A COURSE IN PHONETICS

Fourth Edition

PETER LADEFOGED
University of California, Los Angeles

HEINLE & HEINLE

THOMSON LEARNING

Australia Canada Mexico Singapore Spain United Kingdom United States

HEINLE & HEINLE

TM

THOMSON LEARNING

A Course in Phonetics/Fourth Edition
Ladefoged

ARTICULATORY PHONETICS

Phonetics is concerned with describing speech. There are many different reasons for wanting to describe speech, which means that there are many different kinds of phoneticians. Some are concerned with the sounds that occur in the languages of the world. Others are more concerned with helping people speak a particular form of English. Yet others are looking for ways to make computers talk more intelligibly, or to recognize whatever is said to them. For all these purposes phoneticians need to find out what people are doing when they are talking and how the sounds of speech can be described.

THE VOCAL ORGANS

We will begin by describing how speech sounds are made. In nearly all speech sounds, the basic source of power is the respiratory system pushing air out of the lungs. Try to talk while breathing in instead of out. You will find that you can do it, but it is much harder than talking when breathing out.

Air from the lungs goes up the windpipe (the trachea, to use the more technical term) and into the larynx, at which point it must pass between two small muscular folds called the vocal folds. If the vocal folds are apart, as they normally are when breathing out, the air from the lungs will have a relatively free passage into the pharynx and the mouth. But if the vocal folds are adjusted so that there is only a narrow passage between them, the airstream will cause them to vibrate. Sounds produced when the vocal folds are vibrating are said to be **voiced,** as opposed to those in which the vocal folds are apart, which are said to be **voiceless.**

In order to hear the difference between a voiced and a voiceless sound, try saying a long v sound, which we will symbolize as [vvvvv]. Now compare this with a long f sound [fffff], saying each of them alternately—[fffffvvvvvfffff-vvvvv]. Both of these sounds are formed in the same way in the mouth. The difference between them is that [v] is voiced but [f] is voiceless. You can feel the vocal fold vibrations in [v] if you put your fingertips against your larynx. You can also hear the buzzing of the vibrations in [v] more easily if you stop up your ears while contrasting [fffffvvvvv].

The difference between voiced and voiceless sounds is often important in distinguishing sounds. In each of the pairs of words "fat, vat; thigh, thy; Sue, zoo" the first consonant in the first word of each pair is voiceless, whereas in the second word, it is voiced. To check this for yourself, say just the consonant at the beginning of each of these words and try to feel and hear the voicing as suggested above. Try to find other pairs of words that are distinguished by one having a voiced and the other having a voiceless consonant.

The air passages above the larynx are known as the **vocal tract.** Figure 1.1 shows their location within the head (actually within my head, many years ago). The shape of the vocal tract is a very important factor in the production of

---

FIGURE 1.1   The vocal tract.
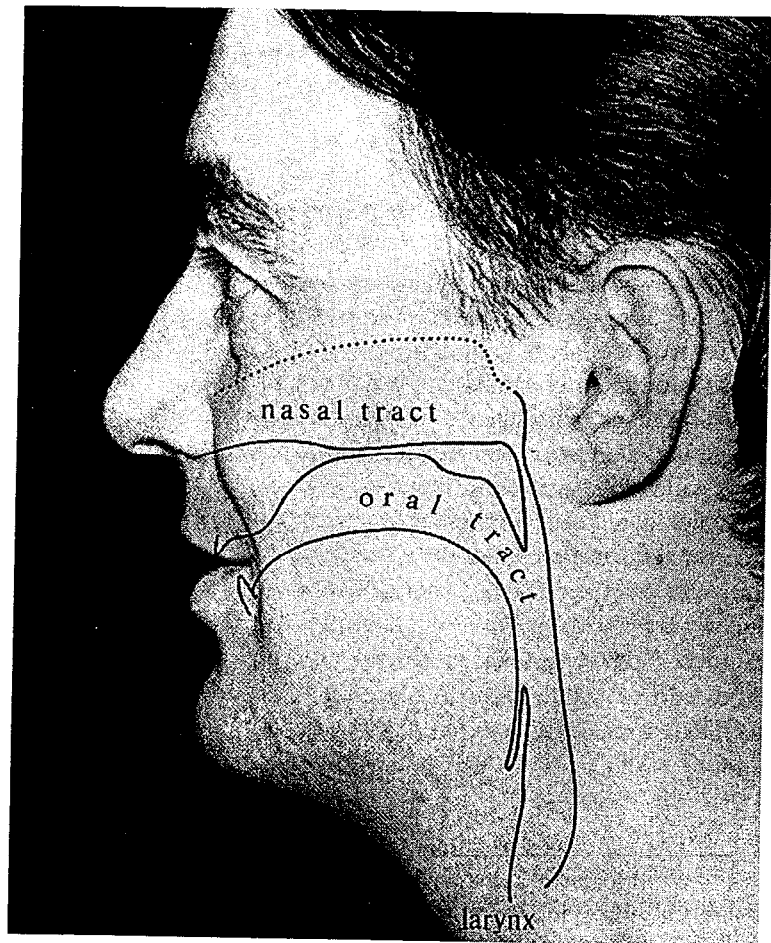
speech, and we will often refer to a diagram of the kind that has been superimposed on the photograph in Figure 1.1. Learn to draw the vocal tract by tracing the diagram in this figure. Note that the air passages that make up the vocal tract may be divided into the oral tract within the mouth and pharynx, and the nasal tract within the nose. The upper limit of the nasal tract has been marked with a dotted line since the exact boundaries of the air passages within the nose depend on soft tissues of variable size.

The parts of the vocal tract that can be used to form sounds are called articulators. The articulators that form the lower surface of the vocal tract often move toward those that form the upper surface. Try saying the word "capital" and note the major movements of your tongue and lips. You will find that the back of the tongue makes contact with the roof of the mouth for the first sound and then comes down for the following vowel. The lips come together in the formation of p and then come apart again in the vowel. The tongue tip comes up for the t and again, for most people, for the final l.

The names for the principal parts of the upper surface of the vocal tract are given in Figure 1.2. The upper lip and the upper teeth (notably the frontal incisors) are familiar enough structures. Just behind the upper teeth is a small protuberance that you can feel with the tip of the tongue. This is called the **alveolar ridge.** You can also feel that the front part of the roof of the mouth is formed by a bony structure. This is the **hard palate.** You will probably have to use a fingertip to feel further back. Most people cannot curl the tongue up far enough to touch the **soft palate,** or **velum,** at the back of the mouth. The soft palate is a muscular flap that can be raised to press against the back wall of the pharynx and shut off the nasal

| FIGURE 1.2 | The principal parts of the upper surface of the vocal tract. |
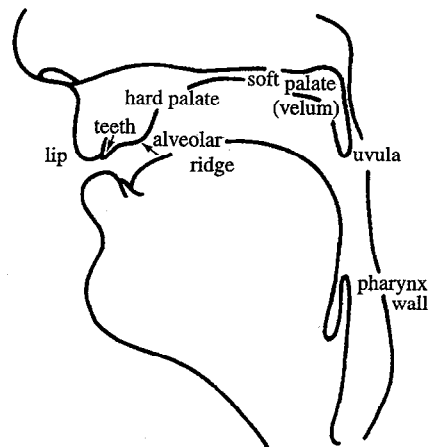
tract, preventing air from going out through the nose. In this case there is said to be a **velic closure.** This action separates the nasal tract from the oral tract so that the air can go out only through the mouth. At the lower end of the soft palate is a small appendage hanging down that is known as the uvula. The part of the vocal tract between the uvula and the larynx is the pharynx. The back wall of the pharynx may be considered to be one of the articulators on the upper surface of the vocal tract.

Figure 1.3 shows the lower lip and the specific names for different parts of the tongue that form the lower surface of the vocal tract. The tip and blade of the tongue are the most mobile parts. Behind the blade is what is technically called the front of the tongue: it is actually the forward part of the body of the tongue, and it lies underneath the hard palate when the tongue is at rest. The remainder of the body of the tongue may be divided into the center, which is partly beneath the hard palate and partly beneath the soft palate; the back, which is beneath the soft palate; and the root, which is opposite the back wall of the pharynx. The epiglottis is attached to the lower part of the root of the tongue.

Bearing all these terms in mind, say the word "peculiar" and try to give a rough description of the actions of the vocal organs during the consonant sounds. You should find that the lips come together for the first sound. Then the back and center of the tongue are raised. But is the contact on the hard palate or on the velum? (For most people, it is centered between the two.) Then note the position in the formation of the *l.* Most people make this sound with the tip of the tongue on the alveolar ridge.

Now compare the words "true" and "tea." In which word is the tongue contact further forward in the mouth? Most people make contact with the tip or blade of

FIGURE 1.3   The principal parts of the lower surface of the vocal tract.
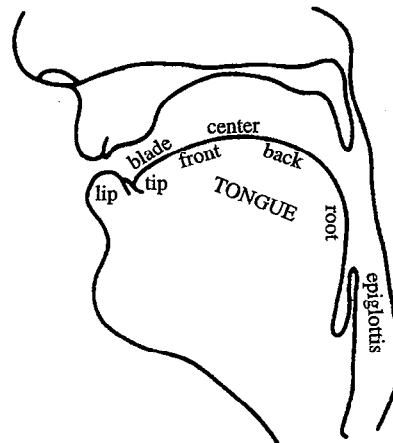
the tongue on the alveolar ridge when saying "tea," but slightly farther back in "true." Try to distinguish the differences in other consonant sounds, such as those in "sigh" and "shy" and those at the beginning of "fee" and "thief."

When considering diagrams such as those we have been discussing, it is important to remember that they show only two dimensions. The vocal tract is a tube, and the positions of the sides of the tongue may be very different from that of the center. In saying "sigh," for example, there is a deep hollow in the center of the tongue that is not present when saying "shy." We cannot represent this difference in a two-dimensional diagram showing just the midline of the tongue—a so-called mid-sagittal view. We will be relying on mid-sagittal diagrams of the vocal organs to a considerable extent in this book. But we should never let this simplified view become the sole basis for our conceptualization of speech sounds.

PLACES OF ARTICULATION

In order to form consonants, the airstream through the vocal tract must be obstructed in some way. Consonants can be classified according to the place and manner of this obstruction. The primary articulators that can cause an obstruction in most languages are the lips, the tongue tip and blade, and the back of the tongue. Speech gestures using the lips are called **labial** articulations; those using the tip or blade of the tongue are called **coronal** articulations; and those using the back of the tongue are called **dorsal** articulations.

If we do not need to specify the place of articulation in great detail, then the articulators for the consonants of English (and of many other languages) can be described using these terms. The word "topic," for example, begins with a coronal consonant; in the middle there is a labial consonant; and at the end a dorsal consonant. (Check this by feeling that the tip or blade of your tongue is raised for the first, coronal, consonant; your lips close for the second, labial, consonant; and the back of the tongue is raised for the final, dorsal, consonant.)

These terms, however, do not specify articulatory gestures in sufficient detail for many phonetic purposes. More specific places of articulation are indicated by the arrows going from one of the lower articulators to one of the upper articulators in Figure 1.4. Because there are so many possibilities in the coronal region, this area has been shown in more detail on the right of the figure. The principal terms for the particular types of obstruction required in the description of English are as follows:

1. **Bilabial**

   (Made with the two lips.) Say words such as "pie, buy, my" and note how the lips come together for the first sound in each of these words. Find a comparable set of words with bilabial sounds at the end.

FIGURE 1.4 A sagittal section of the vocal tract, showing the places of articulation that occur in English. The coronal region is shown in more detail on the right.
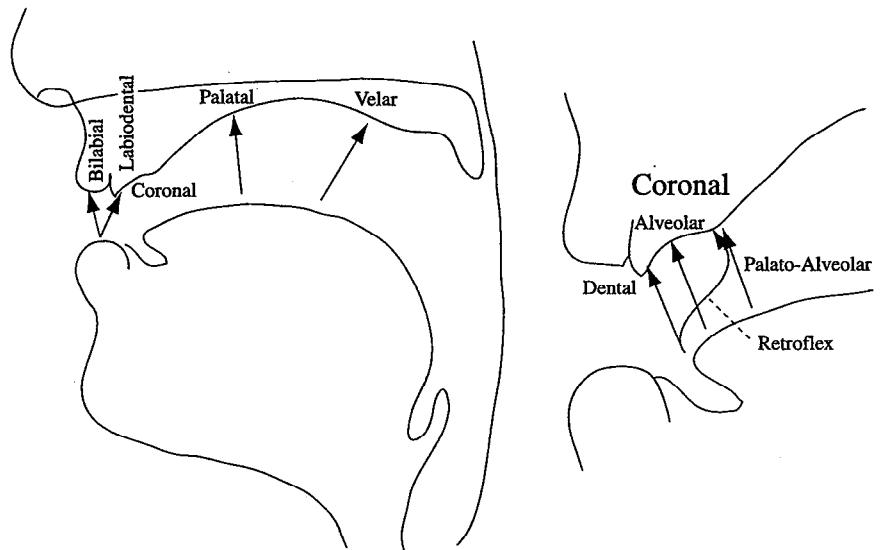
2. Labiodental

(Lower lip and upper front teeth.) Most people, when saying words such as "fie, vie," raise the lower lip until it nearly touches the upper front teeth.

3. Dental

(Tongue tip or blade and upper front teeth.) Say the words "thigh, thy." Some people (most speakers of American English) have the tip of the tongue protruding between the upper and lower front teeth; others (most speakers of British English) have it close behind the upper front teeth. Both these kinds of sounds are normal in English, and both may be called dental. If a distinction is needed, sounds in which the tongue protrudes between the teeth may be called **interdental.**

4. Alveolar

(Tongue tip or blade and the alveolar ridge.) Again there are two possibilities in English, and you should find out which you use. You may pronounce words such as "tie, die, nigh, sigh, zeal, lie" using the tip of the

tongue or the blade of the tongue. Feel how you normally make the alveolar consonants in each of these words, and then try to make them in the other way. A good way to appreciate the difference between dental and alveolar sounds is to say "ten" and "tenth" (or "n" and "nth"). Which *n* is farther back? (Most people make the one in the first of each of these pairs of words on the alveolar ridge and the second as a dental sound with the tongue touching the upper front teeth.)

5. **Retroflex**

(Tongue tip and the back of the alveolar ridge.) Many speakers of English do not use retroflex sounds at all. But for some, retroflex sounds occur initially in words such as "rye, row, ray." Note the position of the tip of your tongue in these words. Speakers who pronounce *r* at the ends of words may also have retroflex sounds with the tip of the tongue raised in "ire, hour, air."

6. **Palato-Alveolar**

(Tongue blade and the back of the alveolar ridge.) Say words such as "shy, she, show." During the consonants, the tip of your tongue may be down behind the lower front teeth, or it may be up near the alveolar ridge, but the blade of the tongue is always close to the back part of the alveolar ridge. Because these sounds are made further back in the mouth than those in "sigh, sea, sew," they can also be called **post-alveolar.** You should be able to pronounce them with the tip or blade of the tongue. Try saying "shipshape" with your tongue tip up on one occasion and down on another. Note that the blade of the tongue will always be raised. You may be able to feel the place of articulation more distinctly if you hold the position while taking in a breath through the mouth. The incoming air cools the blade of the tongue and the back part of the alveolar ridge.

