a sec-
ond.) However, when a vowel is prolonged to around twice its normal length, it is considered in some languages a different vowel, and it can make a difference between words. In Japanese the word biru with a short i means “building,” but with the i prolonged, spelled bi:ru or biiru, the meaning is “beer.” We have just illustrated the two ways of denoting a long, or geminate, vowel: add a colon or simply write it twice.

Japanese, and many other languages such as Finnish and Italian, also have long (geminate) consonants that make a difference in words. When a consonant is long, either the closure or obstruction is prolonged. Pronounced with a short k, the word saki, means “ahead” in Japanese; pronounced with a long k — prolonging the velar closure — the word sakki means “before.”

English is not a language in which vowel or consonant length can change a word. You might say “stooooooop!” to emphasize your desire to make someone stop, but the word is not changed. You may also say in English “Whatttttt a dump!” to express your dismay at a hotel room, prolonging the t-closure, but the word what is not changed.

When we speak, we also change the pitch of our voice. The pitch depends on how fast the vocal cords vibrate; the faster they vibrate, the higher the pitch. If the larynx is small, as in women and children, the shorter vocal cords vibrate faster and the pitch is higher, all other things being equal. That is why women and children have higher pitched voices than men, in general.

In many languages, certain syllables in a word are louder, slightly higher in pitch, and somewhat longer in duration (but not geminate) than other syllables in the word. They are stressed syllables. For example, the first syllable of digest, the noun meaning “summation of articles” is stressed, while in digest, the verb meaning “to absorb food,” the second syllable receives greater stress. Stress can be marked in a number of ways: for example, by putting an accent mark over the stressed vowel in the syllable, as in digest versus digést.

English is a “stress” language. In general, at least one syllable is stressed in an English word. French is not a stress language. The syllables have approximately the same loudness, length, and pitch. When native English speakers attempt to speak French, they often stress syllables, so that native French speakers hear French with “an English accent.” When French speakers speak English, they fail to put stress where a native English speaker would, and that contributes to what English speakers would call a “French accent.”

Length, pitch, and the complex feature stress are prosodic, or suprasegmental, features. They are features over and above the segmental values such as voicing or place of articulation, thus the “supra” in suprasegmental. The term prosodic comes from poetry, where it refers to the metrical structure of verse. One of the essential characteristics of poetry is the placement of stress on particular syllables, which defines the versification of the poem.

TONE AND INTONATION

We have already seen how length and stress can make sounds with the same segmental properties different. In some languages, these differences make different words, such as the two digest. Pitch, too, can make a difference in certain languages.
Speakers of all languages vary the pitch of their voices when they talk. The effect of pitch on a syllable differs from language to language. In English, it doesn’t matter whether you say cat with a high pitch or a low pitch. It will still mean “cat.” But if you say [ba] with a high pitch in Nupe (a language spoken in Nigeria), it will mean “to be sour,” whereas if you say [ba] with a low pitch, it will mean “to count.” Languages that use the pitch of individual vowels or syllables to contrast meanings of words are called tone languages.

The majority of the languages in the world are tone languages. There are more than one thousand tone languages in Africa alone. Many languages of Asia, such as Chinese, Thai, and Burmese, are tone languages, as are many Native American languages.

Thai is a language that has contrasting pitches or tones. The same string of segmental sounds represented by [naa] will mean different things if one says the sounds with a low pitch, a mid pitch, a high pitch, a falling pitch from high to low, or a rising pitch from low to high. Thai therefore has five linguistic tones.

- [naa] low tone “a nickname”
- [naa] mid tone “rice paddy”
- [naa] high tone “young maternal uncle or aunt”
- [naa] falling tone “face”
- [naa] rising tone “thick”

Diacritics are used to represent distinctive tones in the phonetic transcriptions.

- [\^] L low tone
- [\_\_] M mid tone
- [\^] H high tone
- [\_\_\_] HL falling tone (High to Low)
- [\^\_-\_] LH rising tone (Low to High)

We can use these diacritics placed above the vowels to represent the tonal contrasts in any language where the pitch of the vowel is important in conveying meaning as illustrated by the three contrastive tones in Nupe:

- H M L

Akan, sometimes called Twi, the major language of Ghana, has two tones, which are shown in these contrasting two-syllables words.