7. **Palatal**

(Front of the tongue and hard palate.) Say the word "you" very slowly so that you can isolate the consonant at the beginning. If you say this consonant by itself, you should be able to feel that the front of the tongue is raised toward the hard palate. Try to hold the consonant position and breathe inward through the mouth. You will probably be able to feel the rush of cold air between the front of the tongue and the hard palate.

8. **Velar**

(Back of the tongue and soft palate.) The consonants that have the farthest back place of articulation in English are those that occur at the end of "hack, hag, hang." In all these sounds, the back of the tongue is raised so that it touches the velum.

As you can tell from the descriptions of these articulations, the first two, bilabial and labiodental, can be classified as labial, involving at least the lower lip; the next four, dental, alveolar, retroflex, and palato-alveolar (post-alveolar), are

coronal articulations, with the tip or blade of the tongue raised; and the last, velar, is a dorsal articulation, using the back of the tongue. Palatal sounds are sometimes classified as coronal articulations and sometimes as dorsal articulations, a point to which we shall return.

To get the feeling of different places of articulation, consider the consonant at the beginning of each of the following words: "fee, theme, see, she." Say these consonants by themselves. Are they voiced or voiceless? Now note that the place of articulation moves back in the mouth in making this series of voiceless consonants, going from labiodental, through dental and alveolar, to palato-alveolar.

THE ORO-NASAL PROCESS

Consider the consonants at the ends of "rang, ran, ram." When you say these consonants by themselves, note that the air is coming out through the nose. In the formation of these sounds, the point of articulatory closure moves forward, from velar in "rang," through alveolar in "ran," to bilabial in "ram." In each case, the air is prevented from going out through the mouth, but is able to go out through the nose because the soft palate, or velum, is lowered.

In most speech, the soft palate is raised so that there is a velic closure. When it is lowered and there is an obstruction in the mouth, we say that there is a nasal consonant. Raising or lowering the velum controls the oro-nasal process, the distinguishing factor between oral and nasal sounds.

MANNERS OF ARTICULATION

At most places of articulation there are several basic ways in which articulation can be accomplished. The articulators may close off the oral tract for an instant or a relatively long period; they may narrow the space considerably; or they may simply modify the shape of the tract by approaching each other.

STOP

(Complete closure of the articulators involved so that the airstream cannot escape through the mouth.) There are two possible types of stop.

Oral stop If, in addition to the articulatory closure in the mouth, the soft palate is raised so that the nasal tract is blocked off, then the airstream will be completely obstructed. Pressure in the mouth will build up and an **oral stop** will be formed. When the articulators come apart, the airstream will be released in a small burst of sound. This kind of sound occurs in the consonants in the words "pie, buy" (bilabial closure), "tie, dye" (alveolar closure), and "kye, guy" (velar closure). Figure 1.5 shows the position of the vocal organs in the bilabial stop in "buy."

The positions of the vocal organs in the bilabial stop in "buy."
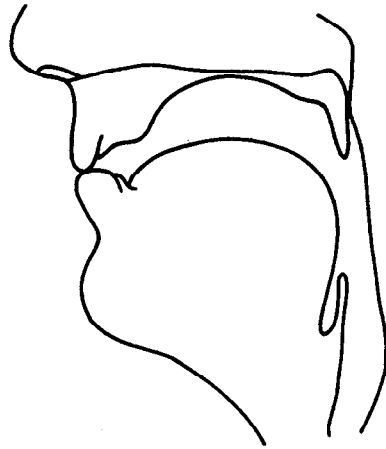
Nasal stop If the air is stopped in the oral cavity but the soft palate is down so that it can go out through the nose, the sound produced is a **nasal stop.** Sounds of this kind occur at the beginning of the words "my" (bilabial closure) and "nigh" (alveolar closure) and at the end of the word "sang" (velar closure). Figure 1.6 shows the position of the vocal organs during the bilabial nasal stop in

FIGURE 1.6 The position of the vocal organs in the bilabial nasal stop in "my."
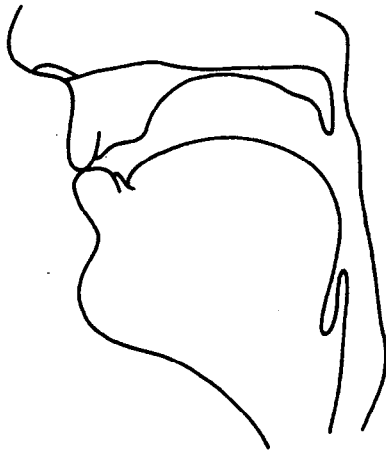
"my." Apart from the presence of a velic closure, there is no difference between this stop and the one in "buy"—shown in Figure 1.5. Although both the nasal sounds and the oral sounds can be classified as stops, the term **stop** by itself is almost always used by phoneticians to indicate an oral stop, and the term **nasal** is used to indicate a nasal stop. Thus the consonants at the beginnings of the words "day" and "neigh" would be called an alveolar stop and an alveolar nasal, respectively. Although the term *stop* may be defined so that it applies only to the prevention of air escaping through the mouth, it is commonly used to imply a complete stoppage of the airflow through both the nose and the mouth.

FRICATIVE

(Close approximation of two articulators so that the airstream is partially obstructed and turbulent airflow is produced.) The mechanism involved in making these slightly hissing sounds may be likened to that involved when the wind whistles around a corner. The consonants in "fie, vie" (labiodental), "thigh, thy" (dental), "sigh, zoo" (alveolar), and "shy" (palato-alveolar) are examples of fricative sounds. Figure 1.7 illustrates one pronunciation of the palato-alveolar fricative consonant in "shy." Note the narrowing of the vocal tract between the blade of the tongue and the back part of the alveolar ridge. The higher-pitched sounds with a more obvious hiss, such as those in "sigh, shy," are sometimes called **sibilants.**

FIGURE 1.7   The positions of the vocal organs in the palato-alveolar (post-alveolar) fricative in "shy."

APPROXIMANT

(An articulation in which one articulator is close to another, but without the vocal tract being narrowed to such an extent that a turbulent airstream is produced.) In saying the first sound in "yacht," the front of the tongue is raised toward the palatal area of the roof of the mouth, but it does not come close enough for a fricative sound to be produced. The consonants in the word "we" (approximation between the lips and in the velar region) and, for some people, in the word "raw" (approximation in the alveolar region) are also examples of approximants.

LATERAL (APPROXIMANT)

(Obstruction of the airstream at a point along the center of the oral tract, with incomplete closure between one or both sides of the tongue and the roof of the mouth.) Say the word "lie" and note how the tongue touches near the center of the alveolar ridge. Prolong the initial consonant and note how, despite the closure formed by the tongue, air flows out freely, over the side of the tongue. Because there is no stoppage of the air, and not even any fricative noises, these sounds are classified as approximants. The consonants in words such as "lie, laugh" are alveolar lateral approximants, but they are usually called just alveolar laterals, their approximant status being assumed. You may be able to find out which side of the tongue is not in contact with the roof of the mouth by holding the consonant position while you breathe inward. The tongue will feel colder on the side that is not in contact with the roof of the mouth.

ADDITIONAL CONSONANTAL ARTICULATIONS

In this preliminary chapter, it will not be necessary to discuss all of the manners of articulation used in the various languages of the world—nor, for that matter, in English. But it might be useful to know the terms **trill** (sometimes called roll) and **tap** (sometimes called flap). Tongue-tip trills occur in some forms of Scottish English in words such as "rye" and "raw." Taps, in which the tongue makes a single tap against the alveolar ridge, occur in the middle of a word such as "pity" in many forms of American English.

The production of some sounds involves more than one of these manners of articulation. Say the word "cheap" and think about how you make the first sound. At the beginning, the tongue comes up to make contact with the back part of the alveolar ridge to form a stop closure. This contact is then slackened so that there is a fricative at the same place of articulation. This kind of combination of a stop immediately followed by a fricative is called an **affricate,** in this case a palato-alveolar (or post-alveolar) affricate. There is a voiceless affricate at the beginning and end of the word "church." The corresponding voiced affricate occurs at the

| FIGURE 1.9 | The relative positions of the highest points of the tongue in the vowels in 1 heed, 2 hid, 3 head, 4 had, 5 father, 6 good, 7 food. |

```
                              front          back


       high               1●


                          2●        ●7
       mid
                            3●              ●6

       low                  4●
                               ●5
```
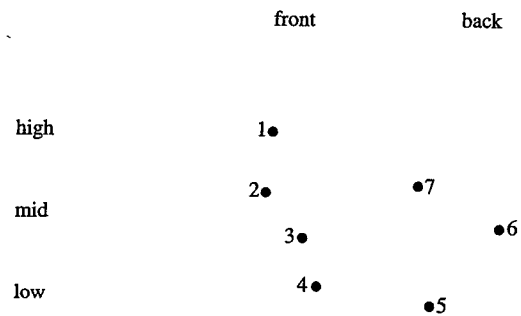
in Figure 1.9. Say just the vowels in the words given in the caption for this figure and check that your tongue moves in the pattern described by the points. It is very difficult to become aware of the position of the tongue in vowels, but you can probably get some impression of tongue height by observing the position of your jaw while saying just the vowels in the four words, "heed, hid, head, had." You should also be able to feel the difference between front and back vowels by contrasting words such as "he" and "who." Say these words silently and concentrate on the sensations involved. You should feel the tongue going from front to back as you say "he, who." You can also feel your lips becoming more rounded.

As you can see from Figure 1.9, the specification of vowels in terms of the position of the highest point of the tongue is not entirely satisfactory for a number of reasons. First, the vowels that are called high do not have the same tongue height. The back high vowel (point 7) is nowhere near as high as the front vowel (point 1). Second, the so-called back vowels vary considerably in their degree of backness. Third, as you can see by looking at Figure 1.8, this kind of specification disregards considerable differences in the shape of the tongue in front vowels and in back vowels. Furthermore, it does not take into account the fact that the width of the pharynx varies considerably with, and to some extent independently of, the height of the tongue in different vowels. We will discuss better ways of describing vowels in Chapters 4 and 9.

SUPRASEGMENTALS

Vowels and consonants can be thought of as the segments of which speech is composed. Together they form the syllables, which go to make up utterances.

Superimposed on the syllables are other features known as suprasegmentals. These include variations in stress and pitch. Variations in length are also usually considered to be suprasegmental features, although they can affect single segments as well as whole syllables.

Variations in stress are used in English to distinguish between a noun and a verb, as in "(an) insult" versus "(to) insult." Say these words yourself, and check which syllable has the greater stress. Then compare similar pairs, such as "(a) pervert, (to) pervert" or "(an) overflow, (to) overflow." You should find that in the nouns the stress is on the first syllable, but in the verbs it is on the last. Thus, stress can have a grammatical function in English. It can also be used for contrastive emphasis (as in "I want a *red* pen, not a black one"). Variations in stress are caused by an increase in the activity of the respiratory muscles (so that a greater amount of air is pushed out of the lungs) and in the activity of the laryngeal muscles (so that there is a significant change in pitch).

You can usually find where the stress occurs on a word by trying to tap with your finger in time with each syllable. It is much easier to tap on the stressed syllable. Try saying "abominable" and tapping first on the first syllable, then on the second, then on the third, and so on. If you say the word in your normal way you will find it easiest to tap on the second syllable. Many people cannot tap on the first syllable without altering their normal pronunciation.

Pitch changes due to variations in laryngeal activity can occur independently of stress changes. When they do, they can affect the meaning of the sentence as a whole. The pitch pattern in a sentence is known as the **intonation.** Listen to the intonation (the variations in the pitch of your voice) when you say the sentence "This is my father." Try to find out which syllable has the highest pitch and which the lowest. In most people's speech, the highest pitch will occur on the first syllable of "father" and the lowest on the second, the last syllable in the sentence. Now observe the pitch changes in the question "Is this your father?" In this sentence the first syllable of "father" is usually on a lower pitch than the last syllable. In English it is even possible to change the meaning of a sentence such as "That's a cat" from a statement to a question without altering the order of the words. If you substitute a mainly rising for a mainly falling intonation, you will produce a question spoken with an air of astonishment: "That's a *cat?*"

All the suprasegmental features are characterized by the fact that they must be described in relation to other items in the same utterance. It is the relative values of pitch, length, or degree of stress of an item that are significant. You can stress one syllable as opposed to another irrespective of whether you are shouting or talking softly. Children can also use the same intonation patterns as adults, although their voices have a higher pitch. The absolute values are never linguistically important. But they do, of course, convey information about the speaker's age, sex, emotional state, and attitude toward the topic under discussion.

EXERCISES
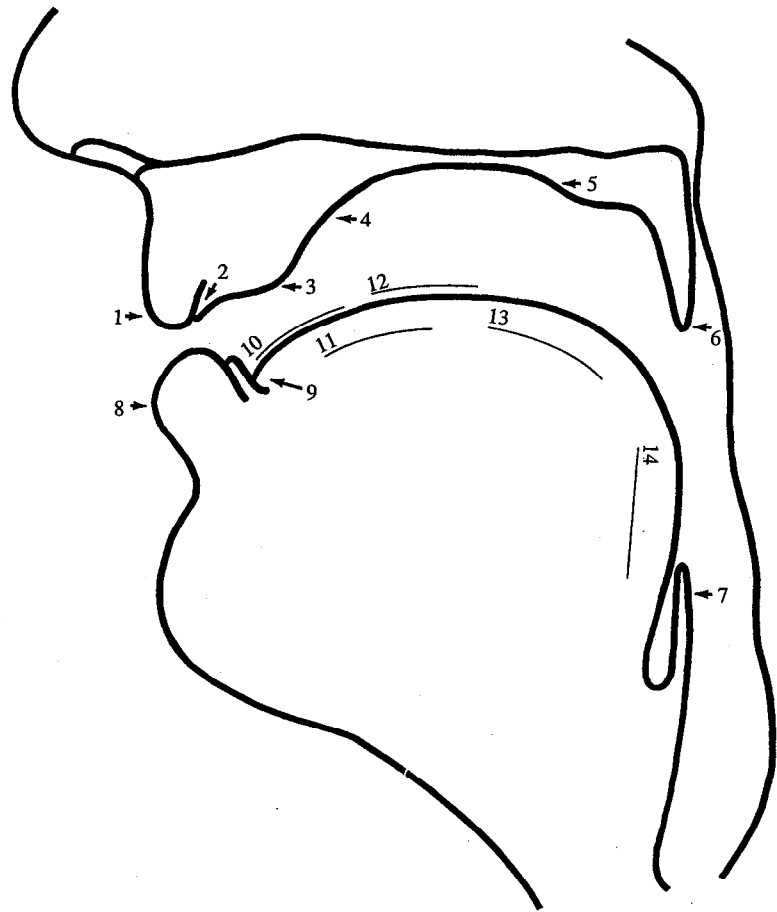
A Fill in the names of the vocal organs numbered in Figure 1.10 above.

1. _____   8. _____

2. _____   9. _____

3. _____   10. _____

4. _____   11. _____

5. _____   12. _____

6. _____   13. _____

7. _____   14. _____

B Describe the consonants in the word "skinflint" using the following chart. Fill in all five columns, and put parentheses around the terms that may be left out, as shown for the first consonant.

	1 voiced or voiceless	2 place of articulation	3 central or lateral	4 oral or nasal	5 articulatory action
s	*voiceless*	*alveolar*	*(central)*	*(oral)*	*fricative*
k					
n					
f					
l					
t					

C Figure 1.11a–g illustrates all the places of articulation we have discussed so far, except for retroflex sounds (which will be illustrated in Chapter 7). In the following spaces, state (1) the place of articulation and (2) the manner of articulation of each sound. In addition, give (3) an example of an English word beginning with the sound illustrated.

	(1) Place of articulation	(2) Manner of articulation	(3) Example
a	_____	_____	_____
b	_____	_____	_____
c	_____	_____	_____
d	_____	_____	_____
e	_____	_____	_____
f	_____	_____	_____
g	_____	_____	_____

FIGURE 1.11 Sounds illustrating all the places of articulation discussed so far, except for retroflex sounds.

(a) (b)

(c) (d)

(e) (f)

(g)
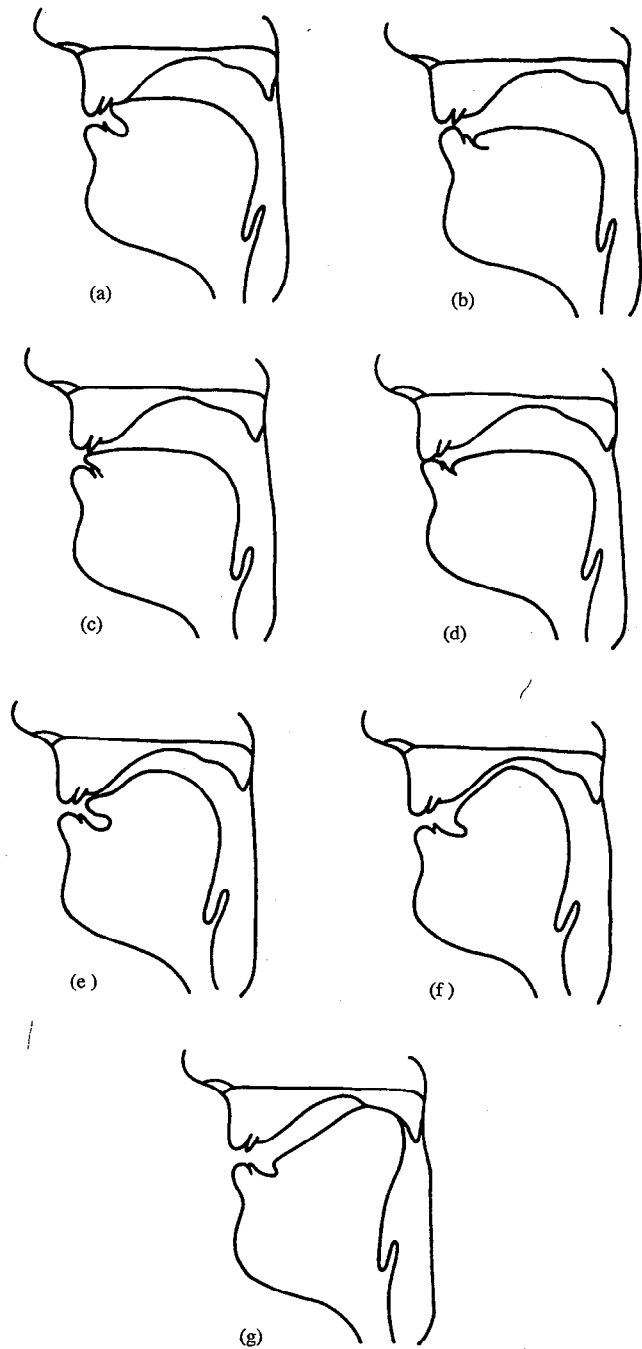
D Studying a new subject often involves learning a large number of technical terms. Phonetics is particularly difficult in this respect. Read over the definitions of the terms in this chapter, and then try the exercises that follow. Listen to the sounds of the words, and be careful not to be confused by spellings. Using a mirror may be helpful.

1. Circle the words that begin with a bilabial consonant:

    mat   gnat   sat   bat   rat   pat

2. Circle the words that begin with a velar consonant:

    knot   got   lot   cot   hot   pot

3. Circle the words that begin with a labiodental consonant:

    fat   cat   that   mat   chat   vat

4. Circle the words that begin with an alveolar consonant:

    zip   nip   lip   sip   tip   dip

5. Circle the words that begin with a dental consonant:

    pie   guy   shy   thigh   thy   high

6. Circle the words that begin with a palato-alveolar consonant:

    sigh   shy   tie   thigh   thy   lie

7. Circle the words that end with a fricative:

    race   wreath   bush   bring   breathe   bang

    rave   real   ray   rose   rough

8. Circle the words that end with a nasal:

    rain   rang   dumb   deaf

9. Circle the words that end with a stop:

    pill   lip   lit   graph   crab   dog   hide

    laugh   back

10. Circle the words that begin with a lateral:

    nut   lull   bar   rob   one

11. Circle the words that begin with an approximant:

    we   you   one   run

12. Circle the words that end with an affricate:

    much   back   edge   ooze

13. Circle the words in which the consonant in the middle is voiced:

    tracking   mother   robber   leisure   massive

    stomach   razor

14. Circle the words that contain a high vowel:

    sat   suit   got   meet   mud

15. Circle the words that contain a low vowel:

weed   wad   load   lad   rude

16. Circle the words that contain a front vowel:

gate   caught   cat   kit   put

17. Circle the words that contain a back vowel:

maid   weep   coop   cop   good

18. Circle the words that contain a rounded vowel:

who   me   us   but   him

E Define the consonant sounds in the middle of each of the following words as indicated in the example:

	Voiced or voiceless	Place of articulation	Manner of articulation
adder	*voiced*	*alveolar*	*stop*
father			
singing			
etching			
robber			
ether			
pleasure			
hopper			
selling			
sunny			
lodger			

F Complete the diagrams so as to illustrate the position of the vocal organs during the first consonants in each of the following words. If the sound is voiced, schematize the vibrating vocal folds by a wavy line at the glottis. If it is voiceless, use a straight line.

Example:
mat

day

cat

think

nut

2
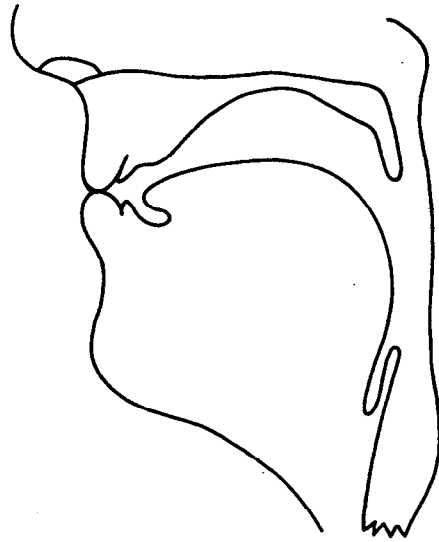
PHONOLOGY AND PHONETIC TRANSCRIPTION

Many people think that learning phonetics means simply learning to use phonetic transcription. But there is really much more to the subject than learning to use a set of symbols. A phonetician is a person who can describe speech, who understands the mechanisms of speech production and speech perception, and who knows how languages use these mechanisms. Phonetic transcription is no more than a useful tool that phoneticians use in the description of speech.

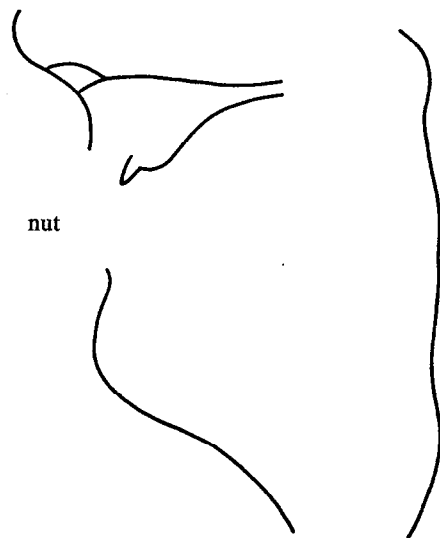When phoneticians transcribe an utterance, they usually do so by noting how the sounds convey differences in meaning. For the most part, they concern themselves with describing only the significant articulations rather than the total set of movements of the vocal organs. For example, when saying the English word "tie," some people pronounce the consonant with the blade of the tongue against the alveolar ridge, others with the tip of the tongue. This kind of difference in articulation does not affect the meaning of the word and is not usually transcribed.

In order to understand what we transcribe and what we don't, it is necessary to understand the basic principles of phonology. **Phonology** is the description of the systems and patterns of sounds that occur in a language. It involves studying a language to determine its distinctive sounds and to find out which sounds convey a difference in meaning. Children have to do this when they are learning to speak. To begin with they might not realize that, for example, there is a difference between the consonants at the beginnings of words such as "white" and "right." Later they realize that these words begin with two distinct sounds. Eventually they learn to distinguish all the sounds that can change the meanings of words.

When two sounds can be used to differentiate words, they are said to belong to different **phonemes.** There must be a phonemic difference if two words (such as "white" and "right" or "cat" and "bat") differ in only a single sound. There are, however, small shades of sounds that cannot be used to distinguish words, such as the differences between the consonants at the beginning and end of the word "pop." For the first of these sounds, the lips must open and there must be a puff of air before the vowel begins. After the final consonant there may be a puff of

air, but it is not necessary. In fact, you could say "pop" and never open your lips for hours, if it happened to be the last word you said before going to sleep. The sound at the end would still be a *p*. Both consonants in this word are voiceless bilabial stops. They are different, but the differences between them cannot be used to change the meaning of a word in English. They both belong to the same phoneme.

We cannot rely on the spelling to tell us whether two sounds are members of different phonemes. For example, if you think about the two words "key" and "car," you realize that they both begin with the same sound, despite the fact that one is spelled with the letter *k* and the other with the letter *c*. But this example also shows that there may be very subtle differences between members of a phoneme. The two words "key" and "car" do not begin with *exactly* the same sound. If you whisper just the first consonants in these two words, you can probably hear a difference; and you may be able to feel that your tongue touches the roof of your mouth in a different place. The two sounds at the beginning of "key" and "car" are slightly different, but it is not the kind of difference that changes the meaning of a word in English. They are both members of the same phoneme.

We noted other small changes in sounds that do not affect the meaning in Chapter 1, where we saw that the tongue is further back in "true" than in "tea," and the *n* in "tenth" is likely to be dental, whereas the one in "ten" is usually alveolar. In some cases the members of a phoneme are more different from one another. For example, most Americans (and many younger speakers of British English) have a *t* in the middle of "pity" that is very different from the *t* at the end of the word "pit." The one in "pity" sounds more like a *d*. Consider also the *l* in "play." You can say just the first two consonants in this word without any voicing, but you will still hear the *l* (try doing this). When you say the whole word "play" the *l* is typically voiceless, and very different from the one in "lay." Say the *l* at the beginning of "lay," and you'll hear that it is definitely voiced.

It follows from these examples that a phoneme is not a single sound but a name for a group of sounds. There is a group of *t* sounds and a group of *l* sounds that occur in English. It is as if you had in your mind an ideal *t* or *l*, and the ones that were actually produced were variations of it, which differed in small ways that did not affect the meaning of English words. These groups of sounds—the phonemes—are abstract units that form the basis for writing down a language systematically and unambiguously.

We often want to record all and only the variations between sounds that cause a difference in meaning. Transcriptions of this kind are called phonemic transcriptions. Languages that have been written down only comparatively recently (such as Swahili and most of the other languages of Africa) have a fairly phonemic spelling system. There is very little difference between a written version of a Swahili sentence and a phonemic transcription of that sentence. But because English pronunciation has changed over the centuries while the spelling has remained basically the same, phonemic transcriptions of English are different from written texts.

THE TRANSCRIPTION OF CONSONANTS

We can begin searching for phonemes by considering the contrasting consonant sounds in English. A good way is to find sets of words that rhyme. Take, for example, all the words that rhyme with "pie" and have only a single consonant at the beginning. A set of words, each of which differs from all the others by only one sound, is called a minimal set. The second column of Table 2.1 lists a minimal set of this kind. There are obviously many other words that rhyme with "pie," such as "spy, try, spry," but these words begin with sequences of two or more of the sounds already in the minimal set. Some of the words in the list begin with two consonant letters ("thigh, thy, shy"), but they each begin with a single consonant sound. "Shy," for example, does not contain a sequence of two consonant sounds in the way that "spy" and "try" do.

Some consonants do not occur in words rhyming with "pie." If we allow the names of the letters as words, then we can find a large set of consonants beginning words rhyming with "pea." A list of such words is shown in the third column of Table 2.1. (Speakers of British English will have to remember that in American English the name of the last letter of the alphabet belongs in this set rather than in the set of words rhyming with "bed.")

Even in this set of words, we are still missing some consonant sounds that contrast with others only in the middle or at the end of words. The letters *ng* often represent a single consonant sound that does not occur at the beginning of a word. You can hear this sound at the end of the word "rang," where it contrasts with other nasals in words such as "ram, ran." There is also a contrast between the consonants in the middle of "mission" and "vision," although there are very few pairs of words that are distinguished by this contrast in English. (One such pair in my English involves the name of a chain of islands—"Aleutian" versus "allusion.") Words illustrating these consonants are given in the fourth column of Table 2.1.

Most of the symbols in Table 2.1 are the same letters we use in spelling these words, but there are a few differences. One variation between spelling and phonetic usage occurs with the letter *c,* which is sometimes used to represent a [k] sound, as in "cup" or "bacon," and sometimes to represent an [s] sound, as in "cellar" or "receive." Two *c*'s may even represent a sequence of these sounds in the same word, as in "accent, access." A symbol that differs from the corresponding letter is [g], which is used for the sound "guy" and "guess" but never for the sound in "age" or the sound in the name of the letter *g.* A few other symbols are needed to supplement the regular alphabet. The phonetic symbols we will use are part of the set approved by the International Phonetic Association, a body founded in 1886 by a group of leading phoneticians from France, Germany, Britain, and Denmark. The complete set of IPA symbols is given in the chart in the inside covers of this book. It will be discussed in detail later in this book. Because we often need to talk about the symbols, the names that have been given to them are shown in the last column of Table 2.1.

| TABLE 2.1 | Symbols for transcribing English consonants. (Alternative symbols that may be found in other books are given in parentheses.) The last column gives the conventional names for the phonetic symbols in the first column. |

p	pie	pea		lowercase *p*
t	tie	tea		lowercase *t*
k	kye	key		lowercase *k*
b	by	bee		lowercase *b*
d	dye	D		lowercase *d*
g	guy			lowercase *g*
m	my	me	ra*m*	lowercase *m*
n	nigh	knee	ra*n*	lowercase *n*
ŋ			ra*ng*	eng (or angma)
f	fie	fee		lowercase *f*
v	vie	V		lowercase *v*
θ	thigh			theta
ð	thy	thee		eth
s	sigh	sea	li*s*ten	lowercase *s*
z		Z	mi*zz*en	lowercase *z*
ʃ (š)	shy	she	mi*ss*ion	esh (or long *s*)
ʒ (ž)			vi*s*ion	long *z* (or yogh)
l	lie	lee		lowercase *l*
w	why	we		lowercase *w*
r (ɹ)	rye			lowercase *r*
j (y)		ye		lowercase *j*
h	high	he		lowercase *h*

Note also the following:

tʃ (tš)	chi(me)	chea(p)
dʒ (dž)	ji(ve)	G

The velar nasal at the end of "rang" is written with [ŋ], a letter *n* combined with the tail of the letter *g* descending below the line. Some people call this symbol "eng" and others "angma." The symbol [θ], an upright version of the Greek letter theta, is used for the voiceless dental fricative in words such as "thigh, thin, thimble, ether, breath, mouth." The symbol [ð], called "eth," is derived from an Anglo-Saxon letter. It is used for the corresponding voiced sound in words such as "thy, then, them, either, breathe." Both these symbols are ascenders (letters that go up from the line of writing rather than descending below it). The spelling system of the English language does not distinguish between [θ] and [ð]. They are both written with the letters *th* in pairs such as "thigh, thy."