- dù à [\_\_] “tail”
- [\_-\_] “tree”
- L L
- kò tó [\_-\_] “go buy”
- kò tó [\_-\_] “crab”
- L H H L
In some tone languages the pitch of each tone is level; in others, the direction of the pitch (whether it glides from high to low or from low to high) is important. Tones that glide are called contour tones; tones that do not are called level, or register, tones. The contour tones of Thai are represented by using a high tone followed by a low tone for a falling glide, and a low followed by a high for a rising tone.

In a tone language, it is not the absolute pitch of the syllables that is important but the relations among the pitches of different syllables. After all, some individual speakers have high-pitched voices, others low-pitched, and others medium-pitched. In many tone languages we find a falling-off of the pitch, a continual downdrift of the tones.

In the following sentence in Twi, the relative pitch rather than the absolute pitch is important.

"Kofi searches for a little food for his friend’s child."

Kòfì hwèhwé áduàn kàkrá mà n’ ádámfo bá

L H L H H L L H L L H

The actual pitches of these syllables would be rather different from each other, as shown in the following musical staff-like figure (the higher the number, the higher the pitch):

7  fì
6  hwè  á
5  Kò  krá
4  hwè  á
3  duàn  kà  bá
2  mà  n’
1  dámfo

The lowering of the pitch is called downdrift. In languages with downdrift — and many tone languages in Africa are downdrift languages — a high tone that occurs after a low tone, or a low tone after a high tone, is lower in pitch than the preceding similarly marked tone. Notice that the first high tone in the sentence is given the pitch value 7. The next high tone (which occurs after an intervening low tone) is 6; that is, it is lower in pitch than the first high tone.

This example shows that in analyzing tones, just as in analyzing segments, all the physical properties need not be considered; only essential features are important in language — in this case, whether the tone is “high” or “low” in relation to the other pitches, but not the specific pitch of that tone.

Languages that are not tone languages, such as English, are called intonation languages. The pitch contour of the utterance varies, but in an intonation language as opposed to a tone language, pitch is not used to distinguish words from each other.
Diacritics

In the discussions of vowel nasalization, prosodic features, and tone, we presented a number of diacritic marks that modify the basic phonetic symbols. A tilde [~] over the vowel symbol means nasalization; a colon [:] after the vowel symbol marks length or gemination; an acute accent shows stress; and various other marks indicate the various tones.

Other diacritics provide additional ways of showing phonetic differences of speech sounds. For example, to differentiate the voiceless lateral liquid that occurs in some pronunciations of words like place [ples], or the voiceless alveolar nasal that may occur in a word like snow [sno], a little round “o” is written under the symbol, as you can see.

Linguists often use cover symbols to refer to classes of sounds. A capital C represents the class of consonants, V the class of vowels, G the glides, and L the liquids. (It’s true. Linguists are just as lazy as the government bureaucrats who make up all those acronyms.)

We can summarize these diacritics and additional symbols as follows:

\[
\begin{align*}
C &= \text{Consonant} & C: &= \text{long C} \\
V &= \text{Vowel} & V: &= \text{long V} \\
\bar{V} &= \text{nasalized V} & \bar{V} &= \text{stressed V} \\
L &= \text{Liquid} & L^c &= \text{voiceless L} \\
G &= \text{Glide} & G^c &= \text{voiceless glide} \\
\text{Tones} &= & \hat{V} &= \text{High} \\
\hat{V} &= \text{Rising} & \hat{V} &= \text{Low} \\
\bar{V} &= \text{Falling} & \bar{V} &= \text{Mid} \\
\end{align*}
\]

Phonetic Symbols and Spelling Correspondences

Table 6.6 shows the sound/spelling correspondences for American English consonants and vowels. (We have not given all possible spellings for every sound, however these examples should help you relate English orthography to the English sound system.) We have included the symbols for the voiceless aspirated stops to illustrate that what speakers usually consider one sound — for example [p] — may phonetically be two (or more) sounds, [p], [pʰ].

Some of these pronunciations may differ from yours, making some of the examples confusing. For example, as we mentioned, some speakers of American English pronounce the words cot and caught identically. In the dialect described here, cot and caught are pronounced differently, so cot is one of the examples of the vowel sound [a]. Many speakers who pronounce cot and caught the same pronounce car and core with
different vowels. If you use the vowel of car to say cot and the vowel of core to say caught, you will be approximating the dialect that distinguishes the two words. In addition, some speakers of English pronounce an r sound only when it occurs before a vowel, and would therefore pronounce the words car and fear without the [r], although we mention fear as one example of [r].