The voiceless palato-alveolar (post-alveolar) fricative [ʃ] ("long *s*") in "shy, sheep, rash" is both an ascender and a descender. It is like a long, straightened *s* going both above and below the line of writing. The corresponding voiced

symbol [ʒ] is like a long *z* descending below the line. This sound occurs in the middle of words such as "vision, measure, leisure" and at the beginning of foreign words such as the French "Jean, gendarme," and foreign names such as "Zsa Zsa."

In earlier editions of this book, the sound at the beginning of the word "rye" was symbolized by [ɹ], an upside-down letter *r*. As the two major dictionaries of American and British English pronunciation (see Further Reading) use a regular [r] for this sound, I have done so here.

It is unfortunate that different books on phonetics use different forms of phonetic transcription. This is not because phoneticians cannot agree on which symbols to use, but because different styles of transcription are more appropriate in one circumstance than in another. Thus in this book, where we are concerned with general phonetics, I have used the IPA symbol [j] for the initial sound in "yes, yet, yeast" because I wish to reserve the symbol [y] for another sound, the vowel in the French word "tu." Another reason for using [j] is that in many languages (German, Dutch, Norwegian, Swedish, and others) this letter is used in words such as "ja," which are pronounced with a sound that in our spelling system would be written with the letter *y*. Books that are concerned only with the phonetics of English often use [y] where this one uses [j]. Some books on phonetics also use [š] and [ž] in place of the IPA symbols [ʃ] and [ʒ], respectively.

There are also disagreements among texts on phonetics on how to transcribe sounds such as the first and last sounds in both "church" and "judge." I have taken the position that these sounds are each sequences of two other consonants and should be written [tʃ] and [dʒ]. They are like other words that begin or end with a sequence of two consonants that occur in other words such as "tray" and "adds."

You can see that a word such as "choose" might be said to begin with [tʃ] if you compare your pronunciation of the phrases "white shoes" and "why choose." In the first phrase, the [t] is at the end of one word and the [ʃ] at the beginning of the next; but in the second phrase, these two sounds occur together at the beginning of the second word. The difference between the two phrases is simply one of the timing of the articulations involved rather than the use of different articulations. Other pairs of phrases that demonstrate this point are "heat sheets" versus "he cheats" and "might shop" versus "my chop." There are no pairs of phrases illustrating the same point for the voiced counterpart [dʒ] found in "jar, gentle, age" because no English word begins with [ʒ].

Some other books on phonetics take the view that the sounds [tʃ] and [dʒ] (as in "church" and "judge") are really single units and are better transcribed with a single symbol, such as [č] and [ǰ]. This view has much to commend it, since the consonants [ʃ] and [ʒ] are not like other consonants such as [r] and [l]. Each of the latter pair of consonants can occur as the second element in many clusters (for example, in "priest, tree, cream, play, clay"). But [ʃ] and [ʒ]

cluster only with [t] and [d], respectively. However, as this is a book on phonetics, it seems appropriate to use two symbols for the consonants in words such as "jay" and "age" to show that there are two elements in each of them, just as there are in other words containing consonant clusters, such as "tree" and "eats."

There is one minor matter still to be considered in the transcription of the consonant contrasts of English. In most forms of both British and American English, "which" does not contrast with "witch." Accordingly, both "why" and "we" in Table 2.1 are said to begin simply with [w]. But some speakers of English contrast pairs of words such as "which, witch; why, Wye; whether, weather." These speakers will have to transcribe the first of each of these pairs of words with an initial [h]. Note that phonetically the [h] is transcribed before [w] in that it is the first part of each of these words that is voiceless.

THE TRANSCRIPTION OF VOWELS

The transcription of the contrasting vowels (the vowel phonemes) in English is more difficult than the transcription of consonants for two reasons. First, accents of English differ more in their use of vowels than in their use of consonants. Second, authorities differ widely in their views of what constitutes an appropriate description of vowels.

Taking the same approach in looking for contrasting vowels as we did in looking for consonant contrasts, we might try to find a minimal set of words that differ only in the vowel sounds. We could, for example, look for monosyllables that begin with [h] and end with [d], and we could supplement this minimal set with other lists of monosyllables that contrast only in their vowel sounds. Table 2.2 shows four such sets of words.

We will consider one form of British and one form of American English. The major difference between the two is that speakers of American English pronounce [r] sounds after vowels, as well as before them, whereas in most forms of British English [r] can occur only before a vowel. American English speakers distinguish between words such as "heart" and "hot" not by making a difference in vowel quality (as I do, speaking a form of British English) but by pronouncing "heart" with an [r] and "hot" with the same vowel but without an [r] following it. In "here, hair, hire," these speakers may use vowels similar to those in "he, head, high," respectively, but in each case with a following [r]. Most speakers of British English distinguish these words by using different **diphthongs**—movements from one vowel to another within a single syllable.

Even within American English there are variations in the number of contrasting vowels that occur. Many Midwestern speakers and most Far Western speakers do not distinguish between the vowels in pairs of words such as "odd, awed" and "cot, caught." Some forms of American English make additional distinctions

| TABLE 2.2 | Symbols for transcribing contrasting vowels in English. Column 1 applies to many speakers of American English, Column 2 to most speakers of British English. The last column gives the conventional names for the phonetic symbols in the second column, some of which are also accompanied by the length mark [ː] or by [r]. |

1	2						
iː	iː	heed	he	bead	heat	keyed	lowercase *i*
ɪ	ɪ	hid		bid	hit	kid	small capital *I*
eɪ	eɪ	hayed	hay	bayed	hate	Cade	lowercase *e*
ɛ	ɛ	head		bed			epsilon
æ	æ	had		bad	hat	cad	ash
ɑːr	ɑː	hard		bard	heart	card	script *a*
ɑː	ɒ	hod		bod	hot	cod	(2) turned script *a*
ɔː	ɔː	hawed	haw	bawd		cawed	open *o*
ʊ	ʊ	hood				could	upsilon
oʊ	əʊ	hoed	hoe	bode		code	lowercase *o*
uː	uː	who'd	who	booed	hoot	cooed	lowercase *u*
ʌ	ʌ	Hudd		bud	hut	cud	turned *v*
ɝ	ɜː	herd	her	bird	hurt	curd	reversed epsilon
aɪ	aɪ	hide	high	bide	height		lowercase *a* (+ ɪ)
aʊ	aʊ		how	bowed		cowed	(as noted above)
ɔɪ	ɔɪ		(a)hoy	Boyd			(as noted above)
ɪr	ɪə		here	beard			(as noted above)
ɛr	ɛə		hair	bared		cared	(as noted above)
aɪr	aə	hired	hire				(as noted above)
Note also the following:							
ju	ju	hued	hue	Bude		cued	(as noted above)

not shown in Table 2.2. For example, some speakers (mainly from the East Coast) distinguish the auxiliary verb "can" from the noun "can," the latter being more diphthongal. But we will have to overlook these small differences in this introductory textbook.

There are several possible ways of transcribing the contrasting vowels in Table 2.2. The two principal forms that will be used in this book are shown in the first and second columns. The first column is suitable for many forms of American English and the second for many forms of British English. The two columns have been kept as similar as possible; as you will see in Chapter 4, I have tried to make the transcriptions compatible with those of well-known authorities on the phonetics of English.

As in the case of the consonant symbols, the vowel symbols in Table 2.2 are used in accordance with the principles of the IPA. Those symbols that have the same shape as ordinary letters of the alphabet represent sounds similar to the sounds these letters have in French or Spanish or Italian. Actually, the IPA usage

of the vowel letters is that of the great majority of the world's languages when they are written with the Roman alphabet, including such diverse languages as Swahili, Turkish, and Navajo. The present spelling of English reflects the way it used to sound many centuries ago when it still had vowel letters with values similar to those of the corresponding letters in all these other languages.

One of the principal problems in transcribing English phonetically is that there are more vowel sounds than there are vowel letters in the alphabet. In a transcription of the English word "sea" as [siː], the [iː] represents a similar (but not identical) sound to that in the Spanish or Italian "si." But unlike Spanish and Italian, English differentiates between vowels such as those in "seat" and "sit" and "heed" and "hid." The vowels in "seat, heed" differ from those in "sit, hid" in two ways: they have a slightly different quality and they are longer. Because the vowels in "sit, hid" are somewhat like those in "seat, heed," they are represented by the symbol [ɪ], a small capital *I*. The difference in length is shown by the symbol [ː].

The vowels in words such as "hay, bait, they" are transcribed with a sequence of two symbols [eɪ], indicating that for most speakers of English these words contain a diphthong. The first element in this diphthong is similar to sounds in Spanish or Italian that use the letter *e*, such as the Spanish word for milk, which is written "leche" and pronounced [letʃe]. The second element in the English words "hay, bait, they" is [ɪ], the symbol used for transcribing the vowel in "hid."

Two symbols that are not ordinary letters of the alphabet, [ɛ] and [æ], are used for the vowels in "head" and "had," respectively. The first is based on the Greek letter epsilon, and the second on the letters *a* and *e* joined together. They may be referred to by the names epsilon and digraph.

Most Americans have the same vowel sound in the words "heart" and "hot" and can use one form of the letter *a*. They would transcribe these words as [hɑːrt] and [hɑːt], with the length mark [ː] indicating that it is a comparatively long vowel. But some East Coast Americans and speakers of British English who do not pronounce [r] sounds after a vowel distinguish between these words by the qualities of the vowels and have to use two different forms of the letter *a*. They would transcribe these words as [hɑːt] and [hɒt].

Most speakers of British forms of English, and many American speakers, distinguish between pairs of words such as "cot, caught; not, naught." The symbols [ɔː], an open letter *o* followed by a length mark, may be used in the second of each of these pairs of words and in words such as "bawd, bought, law." Many Midwestern and Far Western American speakers do not need to use this symbol in any of these words, as they do not distinguish between the vowels in words such as "cot" and "caught." They may, however, have different vowels in words such as "moss" and "morse," so that they, like many other Americans, would use the symbol [ɔː] when there is a following [r] sound.

Another special symbol is used for the vowel in "hood, could, good." This symbol [ʊ] may be thought of as a letter *u* with the ends curled out.

The vowel in "hoe, dough, code" is a diphthong. For most American English speakers, the first element is very similar to sounds that are written in Spanish, or Italian with a letter *o*. Many speakers of English from the southern parts of Britain use a different sound for the first element of the diphthong in these words (which we will symbolize with [ə], an upside down letter *e*). We will discuss this sound more fully in a later section. The final element of the diphthong in words such as "hoe, code" is somewhat similar to the vowel [ʊ] in "hood."

An upside down letter *v* [ʌ] is used for the vowel in words such as "bud, hut." This symbol is sometimes referred to by the name *wedge*.

Another symbol, [ɜː], a reversed form of the Greek letter epsilon followed by a length mark, is used for the sound in "pert, bird, curt," as pronounced by most speakers of British English and those speakers of American English who do not have an [r] in these words. In most forms of American English, the *r* is fully combined with the vowel, and the symbol [ɝ] is used. The little hook [˞] indicates the *r*-coloring of the vowel.

The next three words in Table 2.2 contain diphthongs composed of elements that have been discussed already. The vowel in "hide" [haɪd] begins with a sound between that of the vowel in "cat" [kæt] and that in "far" [fɑː(r)], and moves toward the vowel [ɪ] as in "hid" [hɪd]. The symbol [a] is used for the first part of this diphthong. The vowel in "how" [aʊ] begins with a similar sound but moves toward [ʊ] as in "hood." The vowel in "boy" [bɔɪ] is a combination of the sound [ɔ] as in "bawd" and [ɪ] as in "hid."

Most Americans pronounce the remaining words in Table 2.2 with one of the other vowels followed by [r], while most British English speakers have additional diphthongs in these words. In each case, the end of the diphthong is [ə], the same symbol we used for the beginning of the diphthong in "hoe" for most British English speakers. We will discuss this symbol further in the next paragraph. Some British English speakers also use a diphthong in words like "poor, cure" that can be transcribed as [ʊə]. Some people (myself included) have a diphthong [aə] in words such as "fire, hire." Others pronounce these words as two syllables (like "higher" and "liar"), transcribing them as [faɪə, haɪə].

The words in Table 2.2 are generally monosyllables. Consequently, none of them contains both stressed and unstressed vowels. By far the most common unstressed vowel is [ə], the one we noted at the end of the diphthongs in most forms of British English. It is often called by its German name, *schwa*. It occurs at the ends of words such as "sofa, soda" [ˈsoʊfə, ˈsoʊdə], in the middle of words such as "emphasis, deprecate" [ˈɛmfəsɪs, ˈdɛprəkeɪt], and at the beginnings of words such as "around, arise" [əˈraʊnd, əˈraɪz]. (In all these words, the symbol ['] is a stress mark that has been placed before the syllable carrying the main stress. Stress should always be marked in words of more than one syllable.)

In British English, [ə] is usually the sole component of the "-er" part of words such as "father, brotherhood, simpler" [ˈfɑːðə, ˈbrʌðəhʊd, ˈsɪmplə]. In forms of American English with *r*-colored vowels, these words are usually [ˈfɑːðɚ, ˈbrʌðɚhʊd, ˈsɪmplɚ]. As with the symbol [ɝ], the small hook on [ɚ] symbolizes the *r*-coloring. Both [ə] and [ɚ] are very common vowels, [ə] occurring very frequently in unstressed monosyllables such as the grammatical function words "the, a, to, and, but." In connected speech these words are usually [ðə, ə, tə, ənd, bət].

Some of the other vowels also occur in unstressed syllables, but because of differences in accents of English, it is a little more difficult to say which vowel occurs in which word. For example, nearly all speakers of English differentiate between the last vowels in "Sophie" and "sofa" or "pity" and "patter." But some accents have the vowel [i] as in "heed" at the end of "Sophie, pity, only, corny." Others have [ɪ] as in "hid." Similarly, most accents make the vowel in the second syllable of "taxis" different from that in "Texas." Some have [i] and some have [ɪ] in "taxis." Nearly everybody pronounces "Texas" as [ˈtɛksəs]. (Note that in English the letter *x* often represents the sounds [ks].)