The English used for examples in this book is rather arbitrary. It is in fact a mixture of several dialects. Our aim is to teach phonetics in general, and to show you how phonetics might describe the speech sounds of any of the world’s languages with the proper symbols and diacritics. We illustrate how to do this using American English, and providing the major phonetic symbols for describing most varieties of American English. We are aware that this may present problems for speakers of different dialects. We apologize for this, but we have not figured out a way to solve this problem satisfactorily without introducing complexities unsuitable for this introduction to language.

The symbols listed in Table 6.6 are not sufficient to represent the pronunciation of words in all languages, or even all dialects of English around the world. The symbol [x], for example, is needed for the voiceless velar fricative in the German word Bach, or the Scottish English word for “lake,” loch. The symbol [ʁ] is needed for the French uvular
Table 6.6  Phonetic Symbol/English Spelling Correspondences

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>spit tip apple ample</td>
</tr>
<tr>
<td>ph</td>
<td>pit prick plaque appear</td>
</tr>
<tr>
<td>b</td>
<td>bit tab brat bubble</td>
</tr>
<tr>
<td>m</td>
<td>mitt tam smack Emmy camp comb</td>
</tr>
<tr>
<td>t</td>
<td>stick pit kissed write</td>
</tr>
<tr>
<td>th</td>
<td>tick intend pterodactyl attack</td>
</tr>
<tr>
<td>d</td>
<td>Dick cad drip loved ride</td>
</tr>
<tr>
<td>n</td>
<td>nick kin snow mnemonic gnostic pneumatic know</td>
</tr>
<tr>
<td>k</td>
<td>skin stick scat critique elk</td>
</tr>
<tr>
<td>kh</td>
<td>curl kin charisma critic mechanic close</td>
</tr>
<tr>
<td>g</td>
<td>girl burg longer Pittsburgh</td>
</tr>
<tr>
<td>ng</td>
<td>sing think finger</td>
</tr>
<tr>
<td>f</td>
<td>fat philosophy flat phlogiston coffee reef cough</td>
</tr>
<tr>
<td>v</td>
<td>vat dove gravel</td>
</tr>
<tr>
<td>s</td>
<td>sip skip psychology pass pats democracy scissors fasten deceive descent</td>
</tr>
<tr>
<td>z</td>
<td>zip jazz razor pads kisses Xerox design lazy scissors maize</td>
</tr>
<tr>
<td>th</td>
<td>thigh through wrath ether Matthew</td>
</tr>
<tr>
<td>ð</td>
<td>thy their weather lathe either</td>
</tr>
<tr>
<td>ñ</td>
<td>shoe mush mission nation fish glacial sure</td>
</tr>
<tr>
<td>õ</td>
<td>measure vision azure casual decision rouge (for those who do not pronounce this word with the final sound of judge)</td>
</tr>
<tr>
<td>ċ</td>
<td>choke match feature rich righteous</td>
</tr>
<tr>
<td>j</td>
<td>judge midget George magistrate residual</td>
</tr>
<tr>
<td>l</td>
<td>leaf feel call single</td>
</tr>
<tr>
<td>r</td>
<td>reef fear Paris singer</td>
</tr>
<tr>
<td>j</td>
<td>you yes feud use</td>
</tr>
<tr>
<td>w</td>
<td>witch swim queen</td>
</tr>
<tr>
<td>w̃</td>
<td>which where whale (for speakers who pronounce which differently than witch)</td>
</tr>
<tr>
<td>h</td>
<td>hat who whole rehash</td>
</tr>
<tr>
<td>ß</td>
<td>bottle button glottal (for some speakers)</td>
</tr>
</tbody>
</table>
Table 6.6  (Continued)

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>beet beat be receive key believe amoeba people Caesar Vaseline serene</td>
</tr>
<tr>
<td>i</td>
<td>bit consist injury bin</td>
</tr>
<tr>
<td>e</td>
<td>bate bait ray great eight gauge reign they</td>
</tr>
<tr>
<td>e</td>
<td>bet serenity says guest dead said</td>
</tr>
<tr>
<td>æ</td>
<td>pan act laugh comrade</td>
</tr>
<tr>
<td>u</td>
<td>boot lute who sewer through to too two move Lou</td>
</tr>
<tr>
<td>u</td>
<td>put foot butcher could</td>
</tr>
<tr>
<td>a</td>
<td>cut tough among oven does cover flood</td>
</tr>
<tr>
<td>o</td>
<td>coat go beau grow though toe own over</td>
</tr>
<tr>
<td>ɔ</td>
<td>caught stalk core saw ball awe</td>
</tr>
<tr>
<td>a</td>
<td>cot father palm sergeant honor hospital melodic</td>
</tr>
<tr>
<td>ø</td>
<td>sofa alone symphony suppose melody tedious the America</td>
</tr>
<tr>
<td>aj</td>
<td>bite sight by die dye Stein aisle choir liar island height sign</td>
</tr>
<tr>
<td>aw, æw</td>
<td>about brown doubt coward</td>
</tr>
<tr>
<td>æj</td>
<td>boy doily</td>
</tr>
</tbody>
</table>

Fricative. English does not have rounded front vowels, but languages such as French and Swedish do. Front rounded vowels are symbolized as follows:

- [y] as in French tu [ty] “you” The tongue position for [i] but with the lips rounded
- [ø] as in French bleu [blø] “blue” The tongue position for [e] but with the lips rounded
- [œ] as in French heure “hour” The tongue position for [e] but with the lips rounded

**Sign-Language Primes**

Just as sign languages have their own morphological, syntactic, and semantic systems, they also have their equivalent of phonetics and phonology. The formal units corresponding to phonetic elements of spoken language are referred to as **primes**. The signs of the language that correspond to morphemes or words can be specified by primes of three classes: hand configuration; the motion of the hand(s) toward or away from the
body; and the place of articulation, or the locus, of the sign's movement relative to the body. For example, the sign meaning "arm" is a flat hand, moving to touch the upper arm. It has three prime features: flat hand, motion upward, upper arm.

Figure 6.6 illustrates the hand-configuration primes.

Figure 6.6  ASL hand configuration (with descriptive phrases that are used to refer to them). The letters and numbers refer to the signs used for these symbols when words are finger-spelled. Reprinted from The Signs of Language by Edward Klima and Ursula Bellugi. Cambridge, Mass.: Harvard University Press, Copyright © 1979 by the President and Fellows of Harvard College.
Summary

The science of speech sounds is called phonetics. It aims to provide the set of features or properties to describe and distinguish all the sounds in human languages throughout the world.

When we speak, the physical sounds we produce are continuous stretches of sound, which are the physical representations of strings of discrete linguistic segments. Knowledge of a language permits one to separate the continuous sound into linguistic units — words, morphemes, and sounds.

The discrepancy between spelling and sounds in English and other languages motivated the development of phonetic alphabets in which one letter corresponds to one sound. The major phonetic alphabet in use is that of the International Phonetic Association (IPA), which includes modified Roman letters and diacritics by means of which the sounds of all human languages can be represented. To distinguish between the orthography, or spelling, of words, and their pronunciations, we write phonetic transcriptions between square brackets, as in [fonetik] for phonetic.

All English speech sounds come from the movement of lung air through the vocal tract. The air moves through the glottis, or between vocal cords, up the pharynx, through the oral (and possibly the nasal) cavity, and out the mouth or nose. Other languages may use different airstream mechanisms.

Human speech sounds fall into classes according to their phonetic properties. All speech sounds are either consonants or vowels, and all consonants are either obstruents or sonorants. Consonants have some obstruction of the airstream in the vocal tract, and the location of the obstruction defines their place of articulation, some of which are bilabial, labiodental, alveolar, palatal, velar, uvular, and glottal.

Consonants are further classified according to their manner of articulation. They may be voiced or voiceless, oral or nasal; long or short. They may be stops, fricatives, affricates, liquids, or glides. During the production of voiced sounds, the vocal cords are together and vibrating, whereas in voiceless sounds they are apart and not vibrating. Voiceless sounds may also be aspirated or unaspirated. In the production of aspirated sounds, the vocal cords remain apart for a brief time after the stop closure is released, resulting in a puff of air at the time of the release. Consonants may be grouped according to certain features to form larger classes such as labials, coronals, anteriors, and sibilants.