TRANSCRIPTION EXERCISES

Find the errors in the transcription of the consonant sounds in the following words. In each word there is one error, indicating an impossible pronunciation of that word for a native speaker of English of any variety. Circle this error, and write the correct symbol in the space provided after the word.

1. "strength"	[ˈstrɛŋθ]	should be	[ŋ	]
2. "crime"	[ˈcraɪm]		[k	]
3. "wishing"	[ˈwɪshɪŋ]		[ʃ	]
4. "wives"	[ˈwaɪvs]		[z	]
5. "these"	[ˈθiz]		[ð	]
6. "hijacking"	[ˈhaɪjækɪŋ]		[dʒ	]
7. "chipping"	[ˈtʃɪppɪŋ]		[p	]
8. "yelling"	[ˈyɛlɪŋ]		[j	]
9. "sixteen"	[ˈsɪxtiːn]		[ks	]
10. "thesis"	[ˈðisɪs]		[θiː	]

Now try another ten words in which the errors are all in the vowels. Again, there is only one possible error; but because of differences in varieties of English, there are sometimes alternative possible corrections.

11. "man-made" [ˈmanmeɪd]   should be   [   ]

12. "football" [ˈfʊtbol]   [   ]

13. "tea chest" [ˈtiːtʃest]   [   ]

14. "tomcat" [ˈtomkæt]   [   ]

15. "tiptoe" [ˈtiptoʊ]   [   ]

16. "avoid" [æˈvɔɪd]   [   ]

17. "remain" [rəˈman]   [   ]

18. "roommate" [ˈrɔmmeɪt]   [   ]

19. "umbrella" [umˈbrɛlə]   [   ]

20. "manage" [ˈmænædʒ]   [   ]

Now correct the following words, again by circling the error and noting the correction. There is still only one per word, but it may be among the vowels, the consonants, or the stress marks.

21. "magnify" [ˈmægnifaɪ]   should be   [   ]

22. "traffic" [ˈtræfɪc]   [   ]

23. "simplistic" [ˈsɪmplɪstɪk]   [   ]

24. "irrigate" [ˈɪrɪgeɪt]   [   ]

25. "improvement" [ɪmˈprʊvmənt]   [   ]

26. "demonstrate" [ˈdemɑnstreɪt]   [   ]

27. "human being" [huːmən ˈbiːŋ]   [   ]

28. "appreciate" [əˈpreʃiːeɪt]   [   ]

29. "joyful" [ˈdʒɔyfʊl]   [   ]

Finally, transcribe the following words or phrases as you pronounce them, or as you think they should be pronounced in one of the well-known accents of English. Give a rough label for your accent or the one you have chosen to use (e.g., BBC English, Southern Californian, New York, London, Scottish), and try to make the transcription illustrate an ordinary conversational style. Do not use a reading pronunciation or one that you think ought to be used. Say the words aloud (or listen to someone else saying them) and notice how they are normally pronounced. Be careful to put in stress marks at the proper places.

accent _____

30. chocolate pudding

31. modern languages

32. impossibility

33. boisterous

34. youngster

35. another

36. diabolical

37. nearly over

38. red riding hood

39. inexcusable

CONSONANT AND VOWEL CHARTS

So far we have been using the consonant and vowel symbols mainly as ways of representing the contrasts that occur among words in English. But they can also be thought of in a completely different way. We may regard them as shorthand descriptions of the articulations involved. Thus [p] is an abbreviation for "voiceless bilabial stop" and [l] is equivalent to "voiced alveolar lateral approximant." The consonant symbols can then be arranged in the form of a chart as in Figure 2.1. The places of articulation are shown across the top of the chart, starting from the most forward articulation (bilabial) and going toward those sounds made in the back of the mouth (velar). The manners of articulation are shown on the vertical axis of the chart. By convention, the voiced–voiceless distinction is shown by putting the voiceless symbols to the left of the voiced symbols.

The symbol [w] is shown in two places in the consonant chart in Figure 2.1. This is because it is articulated with both a narrowing of the lip aperture, which makes it bilabial, and a raising of the back of the tongue toward the soft palate, which makes it velar. The symbol [h] does not appear anywhere on the chart. In English, [h] acts like a consonant, but from an articulatory point of view it is the voiceless counterpart of the surrounding sounds. It does not have a precise place of articulation, and its manner of articulation is similar to that of the vowels before and after it.

The symbols we have been using for the contrasting vowels may also be regarded as shorthand descriptions for different vowel qualities. There are problems in this respect, in that we have been using these symbols somewhat loosely, allowing them to have different values for different accents. But the general values can be indicated by a vowel chart as in Figure 2.2. The symbols have been placed within a quadrilateral, which shows the range of possible vowel qualities but does not show differences in length. Thus [i] is used for a high front vowel, [u] for a high back one, [ɪ] for a lower high front vowel, [e] for a raised mid-front vowel, [ɛ] for a lowered mid-front vowel, and so on.

| FIGURE 2.1 | A phonetic chart of the English consonants we have dealt with so far. Whenever there are two symbols within a single cell, the one on the left represents a voiceless sound. All other symbols represent voiced sounds. Note also the consonant [h], which is not on this chart, and the affricates [tʃ, dʒ], which are sequences of symbols on the chart. |

Place of articulation

Manner of articulation	bilabial	labio-dental	dental	alveolar	palato-alveolar	palatal	velar
nasal (stop)	m			n			ŋ
stop	p b			t d			k g
fricative		f v	θ ð	s z	ʃ ʒ		
(central) approximant	(w)			r		j	w
lateral (approximant)				l			

The simple vowel chart in Figure 2.2 shows only two of the dimensions of vowel quality, and (as we will see in later chapters) even these are not represented very accurately. Furthermore, Figure 2.2 does not show anything about the variations in the degree of lip rounding in the different vowels, nor does it indicate anything about vowel length. It does not show, for example, that in most circumstances [i] and [u] are longer than [ɪ] and [ʊ].

PHONOLOGY

At the beginning of this chapter we discussed another reason why it is only approximately true that in our transcriptions of English the symbols have the values shown in Figures 2.1 and 2.2. In the style of transcription we have been using so far, some of the symbols may represent different sounds when they occur in different contexts. For example, the symbol [t] may represent a wide variety of sounds. In "tap" [tæp] it represents a voiceless alveolar stop. But the

FIGURE 2.2	A vowel chart showing the relative vowel qualities represented by some of the symbols used in transcribing English. The symbols [e, a, o] occur as the first elements of diphthongs.
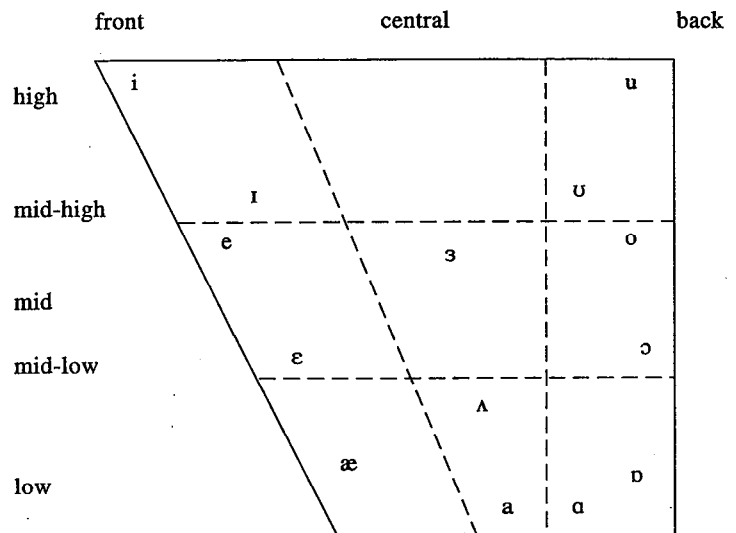
[t] in "eighth" [eɪtθ] may be made on the teeth, because of the influence of the following voiceless dental fricative [θ]. This [t] is more accurately called a voiceless dental stop, and we will later use a special symbol for transcribing it. In most forms of both British and American English, the [t] in "bitten" is accompanied by a glottal stop, and we will later be using a special symbol for this sound. As we saw, for most Americans and for many younger British English speakers, the [t] in "catty" [ˈkæti] symbolizes a voiced, not a voiceless, sound. All these different sounds are part of the / t / phoneme. Each of them occurs in a specific place—before [θ] or whatever may be the case—and none of them is different enough to change the meaning of a word in English. When saying them, what the speaker is trying to produce and what the listener hears is some form of / t /.

Similarly, other symbols represent different sounds in different contexts. The symbols [l] and [r] normally stand for voiced approximants. But in words such as "ply" [plaɪ] and "try" [traɪ] they represent voiceless sounds. Vowel sounds also vary. The [iː] in "heed" [hiːd] is usually very different from the [iː] in "heel" [hiːl].

Many of the variations we have been discussing can be described in terms of simple statements about regular sound patterns. Statements of this kind may be considered rules. In most forms of American English, for example, it is a rule

that [t] becomes voiced not only in "catty," but on all occasions when it occurs between two vowels, the second of which is unstressed (for example, in "pity, matter, utter, divinity," etc.). In English of nearly all kinds, it is also a rule that whenever [t] occurs before a dental fricative, it is pronounced as a dental stop, so that we may use the special symbol [t̪]. The same is true of [d], as in "width" [wɪd̪ð], [n], as in "tenth" [ten̪θ], and [l], as in "wealth" [wel̪θ]. In all these cases, the mark [̪] may be added under the symbol in order to indicate that it represents a dental articulation.

When we describe the sound patterns that occur in English, we want to be able to say that in some sense there are always the same underlying sounds. But these underlying sounds may change depending on the context in which they occur. The phonology of a language is the set of rules that describe the changes in the underlying sounds, the abstract units called phonemes, defined at the beginning of this chapter. When we transcribe a word in a way that shows none of the details of the pronunciation that are predictable by phonological rules, we are making a phonemic transcription. Phonemic segments are usually placed between slanting lines. Thus we may say that the underlying phonemic segments in "ten" and "tenth" are / ten / and / tenθ /. But the phonetic segments that are actually pronounced are [ten] and [ten̪θ].

The variants of the phonemes that occur in detailed phonetic transcriptions are known as **allophones.** They are generated as a result of applying the phonological rules to the underlying phonemes. We have already discussed some of the rules that generate different allophones of the phoneme / t /. For example, we know that in most varieties of American English, / t / has a voiced allophone when it occurs between a stressed vowel and an unstressed vowel.

The transcription used in the first part of this chapter is not, strictly speaking, a simple phonemic transcription. The symbols in Tables 2.1 and 2.2 distinguish all the oppositions that occur in actual pronunciations of English words. But these are not the underlying segments. There are several ways in which we could change our notational system into a more abstract one that uses a simpler set of symbols.

One way would be not to show the differences in length between pairs of vowels. Our present transcription indicates the differences in both quality and length in the vowels in a pair or words such as "bead, bid" [biːd, bɪd] or "fool, full" [fuːl, fʊl]. But the differences in length are entirely predictable from the differences in quality. The symbols [i, u] can be said to represent long vowels, and the symbols [ɪ, ʊ] short vowels (perhaps with an additional stipulation that we have been tacitly assuming: [i] in unstressed position as in "happy" is not long). In this way we would eliminate the length mark [ː].

Alternatively, instead of not showing the length difference, we could regard the quality as predictable from the length and not show the quality difference. In this case, the vowel [i] would have the value appropriate for the vowel in "bead" when it occurred before a length mark as in [biːd]. But it would have the

value appropriate for the vowel in "bid" when it occurred without a length mark, which would then be transcribed as [bid]. Remember that there is nothing sacred about the phonetic value of a symbol. Some phoneticians transcribe "bead, bid" as [biːd, bɪd] as I have been doing in this book so far, while others transcribe these same words as [biːd, bid]. Using the same principle, they might transcribe "cooed, could" not as [kuːd, kʊd], as I have done, but as [kuːd, kud]. Finally, "laid, led" would not be [leɪd, lɛd], but [leːd, led]. In this style of transcription, the differences in quality are treated as if they depended on the differences in length. The symbols / i / and / u / would be said to represent higher vowels when long, and the symbol / e / would be said to represent a diphthong in these circumstances. This style of transcription uses an additional symbol for length, but it more than compensates for that by eliminating the vowel symbols [ɪ, ɛ, ʊ].

We can now see why there are various possible ways of transcribing a language. Even if we consider only one particular accent of the language, it might be appropriate in some cases to symbolize one aspect of a contrast, such as the length, and in other cases to symbolize another, such as the quality. In addition, we may choose to make a transcription that shows only the underlying phonemes, or we may choose to represent some allophonic differences. Even if we are representing only the underlying phonemes in a particular accent of English, we may do so using only simple symbols, such as [r], or more unusual symbols, such as [ɹ], which convey more specific information on the phonetic quality. The term **broad transcription** is often used to designate a transcription that uses a simple set of symbols. Conversely, a **narrow transcription** is one that shows more phonetic detail, either just by using more specific symbols or by also representing some allophonic differences. The use of **diacritics,** small marks that can be added to a symbol to modify its value, is a means of increasing precision. One such diacritic is a small circle, [̥], that can be placed under a symbol to make it represent a voiceless sound, so that "ply" and "try," for instance, can be written [pl̥aɪ] and [tr̥aɪ]. Another useful diacritic is the mark [̪] beneath a consonant, which we have been using to indicate that the sound is dental and not alveolar.

Every transcription should be considered as having two aspects, one of which is often not explicit. There is the text itself, and at least implicitly there is a set of conventions for interpreting the text. These conventions are usually of two kinds. First, there are the conventions that ascribe general phonetic values to the symbols. It was with these conventions in mind that I said earlier that a symbol could be regarded as an approximate specification of the articulations involved. Second, there are the rules that specify the allophones that occur in different circumstances. Thus when I transcribe the word "peels" as [piːlz], I am assuming that the reader knows a number of the rules of English, including those that make / iː / somewhat lower and more central when it occurs before / l / and a final / z / voiceless toward the end.

On a few occasions, a transcription cannot be said to imply the existence of rules accounting for allophones. This is at least theoretically possible in the case of a narrow transcription so detailed that it shows *all* the rule-governed alternations among the sounds. A transcription that shows the allophones in this way is called a completely **systematic phonetic transcription.** In practice, it is difficult to make a transcription so narrow that it shows every detail of the sounds involved. On some occasions, a transcription may not imply the existence of rules accounting for allophones because, in the circumstances when the transcription was made, nothing was known about the rules. When writing down an unknown language or when transcribing a child or a patient not seen previously, one does not know what rules will apply. In these circumstances, the symbols indicate only the phonetic value of the sounds. This kind of transcription is called an **impressionistic transcription.**

EXERCISES

A Which of the two transcriptions below is the narrower?
"Betty cried as she left in the plane."

(a) [ˈbeti ˈkraɪd əz ʃiː ˈleft ɪn ðə ˈpleɪn]

(b) [ˈbɛdi ˈkr̥aɪd əz ʃiː ˈlɛft ɪn̥ ðə ˈpl̥eɪn]

B State rules for converting the transcription in (a) into that in (b). Make your rules as general as possible, so that they cover not only this pair of transcriptions but also other similar sentences (for example, [t] → [d] when it occurs after a stressed vowel and before an unstressed vowel).

C Read the following passages in phonetic transcription. Both passages use a fairly broad style of transcription, showing few allophones. The first represents a form of British English of the kind I speak myself. The second represents an American pronunciation typical of a Midwestern speaker. By this time you should be able to read transcriptions of different forms of English, although you may have difficulty pronouncing each word exactly as it is represented. Nevertheless, read each passage several times and try to pronounce it as indicated. Take care to put the stresses on the correct syllables, and say the unstressed syllables with the vowels as shown. Note any differences between each transcription and your own pronunciation of the corresponding words.

British English

[ɪt ɪz ˈpɒsəbl tə trænˈskraɪb fəˈnɛtɪklɪ ˈɛnɪ ˈʌtrəns, ɪn ˈɛnɪ ˈlæŋgwɪdʒ, ɪn ˈsɛvrəl ˈdɪfrənt ˈweɪz ˈɔl əv ðəm ˈjuːzɪŋ ði ˈælfəbɛt ənd kənˈvɛnʃnz əv ði ˈaɪ piː ˈeɪ. ðə ˈseɪm ˈθɪŋ ɪz ˈpɒsəbl wɪð ˈməʊst ˈʌðə ɪntəˈnæʃənl fəˈnɛtɪk ˈælfəbɛts. ə trænˈskrɪpʃn wɪtʃ ɪz ˈmeɪd baɪ ˈjuːzɪŋ ˈlɛtəz əv ðə ˈsɪmpləst ˈpɒsəbl ˈʃeɪps, ənd ɪn ðə ˈsɪmpləst ˈpɒsɪbl ˈnʌmbə, ɪz ˈkɔːld ə ˈsɪmpl fəʊˈniːmɪk trænˈskrɪpʃn.]

American English

[ɪf ðə 'nʌmbɚ əv 'dɪfrənt 'lɛdɚz ɪz 'mɔːr ðən ðə 'mɪnəməm æz də'faɪnd ə'bʌv ðə træn'skrɪpʃn wɪl 'nɑːt bi ə fə'niːmɪk, bəd ən ælə'fɑːnɪk wʌn. 'sʌm əv ðə 'founiːmz, 'ðæd ɪz tə 'seɪ, wɪl bɪ reprə'zentəd baɪ 'mɔːr ðən 'wʌn 'dɪfrənt 'sɪmbl. ɪn 'ʌðɚ 'wɝdz 'sʌm 'æləfounz əv 'sʌm 'founiːmz wɪl bɪ 'sɪŋgld 'aut fɚ'reprəzen'teɪʃn ɪn ðə træn'skrɪpʃn, 'hens ðə 'tɝm ælə'fɑːnɪk.]

(Both these passages are adapted from David Abercrombie, *English Phonetic Texts* [Salem, NH: Faber & Faber, 1964].)

D Transcribe the following phrases as you pronounce them, or as you think they should be pronounced in one of the well-known accents of English using (a) a fairly broad transcription, and (b) a narrower transcription.

accent _____

"Please come home."

(a)

(b)

"He is going by train."

(a)

(b)

"The angry American."

(a)

(b)

"His knowledge of the truth."

(a)

(b)

"I prefer sugar and cream."

(a)

(b)

"Sarah took pity on the young children."

(a)

(b)

PERFORMANCE EXERCISES

As noted in the preface, it is extremely important to develop practical phonetic skills at the same time as you learn the theoretical concepts. One way to do this is to learn to pronounce nonsense words. You should also transcribe nonsense words that are dictated to you. By using nonsense words you are forced to listen to the sounds that are being spoken. Accordingly, you should find another student to work with, so that you can do the following exercises in pairs.

A Learn to say simple nonsense words. A good way is to start with a single vowel, and then add consonants and vowels one by one at the beginning. In this way you are always reading toward familiar material, rather than having new difficulties ahead of you. Make up sets of words such as the following:

zɑː

ɹˈzɑː

tɹˈzɑː

ˈætɹˈzɑː

ˈmætɹˈzɑː

ʌˈmætɹˈzɑː

tʌˈmætɹˈzɑː

B Choose an order in which to say the following "words" (for example, say the second word first, the fourth word next, and then the fifth, third, and first words). Write this order down as you dictate the words to your partner—whose task is, of course, to write down the order in which you have said them. Reverse roles and repeat the exercises. You may find it advisable to repeat each word twice.

spoken   heard

piˈsuːz

piˈsuːs

piˈzuːs

piˈzuːz

piˈzuːʒ

C Repeat this exercise with the following sets of words:

spoken	heard	spoken	heard
	tɑːˈθɛð		ˈkiːpiːk
	tɑːˈθɛθ		ˈkɪpiːk
	tɑːˈðɛθ		ˈkiːpɪk

spoken	heard	spoken	heard
taɾðɛð		ˈkɪpɪk	
taɾfeð		ˈkɪpɪt	
ˈlæmæm		ˈmʌlʌl	
ˈlæmæn		ˈmʌrʌl	
ˈlænæm		ˈmʌwʌl	
ˈlænæn		ˈnʌlʌl	
ˈlænæŋ		ˈnʌrʌl	

D Look at the following nonsense words, and either say these to your partner or (preferably, since your partner has seen these words too) make up a set similar to them and say these instead. Your words can differ from the sample set in as many sounds as you like. But I suggest that you should not make them much longer at first. You will also find it advisable to write down your words and practice saying them for some time by yourself, so that you can pronounce them fluently when you say them to your partner.

ˈskɑːnziːl

ˈbraɪgbluːzd

ˈdʒɯŋsmæŋ

flɔɪʃˈθraɪðz

pjuːtˈpeɪtʃ

When you have finished saying each word several times and your partner has written the words down, compare notes. Try to decide whether any discrepancies were due to errors in saying the words or in hearing them. If possible, the speaker should try to illustrate discrepancies by pronouncing the word in both ways, saying, for example, "I said [ˈskɑːnziːl] but you wrote [ˈskɑːnsiːl]."

There is no one best way of doing ear training work of this kind. I find it helpful to look carefully at a person pronouncing an unknown word, then try to say the word myself immediately afterward, getting as much of it right as possible, but not worrying if I miss some things on first hearing. I then write down all that I can, leaving blanks to be filled in when I hear the word again. It seems important to me to get at least the number of syllables and the placement of the stress correct on first hearing, so that I have a framework in which to fit later observations.

Repeat this kind of production and perception exercise as often as you can. You should do a few minutes' work of this kind every day, so that you spend at least an hour a week doing practical exercises.

THE CONSONANTS
OF ENGLISH

This chapter will discuss the consonants of English in some detail, showing how they vary in different circumstances.

STOP CONSONANTS

Consider the difference between the words in the first column in Table 3.1 and the corresponding words in the second column. This opposition may be said to be between the set of voiceless stop consonants and the set of voiced stop consonants. But the difference is really not just one of voicing during the consonant closure, as you can see by saying these words yourself. Most people have very little voicing going on while the lips are closed during either "pie" or "buy." Both stop consonants are essentially voiceless. But in "pie," *after* the release of the lip closure, there is a moment of **aspiration,** a period of voicelessness after the stop articulation and before the start of the voicing for the vowel. If you put your hand in front of your lips while saying "pie," you can feel the burst of air that comes out during the period of voicelessness after the release of the stop.

In a narrow transcription, aspiration may be indicated by a small raised *h,* [ʰ]. Accordingly, these words may be transcribed as [pʰaɪ, tʰaɪ, kʰaɪ]. You may not be able to feel the burst of air in "tie, kye" because these stop closures are made well inside the mouth cavity. But listen carefully and notice that you can hear the period of voicelessness after the release of the stop closure in each of them. It is this interval that indicates the fact that the stop is aspirated. The major difference between the words in the first two columns is not that one has voiceless stops and the other voiced stops. It is that the first column has (voiceless) aspirated stops and the second column has (partially voiced) unaspirated stops. The amount of voicing in each of the stops [b, d, g] depends on the context in which it occurs. When it is in the middle of a sentence in which a voiced sound occurs on either side, voicing may occur throughout the stop closure. Most

| TABLE 3.1	Words illustrating allophones of English stop consonants.			
1	**2**	**3**	**4**	**5**
pie	buy	spy	nap	nab
tie	die	sty	mat	mad
kye	guy	sky	knack	nag

speakers of English have no voicing during the closure of so-called voiced stops in sentence initial position.

You should be able to hear these differences (one of the main objects of this book is to teach you to become a phonetician by learning to listen very carefully), but you can also see them in computer produced waveforms. We will discuss acoustic analyses of sounds later in this book, but we can start studying simple waveforms here. Figure 3.1 is a record of the words "tie" and "die." It is quite easy to see the different segments in the sound wave. In the first word, "tie," there is a spike indicating the burst of noise that occurs when the stop closure is released, followed by a period of very small semirandom variations during the aspiration, and then a regular, repeating wave as the vocal folds begin to

| FIGURE 3.1	The waveforms of the words "tie" and "die."
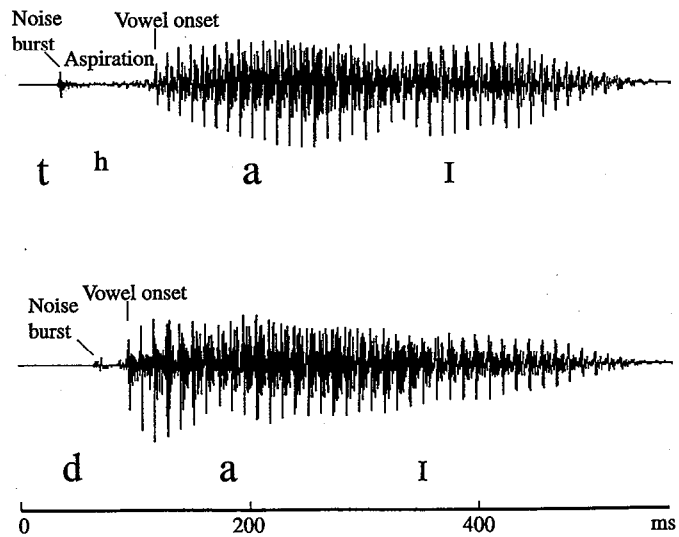
vibrate regularly for the vowel. In "die" there is no real spike at the beginning, just a somewhat irregular waveform before the start of the wave for the vowel. As you can see, the major difference between "tie" and "die" is the increase in time between the release of the stop and the start of the vowel. We will discuss this distinction further in Chapter 6.

Now consider the words in the third column of Table 3.1. Are the sounds of the stop consonants more like those in the first column or those in the second? As in many cases, English spelling is misleading, and they are in fact more like those in the second column. There is no opposition in English between words beginning with / sp / and / sb /, or / st / and / sd /, or / sk / and / sg /. English spelling uses *p, t, k,* but the stops that occur after / s / are really somewhere in between initial / p / and / b /, / t / and / d /, / k / and / g /, and usually more like the so-called voiced stops / b, d, g / in that they are completely unaspirated. Figure 3.2 shows the computer produced waveform in "sty." You can see the small variations in the waveform corresponding to the fricative / s /, followed by a straight line during the period in which there is no sound as there is a complete stop for the / t /. This is followed by a sound wave very similar to that of the / d / in Figure 3.1.

If you have access to a computer that can manipulate sounds and let you see the waveforms of words, you can verify this for yourself. Record words such as "spy, sty, sky, spill, still, skill," each said as a separate word. Now find the beginning and end of each / s /, and cut this part out. When you play the edited recordings to others and ask them to write down the words that they hear, they will almost certainly write "buy, die, guy, bill, dill, gill."

What about the differences between the words in the fourth and fifth columns of Table 3.1? The consonants at the end of "nap, mat, knack" are certainly voiceless. But if you listen carefully to the sounds at the end of the words "nab, mad,

---

**FIGURE 3.2** The waveform of the word "sty."
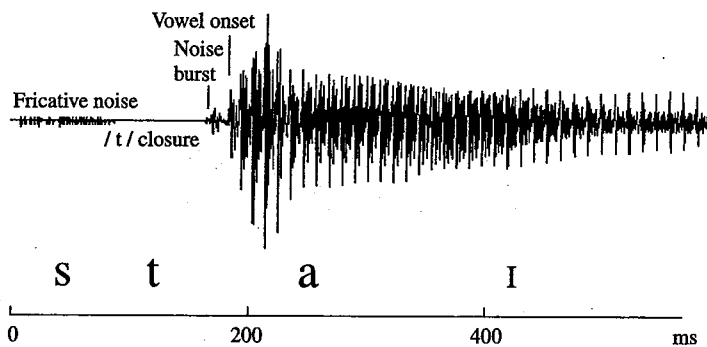
nag," you may find that the so-called voiced consonants / b, d, g / have very little voicing. Try saying these words separately. You can, of course, say each of them with the final consonant exploded and a short vowel-like sound resembling [ə] afterward. But it would be more normal to say each of them without exploding the final consonants. You could even say "cab" and not open the lips for a considerable period of time, if it were the last word of an utterance. In such circumstances it is quite clear that the final consonants are not fully voiced throughout the closure.

There is, however, a clear distinction between the words in the fourth and fifth columns. Say these words in pairs, "nap, nab; mat, mad; knack, nag," and try to decide which has the longer vowel. In these pairs, and in all similar pairs, such as "cap, cab; cat, cad; back, bag" the vowel is much shorter before the voiceless consonants / p, t, k / than it is before the voiced consonants / b, d, g /. The major difference between such pairs of words is in the vowel length, not in the voicing of the final consonants.

This length difference is very evident in Figure 3.3, which shows the waveforms of the words "mat" and "mad." The vowel in "mad" is almost twice as long as the vowel in "mat." You can also see small voicing vibrations during the / d / in "mad," but there is nothing noteworthy at the end of "mat" except the

---

FIGURE 3.3 The waveforms of the words "mat" and "mad."
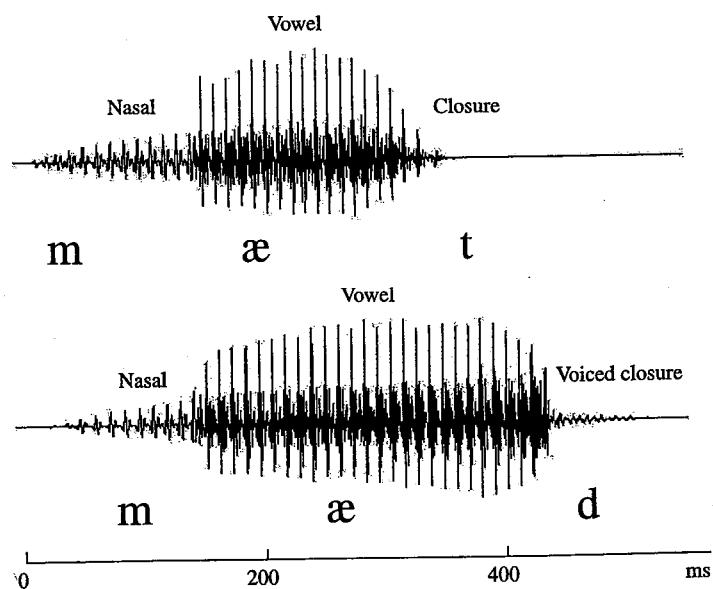
slightly irregular voicing at the time of the closure. We will return to this point later in this section.