Vowels form the nucleus of syllables. They differ according to the position of the tongue and lips: high, mid, or low tongue; front, central, or back of the tongue; rounded or unrounded lips. The vowels in English may be tense or lax. Tense vowels are slightly longer in duration than lax vowels. Vowels may also be stressed (longer, higher in pitch, and louder) or unstressed. Vowels, like consonants, may be nasal or oral, though most vowels in all languages are oral.

Length, pitch, loudness, and stress are prosodic, or suprasegmental, features. They are imposed over and above the segmental values of the sounds in a syllable.

In many languages, the pitch of the vowel or syllable is linguistically significant. For example, two words may contrast in meaning if one has a high pitch and another a low pitch. Such languages are tone languages. There are also intonation languages in
which the rise and fall of pitch may contrast meanings of sentences. In English the statement *Mary is a teacher* will end with a fall in pitch, but as a question, *Mary is a teacher?* the pitch will rise.

English and other languages use **stress** to distinguish different words, such as *content* and *contén*. In some languages, long vowels and long consonants contrast with their shorter counterparts. Thus *biri* and *bitru* (*bi:*ru), *saki* and *sakki* are different words in Japanese. Long sounds are sometimes referred to as **geminates**.

Diacritics to specify such properties as **nasalization**, **length**, **stress**, and **tone** may be combined with the phonetic symbols for more detailed phonetic transcriptions. A phonetic transcription of *main* would use a tilde diacritic to indicate the nasalization of the vowel: [m̩en]

In sign languages, instead of phonetic features there are three classes of **primes**—hand configuration, the motion of the hand(s) toward or away from the body, and the place of articulation, or the locus, of the sign’s movements.

---

**References for Further Reading**


---

**Exercises**

1. Write the phonetic symbol for the first sound in each of the following words according to the way you pronounce it.

   *Examples: ooze* [u]  *psycho* [s]

   a. judge [ ]  f. thought [ ]

   b. Thomas [ ]  g. contact [ ]

   c. though [ ]  h. phone [ ]

   d. easy [ ]  i. civic [ ]

   e. pneumonia [ ]  j. usual [ ]
2. Write the phonetic symbol for the last sound in each of the following words.

*Example:* boy [ɔ] (Diphthongs should be treated as one sound.)

a. fleece [ ]

b. neigh [ ]

c. long [ ]

d. health [ ]

e. watch [ ]

f. cow [ ]

g. rough [ ]

h. cheese [ ]

i. bleached [ ]

j. rags [ ]

3. Write the following words in phonetic transcription, according to your pronunciation.

*Examples:* knot [ŋt], delightful [dɪliətfl] or [dəliətfl]. Some students may pronounce a number of words identically.

a. physics

b. merry

c. marry

d. Mary

e. yellow

f. sticky

g. transcription

h. Fromkin

i. tease

j. weather

k. coat

l. Rodman

m. heath

n. “your name”

4. Below is a phonetic transcription of a verse in the poem "The Walrus and the Carpenter" by Lewis Carroll. The speaker who transcribed it may not have exactly the same pronunciation as you; there are many correct versions. However, there is one major error in each line that is an impossible pronunciation for any American English speaker. The error may consist of an extra symbol, a missing symbol, or a wrong symbol in the word. Note that the phonetic transcription that is given is a narrow transcription; aspiration is marked, as is the nasalization of vowels. This is to illustrate a detailed transcription. However, none of the errors involve aspiration or nasalization of vowels.

Write the word in which the error occurs in the correct phonetic transcription.

**Corrected Word**

a. ðə tʰəʃjəm hæz cəm

b. ðə wələrəs sed

c. tʰu tʰələk əv məni əθəz

d. əv suz ənd ʃɪps

e. ænd sɪliŋ wæx

f. əv kʰæbəgæz ænd kʰəŋz

g. ænd waj ðə si ts əθəlin hæt

h. ænd wəθər pʰɪŋz hæv əθə ̅wəŋz

5. The following are all English words written in phonemic transcription. Write the words using normal English orthography.

a. /hit/

b. /strək/

c. /fez/
d. /ton/
e. /boni/
f. /skrim/
g. /frut/
h. /pričar/
i. /krak/

6. Write the symbol that corresponds to each of the following phonetic descriptions, then give an English word that contains this sound.