Another way of comparing the length differences is to say short sentences such as "Take a cap now" and "Take a cab now." If you say both these sentences with a regular rhythm, you will find that the length of time between "Take" and "now" is about the same on both occasions. This is because the whole word "cap" is only slightly shorter than the whole word "cab." The vowel is much shorter in "cap" than in "cab." But the consonant / p / makes up for this by being slightly longer than the consonant / b /. It is a general rule of English (and of most other languages) that syllable final voiceless consonants are longer than the corresponding voiced consonants after this same vowel.

The phrases "Take a cap now" and "Take a cab now" also illustrate a further point about English stop consonants at the end of a word (or, in fact, at the end of a syllable). Say each of these phrases without a pause before "now." Do your lips open before the [n] of "now" begins, or do they open during the [n]? If they open before the [n], there will be a short burst of aspiration or a short vowel-like sound between the two words. But for most people, final stops are unexploded when the next word begins with a nasal. The same is true if the next word begins with a stop. The final [t] in "cat" is nearly always unexploded in phrases like "the cat pushed. . . ." In a narrow transcription we can symbolize the fact that a consonant is unexploded by adding a small raised mark [˺]. We could therefore transcribe the phrase as [ðə ˈkʰæt˺ ˈpʰʊʃt].

The same phenomenon occurs even within a word such as "apt" [æp˺t] or "act" [æk˺t]. Furthermore, across a word boundary the two consonants involved can even be identical, as in the phrase "white teeth." In order to convince yourself that there are two examples of / t / in this phrase, try contrasting it with "why teeth (are . . .)." Not only is the vowel in "white" much shorter than the vowel in "why" (because the vowel in "white" is in a syllable with a voiceless consonant at the end), but also the stop closure in "white teeth" is very much longer than the stop in the phrase with only one / t /. In "white teeth" there really are two examples of / t / involved, the first of which is unexploded.

Other languages do not have this rule. For example, it is a mark of speakers with an Italian accent (at least as caricatured in films and television) that they explode all their final stop consonants, producing an extra vowel at the end, as they normally would in their own language. Authors trying to indicate an Italian speaking English will write the sentence "It's a big day" as "It's a bigga day." They are presumably trying to indicate the difference between the normal [ɪts ə ˈbɪg˺ ˈdeɪ] and the foreign accent [ɪts ə ˈbɪgə ˈdeɪ].

It is interesting that words such as "rap, rat, rack" are all distinguishable, even when the final consonants are unexploded. The difference in the sounds must therefore be in the way that the vowels end—after all, the rest is silence. The consonants before and after a vowel always affect it, so that there is a slight but noticeable difference in its quality. Compare your pronunciation of words such

as "pip, tit, kick." Your tongue tip is up throughout the word "tit," whereas in "pip" and "kick" it remains behind the lower front teeth. In "kick" it is the back of the tongue that is raised throughout the word, and in "pip" the lip gestures affect the entire vowel. The same is true for words with voiced consonants, such as "bib, did, gig." The consonant gestures are superimposed on the vowel in such a way that their effect is audible throughout much of the syllable.

The sounds [p, t, k] are not the only voiceless stops that occur in English. Many people also pronounce a glottal stop in some words. A glottal stop is the sound (or, to be more exact, the lack of sound) that occurs when the vocal folds are held tightly together. The symbol for it is [ʔ], which is similar to a question mark without the dot.

Glottal stops occur whenever one coughs. You should be able to get the sensation of the vocal folds being pressed together by making small coughing noises. Next, take a deep breath and hold it with your mouth open. Listen to the small plosive sound that occurs when you let the breath go. Now, while breathing out through your mouth, try to check and then release the breath by making and releasing a short glottal stop. Then do the same while making a voiced sound such as the vowel [ɑ]. Practice producing glottal stops between vowels, saying [ɑʔɑ] or [iʔi], so that you get to know what they feel like.

One of the most common occurrences of a glottal stop is in the utterance meaning "no," which is often spelled "uh-uh." If someone asks you a question, you can reply "no" by saying [ˈʔʌʔʌ]. Note that there is a contrast between the utterance meaning "no" and that meaning "yes" that is dependent on the presence of the glottal stop. If you had meant to say "yes," you might well have said [ˈʌhʌ]. We can tell that it is the glottal stop that is important in conveying the meaning by the fact that one could be understood equally well by using a syllabic consonant (shown by putting the mark [ˌ] under the consonant) instead of a vowel, and saying [ˈm̩hm̩] for "yes" and [ˈʔm̩ʔm̩] for "no." As long as there is a glottal stop between the two syllables, the utterance will mean "no," irrespective of what vowel or nasal is used.

Glottal stops frequently occur as allophones of / t /. Probably most Americans and many British speakers have a glottal stop followed by a syllabic nasal in words such as "beaten, kitten, fatten" [ˈbiːʔn̩, ˈkɪʔn̩, ˈfæʔn̩]. London Cockney also has a glottal stop between vowels as in "butter, kitty, fatter" [ˈbʌʔə, ˈkɪʔi, ˈfæʔə]. Many speakers in both countries have a glottal stop just before final voiceless stops in words such as "rap, rat, rack." Usually the articulatory gesture for the other stop is still audible, so these words could be transcribed [ræʔp, ræʔt, ræʔk]. When I recorded the word "mat" for Figure 3.3, I pronounced it as [mæʔt], with the glottal stop and the closure for [t] occurring almost simultaneously.

Practice producing words with and without a glottal stop. After you have some awareness of what a glottal stop feels like, try saying the words "rap, rat,

rack" in several different ways. Begin by saying them with a glottal stop and a final relcase [ræʔpʰ, ræʔtʰ, ræʔkʰ]. Next, say them without a glottal stop and with the final stops unexploded [ræp˺, ræt˺, ræk˺]. Then say them with a glottal stop and a final unexploded consonant [ræʔp˺, ræʔt˺, ræʔk˺]. Finally, say them with a glottal stop and no other final consonant [ræʔ, ræʔ, ræʔ].

When a voiced stop and a nasal occur in the same word, as in "hidden," the stop is not released in the usual way. Both the [d] and the [n] are alveolar consonants. The tongue comes up and contacts the alveolar ridge for [d] and stays there for the nasal, which becomes syllabic [ˈhɪdn̩]. Consequently, as shown in Figure 3.4, the air pressure that is built up behind the stop closure is released through the nose by the lowering of the soft palate (the velum) for the nasal consonant. This phenomenon is known as **nasal plosion.** It is normally used in pronouncing words such as "sadden, sudden, leaden" [ˈsædn̩, ˈsʌdn̩, ˈledn̩]. It is considered a mark of a foreign accent to add a vowel [ˈsædən, ˈsʌdən, ˈledən]. Nasal plosion also occurs in the pronunciation of words with [t] followed by [n], as in "kitten" [ˈkɪtn̩], for those people who do not have a glottal stop instead of the [t].

It is worth spending some time thinking exactly how you and others pronounce words such as "kitten, button," in that it enables you to practice making detailed phonetic observations. There are a number of different possibilities.

FIGURE 3.4   Nasal plosion.
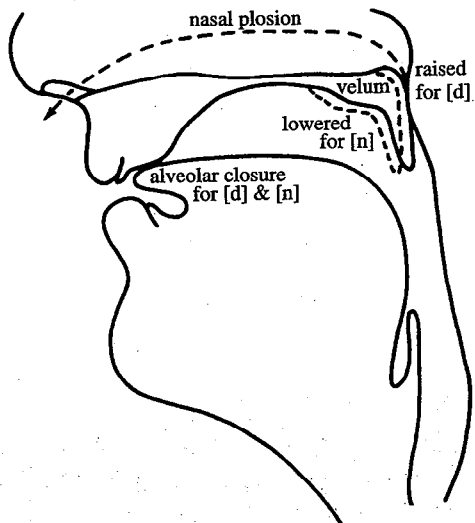
Most British and American English speakers make a glottal stop at the end of the vowel before making an alveolar closure. Then, while still maintaining the glottal stop, they lower the velum and raise the tongue for the alveolar closure. But which comes first? If they lower the velum before making the alveolar closure, there is only [ʔn] and no [t]. If they make the alveolar closure first, we could say that there is [ʔtn]; but there would not be any nasal plosion, as there would be no pressure built up behind the [t] closure. Nasal plosion occurs only if there is no glottal stop, or the glottal stop is released after the alveolar closure has been made and before the velum is lowered.

These are fairly difficult sequences to determine, but there are some simple things you can do to help you find out what articulations you use. First of all, find a drinking straw and something to drink. Put one end of the straw between your lips and hold the other end just below the surface of the liquid. Now say [ɑpɑ], and note how bubbles form during [p]. This is because pressure is built up behind your closed lips. Now push the straw slightly further into your mouth and say [ɑtɑ]. It will not sound quite right because the straw gets in the way of your tongue when it makes the alveolar closure. You may have to try different positions of the straw, but go on until you can see bubbles coming out, and convince yourself that pressure builds up behind the [t]. Now try saying "button." Of course there will be bubbles during the [b], but are there any at the end of the word, or do you have a glottal stop and no [t] behind which pressure builds up?

When two sounds have the same place of articulation, they are said to be **homorganic.** Thus the consonants [d] and [n], which are both articulated on the alveolar ridge, are homorganic. For nasal plosion to occur within a word, there must be a stop followed by a homorganic nasal. Only in these circumstances can there be pressure first built up in the mouth during the stop and then released through the nose by lowering the soft palate. Many forms of English do not have any words with a bilabial stop [p] or [b] followed by the homorganic nasal [m̩] at the end of the word. Nor in most forms of English are there any words in which the velar stops [k] or [g] are normally followed by the velar nasal [ŋ]. Consequently, both bilabial and velar nasal plosion are less common than alveolar nasal plosion in English. But when talking in a rapid conversational style, many people pronounce the word "open" as ['oʊpm̩], particularly if the next word begins with [m], as in "Open my door, please." Quite frequently when counting, people will pronounce "seven" as ['sɛbm̩], and I have also heard "something, captain, bacon" pronounced ['sʌmpm̩, 'kæpm̩, 'beɪkŋ̩]. You should try to pronounce all these words in these ways yourself.

A phenomenon similar to nasal plosion may take place when an alveolar stop [t] or [d] occurs before a homorganic lateral [l] as in "little, ladle" ['lɪtl̩, 'leɪdl̩]. The air pressure that is built up during the stop can be released by lowering the sides of the tongue; this effect is called **lateral plosion.** Say the word "middle" and note the action of the tongue. Many people (particularly British speakers)

maintain the tongue contact on the alveolar ridge through both the stop and the lateral, releasing it only at the end of the word. Others (most Americans) pronounce a very short vowel in the second syllable. For those who have lateral plosion, no vowel sound occurs in the second syllables of "little, ladle." The final consonants in all these words are syllabic. There may also be lateral plosion in words such as "Atlantic," in which the [t] may be resyllabified so that it is at the beginning of the stressed, second syllable. We should also note that most Americans, irrespective of whether they have lateral plosion or not, do not have a voiceless stop in "little." There is a general rule in American English that whenever / t / occurs after a stressed vowel and before an unstressed syllable other than [n̩], it is changed into a voiced sound. For those Americans who have lateral plosion, this will be the stop [d].
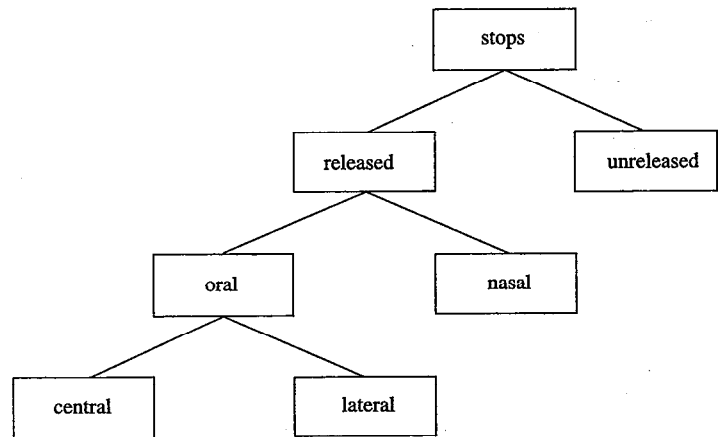
This brings us to another important point about coronal stops and nasals. For many speakers, including most Americans, the consonant between the vowels in words such as "city, better, writer" is not really a stop but a quick tap in which the tongue tip is thrown against the alveolar ridge. This sound can be written with the symbol [ɾ] (instead of a [d] as I have been doing in similar words up till now) so that "city" can be transcribed as ['sɪɾi]. Many Americans also make this kind of tap when / d / occurs after a stressed vowel and before an unstressed vowel. As a result they do not distinguish between pairs of words such as "latter, ladder." But some maintain a distinction by having a shorter vowel in words such as "latter," which have a voiceless consonant in their underlying form. It is as if the statement that vowels are shorter before voiceless consonants had applied first, and then a later rule was applied changing [t] into [ɾ] when it occurred between a stressed and an unstressed syllable.

We can summarize the discussion of stop consonants by thinking of the possibilities that there are in the form of a branching diagram as shown in Figure 3.5. The first question to consider is whether the stop is released (exploded) or not. If it is released, then is it oral plosion or is the release due to the lowering of the velum, with air escaping through the nose, making it nasal plosion? If it is oral plosion, then is the closure in the mouth entirely removed, or is the articulation in the midline retained and one or both sides of the tongue lowered so that air escapes laterally? You should be able to produce words illustrating all these possibilities. For coronal stops, there is an additional point not shown in Figure 3.5, namely, is the [t] or [d] sound produced as a tap [ɾ]?

FRICATIVES

The fricatives of English vary less than the stop consonants, yet the major allophonic variations that do occur are in many ways similar to those of the stops. Earlier we saw that when a vowel occurs before one of the voiceless stops / p, t, k /, it is shorter than it would be before one of the voiced stops / b, d, g /. The

FIGURE 3.5   Stop consonant releases.

```
                          ┌───────┐
                          │ stops │
                          └───────┘
                     ┌────────┴────────┐
              ┌──────────┐        ┌────────────┐
              │ released │        │ unreleased │
              └──────────┘        └────────────┘
              ┌──────┴──────┐
         ┌──────┐       ┌───────┐
         │ oral │       │ nasal │
         └──────┘       └───────┘
         ┌────┴────┐
    ┌─────────┐ ┌─────────┐
    │ central │ │ lateral │
    └─────────┘ └─────────┘
```

same kind of difference in vowel length occurs before voiceless and voiced fricatives. The vowel is shorter in the first word of each of the pairs "strife, strive" [straɪf, straɪv]; "teeth, teethe" [tiːθ, tiːð]; "rice, rise" [raɪs, raɪz]; and "mission, vision" ['mɪʃn̩, 'vɪʒn̩].

Stops and fricatives are the only English consonants that can contrast by being either voiced or voiceless. Consequently, we can revise our statement that vowels are shorter before voiceless stops than before voiced stops. Instead, we can say that vowels are shorter before all voiceless consonants than before all voiced consonants. In this way we can capture a linguistically significant generalization that would have been missed if our statements about English had included two separate statements, one dealing with stops and the other dealing with fricatives.