*Example:* voiced alveolar stop [d] *dough*

a. voiceless bilabial unaspirated stop [ ]
b. low front vowel [ ]
c. lateral liquid [ ]
d. velar nasal [ ]
e. voiced interdental fricative [ ]
f. voiceless affricate [ ]
g. palatal glide [ ]
h. mid lax front vowel [ ]
i. high back tense vowel [ ]
j. voiceless aspirated alveolar stop [ ]

7. In each of the following pairs of words, the bold italicized sounds differ by one or more phonetic properties (features). Give the symbol for each italicized sound, state their differences and, in addition, state what properties they have in common.

*Example:* phone — phonic

The o in *phone* is mid, tense, round. The o in *phonic* is low, unround.

Both are back vowels.

a. bath — bathe
b. reduce — reduction
c. cool — cold
d. wife — wives
e. cats — dogs
f. impolite — indecent

8. Write a phonetic transcription of the italicized words in the following poem entitled “English” published long ago in a British newspaper.

I take it you already *know*

Of *tough* and *bough* and *cough* and *dough*?

Some may stumble, but not *you*,

On *hiccough*, *thorough*, *slough* and *through*?

So now you are ready, perhaps,

To learn of less familiar traps?

Beware of *heard*, a dreadful *word*
That looks like *beard* and sounds like *bird*.
And *dead*, it’s *said* like *bed*, not *bead*;
For goodness’ sake, don’t call it *deed*!
Watch out for *meat* and *great* and *threat*.
(They rhyme with *suite* and *straight* and *debt.*)
A *moth* is not a moth in *mother*,
Nor *both* in *bother*, *broth* in *brother*.

9. For each group of sounds listed, state the phonetic feature(s) they all share.
   *Example:* [p] [b] [m] Features: bilabial, stop, consonant
   a. [g] [p] [t] [d] [k] [b]
   b. [u] [u] [ɒ] [ɒ]
   c. [i] [i] [ɛ] [ɛ] [æ]
   d. [t] [s] [ʃ] [p] [k] [ɛ] [f] [h]
   e. [v] [z] [ʒ] [j] [n] [ɡ] [d] [l] [r] [w] [j]
   f. [t] [d] [s] [ʃ] [n] [ɛ] [j]

10. Write the following sentences in regular English spelling.
   a. nom čamski iz e lingwist hu tičaz æt em aj ti
   b. fonetiks iz da stadi æv spič sawndz
   c. ol spoken længwirz juz sawndz prædust baj ðæ apær respiratori sistem
   d. m wan dajalekt æv inglæ kat ðæ nawm ænd kot ðæ varb ær pronawnst ðæ sem
   e. sam pipæl þæm fonetiks iz veri intærestæ
   f. viktornæ jærmæ rabært radmæn ænd nine hajæmz ær ðæ Chern æv ðæs buk.

11. What phonetic property or feature distinguishes the sets of sounds in column A from those in Column B?

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. [i] [i]</td>
<td>[u] [u]</td>
</tr>
<tr>
<td>b. [p] [t] [k] [s] [f]</td>
<td>[b] [d] [g] [z] [v]</td>
</tr>
<tr>
<td>c. [p] [b] [m]</td>
<td>[t] [d] [n] [k] [ɡ] [ŋ]</td>
</tr>
<tr>
<td>d. [i] [i] [u] [u]</td>
<td>[ɛ] [ɛ] [o] [ɔ] [æ] [a]</td>
</tr>
<tr>
<td>e. [f] [v] [s] [z] [ʃ] [ʒ]</td>
<td>[ɛ] [j]</td>
</tr>
<tr>
<td>f. [i] [i] [ɛ] [æ] [ɛ] [æ]</td>
<td>[u] [u] [ɔ] [ɔ] [a]</td>
</tr>
</tbody>
</table>

12. A. Which of the following sound pairs have the same manner of articulation?
   i. [h] [ʔ]      vi. [f] [ʃ]
   ii. [ɾ] [w]     vii. [k] [θ]
   iii. [m] [ŋ]    viii. [s] [ɡ]
   iv. [ʌ] [v]     ix. [j] [w]
   v. [t] [t]      x. [j] [j]

B. For each sound in part A, identify the manner of articulation.