We also saw that a voiceless stop at the end of a syllable (as in "hit") is longer than the corresponding voiced stop (as in "hid"). Similarly, the voiceless fricatives are longer than their voiced counterparts in each of the pairs "safe, save" [seɪf, seɪv], "lace, laze" [leɪs, leɪz], and all the other pairs of words we have been discussing in this section. Again, because fricatives behave like stops, a linguistically significant generalization would have been missed if we had regarded each class of consonants completely separately.

Fricatives are also like stops in another way. Consider the degree of voicing that occurs in the fricative at the end of the word "ooze," pronounced by itself. In most pronunciations, the voicing that occurs during the final [z] does not last throughout the articulation but changes in the last part to a voiceless sound like [s]. In general, voiced fricatives at the end of a word, as in "prove, smooth,

choose, rouge" [pruːv, smuːð, tʃuːz, ruːʒ] are voiced throughout their articulation only when they are followed by another voiced sound. In a phrase such as "prove it," the [v] is fully voiced because it is followed by a vowel. But in "prove two times two is four" or "try to improve," where the [v] is followed by a voiceless sound [t] or by a pause at the end of the phrase, it is not fully voiced.

Briefly stated, then, fricatives are like stops in three ways: first, stops and fricatives influence vowel length in similar ways—vowels before voiceless stops or fricatives are shorter than before voiced stops or fricatives; second, final voiceless stops and fricatives are longer than final voiced stops and fricatives; and third, the final stops and fricatives that are classified as voiced are not actually voiced throughout the articulation unless the adjacent sounds are also voiced. In addition, both these types of articulation involve an obstruction of the airstream. Because they have an articulatory feature in common and because they act together in phonological statements, we refer to fricatives and stops together as a natural class of sounds called **obstruents.**

However, fricatives do differ from stops in that they sometimes involve actions of the lips that are not immediately obvious. Try saying "fin, thin, sin, shin" [fɪn, θɪn, sɪn, ʃɪn]. There is clearly a lip action in the first word as it involves the labiodental sound [f]. But do your lips move in any of the other three words? Most people find that their lips move slightly in any word containing / s / ("sin, kiss") and quite considerably in any word containing / ʃ / ("shin, dish"), but that there is no lip action in words containing / θ / ("thin, teeth"). There is also lip movement in the voiced sounds corresponding to / s / and / ʃ /, namely / z / as in "zeal, zest" and / ʒ / in "leisure, treasure," but none in / ð /, as in "that, teethe."

The primary articulation in these fricatives is the close approximation of two articulators so that friction can be heard. The lip rounding is a lesser articulation in that the two articulators (the lower lip and the upper lip) approach one another but not sufficiently to cause friction. A lesser degree of closure by two articulators not involved in the primary articulation is called a secondary articulation. This particular one, in which the action of the lips is added to another articulation, is called labialization. The English fricatives / ʃ, ʒ / are strongly labialized, and the fricatives / s, z / are slightly labialized.

AFFRICATES

This is a convenient place to review the status of affricates in English. An affricate is simply a sequence of a stop followed by a homorganic fricative. Some such sequences, for example the dental affricate [tθ] as in "eighth" or the alveolar affricate [ts] as in "cats," have been given no special status in English phonology. They have been regarded just as consonant clusters comparable with those at the end of "lapse" and "sacks" (which are not affricates, as the stops and the fricatives are not homorganic). But, as we noted in the discussion of symbols

for transcribing English, it is appropriate to regard the sequences [tʃ] and [dʒ] as different from other sequences of consonants. They are the only affricates in English that can occur at both the beginning and the end of words. In fact, even the other affricates that can occur at the end of words will usually do so only as the result of the formation of a plural or some other suffix as in "eighth." From the point of view of a phonologist considering the sound pattern of English, the palato-alveolar affricates are plainly single units.

NASALS

The nasal consonants of English vary even less than the fricatives. Nasals, together with [r, l], can be syllabic when they occur at the end of words. As we have seen, the mark [ͺ] under a consonant indicates that it is syllabic. (Vowels, of course, are always syllabic and therefore need no special mark.) In a narrow transcription, we may transcribe the words "sadden, table" as ['sædn̩, 'teɪbl̩]. In most pronunciations, "prism, prison" can be transcribed ['prɪzm̩, 'prɪzn̩], as these words do not usually have a vowel between the last two consonants. Syllabic consonants can also occur in phrases such as "Jack and Kate" ['dʒæk n̩ 'keɪt] and "not until" [nɑt n̩'tɪl].

The nasal [ŋ] differs from the other nasals in a number of ways. No English word can begin with [ŋ]. This sound can occur only within or at the end of a word, and even in these circumstances it does not behave like the other nasals. It can be preceded only by the vowels / ɪ, ɛ, æ, ʊ, ə, ɑ, ɔ /, and it cannot be syllabic (except in slightly unusual pronunciations, such as "bacon" ['beɪkŋ̩], and phrases such as "Jack and Kate" mentioned previously).

One way to consider the different status of [ŋ] is that it is a sequence of the phonemes / n / and / g /. Looking at it this way, "sing" would be phonemically / sɪng / and "sink" would be / sɪnk /. There is a rule saying that / n / has the allophone [ŋ] whenever it occurs before / g / and / k /, turning [sɪng] into [sɪŋg] and [sɪnk] into [sɪŋk]. Another rule allows for the dropping of / g / (but not of / k /) whenever it occurs after [ŋ] at the end of either a word (as in "sing") or a stem followed by a suffix such as "-er" or "ing." In this way the / g / is dropped in "singer," which contains a suffix "-er," but is retained in "finger," in which the "-er" is not a suffix. This second rule would not apply in the case of some speakers from the New York area who make "singer" rhyme with "finger."

APPROXIMANTS

The voiced approximants are / w, r, j, l / as in "whack, rack yak, lack." The first three of these sounds are central approximants, and the last is a lateral approximant. The articulation of each of them varies slightly depending on the

articulation of the following vowel. You can feel that the tongue is in a different position in the first sounds of "we" and "water." The same is true for "reap" and "raw," "lee" and "law," and "ye" and "yaw." Try to feel where your tongue is in each of these words.

These consonants also share the possibility of occurring in consonant clusters with stop consonants. The approximants / r, w, 1 / combine with stops in words such as "pray, bray, tray, dray, cray, gray, twin, dwell, quell, Gwen, play, blade, clay, glaze." The approximants are largely voiceless when they follow one of the voiceless stops / p, t, k / as in "play, twice, clay." This voicelessness is a manifestation of the aspiration that occurs after voiceless stops, which we discussed at the beginning of this chapter. At that time I introduced a small raised h symbol [ʰ], which could be used to show that the first part of the vowel was voiceless. When there is a consonant rather than a vowel immediately following the stop, we can use the diacritic [̥] to indicate that the consonant is voiceless. We can transcribe the words "play, twice, clay," in which there are approximants after initial voiceless plosives, as [pl̥eɪ, tw̥aɪs, kl̥eɪ]. The approximant / j / as in "you" [ju] can occur in similar consonant clusters, as in "pew, cue" [pju, kju] and, for speakers of British English, "tune" [tj̥un]. We will discuss the sequence [ju] again when we consider vowels in more detail.

Lip rounding is an essential part of / w /. Because there is a tendency for the articulations during one sound to anticipate those in the following sound, stops are slightly rounded when they occur in clusters in which / w / is the second element, as in "twice, dwindle, quick" [tw̥aɪs, 'dwɪndl̩, kw̥ɪk]. We will discuss this phenomenon, which is known as anticipatory coarticulation, in a later section.

In many people's speech / r / also has some degree of lip rounding. Try saying words such as "reed" and "heed." Do you get some movement of the lips in the first word but not in the second? Note also whether you get anticipatory lip rounding so that the stops [t, d] are slightly rounded in words such as "tree, dream."

In most forms of British English, there is a considerable difference in the articulation of / l / before a vowel, as in "leaf" or "feeling," as compared with / l / before a consonant or at the end of a word, as in "field" or "feel." In most forms of American English, there is less distinction between these two kinds of / l /. Note the articulation of / l / in your own pronunciation. Try to feel where the tongue is during the / l / in "leaf." You will probably find that the tip is touching the alveolar ridge, and one or both sides are near the upper side teeth, but not quite touching. Now compare this articulation with the / l / in "feel." Most (but not all) speakers make this sound with the tongue tip touching the alveolar ridge. But in both British and American English the center of the tongue is pulled down and the back is arched upward as in a back vowel. If there is contact on the alveolar ridge, it is the primary articulation. The arching upward of the back of the tongue forms a secondary articulation, which we will call velarization. In most forms of American English, all examples of / l / are comparatively velarized, except, perhaps, those that are syllable initial and between high front vowels, as

in "freely." In British English / l / is usually not velarized when it is before a vowel, as in "lamb" or "swelling," but it is velarized when word final or before a consonant, as in "ball" or "swelter."

The symbol for velarization is the mark [~] through the middle of the symbol. Accordingly, a narrow transcription of "feel" would be [fiɫ]. In my own speech, the whole body of the tongue is drawn up and back in the mouth so that the tip of the tongue no longer makes contact with the alveolar ridge. Strictly speaking, therefore, this sound is not an alveolar consonant but more like some kind of back vowel in the speech of some English speakers.

Finally, we must consider the status of / h /. Earlier I suggested that English / h / is somewhat like the voiceless counterpart of the surrounding sounds. At the beginning of a sentence, / h / is like a voiceless vowel, but / h / can also occur between vowels in words or phrases like "behind the head." As you move from one vowel through / h / to another, the articulatory movement is continuous, and the / h / is signaled by a weakening of the voicing, which may not even result in a completely voiceless sound.

In many accents of English, / h / can occur only before stressed vowels or before the approximant / j / as in "hue" [hju]. Some speakers of English also sound / h / before / w /, so that they contrast "which" [hwɪtʃ] and "witch" [wɪtʃ]. The symbol [ʍ] (an inverted w) is sometimes used for this voiceless approximant. The contrast between / w / and / ʍ / is disappearing in most forms of English, so that in those dialects in which it occurs, [ʍ] is more likely to be found only in the less common words such as "whether" rather than in frequently used words such as "what."

RULES FOR ENGLISH CONSONANT ALLOPHONES

A good way of summarizing (and slightly extending) all that I have said about English consonants so far is to list a set of rules describing the allophones. These rules are simply *descriptions* of language behavior. They are not the kind of rules that prescribe what people ought to do. Like most phoneticians, I would not presume to set myself up as an arbiter of fashion and declare what constitutes "good" speech. To the extent that phonetics is part of an exact scientific discipline, I would like to be able to formalize my description of speech in terms of a set of precise statements. But these statements should be regarded as descriptive, not prescriptive, rules.

Some of the rules apply to all the consonants of English. One of the rules dealing with consonant length is the following:

(1) Consonants are longer when at the end of a phrase.

You can see the application of this rule by comparing the consonants in words such as "bib, did, don, nod." If you can, make a recording of these words, and

then play the recording backward. Are the first two words the same backward and forward? Do the third and fourth words sound like each other when played in reverse?

Most of the rules apply to only selected groups of consonants, as in (2) through (18) as follows:

(2) Voiceless stops / p, t, k / are aspirated when they are syllable initial, as in words such as "pip, test, kick" [pʰɪp, tʰɛst, kʰɪk].

(3) Obstruents—stops and fricatives—classified as voiced (that is, / b, d, g, v, ð, z, ʒ /) are voiced through only a small part of the articulation when they occur at the end of an utterance (as the / v / in "Try to improve.") or before a voiceless sound (as the / d / in "add two").

(4) So-called voiced stops and affricates / b, d, g, dʒ / are voiceless when syllable initial, except when immediately preceded by a voiced sound.

(5) The approximants / w, r, j, l / are at least partially voiceless when they occur after initial / p, t, k /, as in "play, twin, cue" [pl̥eɪ, tw̥ɪn, kj̥u].

(6) Voiceless stops / p, t, k / are unaspirated in words such as "spew, stew, skew."

(7) Voiceless stops and affricates / p, t, k, tʃ / are longer than the corresponding voiced stops and affricates / b, d, g, dʒ / when at the end of a syllable.

Words exemplifying this rule are "cap" as opposed to "cab" and "back" as opposed to "bag." Try contrasting these words in sentences, and you may be able to hear the differences more clearly.

(8) Stops are unexploded when they occur before another stop in words such as "apt" [æp̚t] and "rubbed" [rʌb̚d].

(9) In many accents of English, syllable final / p, t, k / are accompanied by a glottal stop, as in pronunciations of "tip, pit, kick" as [tɪʔp, pɪʔt, kɪʔk].

This rule does not apply to all varieties of English. Some people do not have any glottal stops in these circumstances, and others have glottal stops completely replacing some or all of the voiceless stops. In any case, even for those who simply add a glottal stop, the rule is not completely accurate as stated. Many people will have a glottal stop at the end of "cat" in phrases such as "that's a cat" or "the cat sat on the mat," but they will not have this allophone of / t / in "The cat is on the mat." An accurate statement would require my giving a much better account of what is meant by a syllable, but I cannot do this because it is not possible to give a phonetic definition of a syllable. We will return to this point in Chapter 10.

(10) In many accents of English, / t / is replaced by a glottal stop when it occurs before an alveolar nasal in the same word, as in "beaten" [ˈbiːʔn̩].

(11) Nasals are syllabic at the end of a word when immediately after an obstruent as in "leaden, chasm" ['lɛdn̩, 'kæzm̩].

Note that we cannot say that nasals become syllabic whenever they occur at the end of a word and after a consonant. The nasals in "kiln, film" are not syllabic in most accents of English. We can, however, state a rule describing the syllabicity of / l / by saying simply the following:

(12) The lateral / l / is syllabic at the end of a word when immediately after a consonant.

This rule summarizes the fact that / l / is syllabic not only after stops and fricatives (as in "paddle, whistle" ['pædl̩, 'wɪsl̩]), but also after nasals (as in "kennel, channel" ['kɛnl̩, 'tʃænl̩]). The only problem with this rule is what happens after / r /. It is correct for words such as "barrel" ['bærl̩], but it does not work in most forms of American English in words such as "snarl" [snɑrl], when / r / has to be considered as part of the vowel.

When it is not part of the vowel, / r / is like / l / in most forms of American English in that it, too, can be syllabic when it occurs at the end of a word and after a consonant, as in "sabre, razor, hammer, tailor" ['seɪbr̩, 'reɪzr̩, 'hæmr̩, 'teɪlr̩]. If we introduce a new term, **liquid,** which is used simply as a cover term for the consonants / l, r /, we may rephrase the statement in (12) and say:

(12a) The liquids / l, r / are syllabic at the end of a word when immediately after a consonant.

The next rule also applies more to American English than to British English. It accounts for the / t / in "fatty, data" ['fæɾi, 'dæɾə]. But note that these are not the only contexts in which these changes occur. This is not simply a change that affects / t / after a stressed vowel and before an unstressed one, in that / t / between two unstressed vowels (as in "divinity") is also affected. However, not all cases of / t / between vowels change in this way. The / t / in "attack" (i.e., before a stressed syllable) is voiceless, and / t / after another consonant (for example, in "hasty, captive") is also voiceless. Note also that most American English speakers have an articulatory gesture very like that in [ɾ] in words containing / d / and / n / in similar circumstances, such as "daddy, many." The first of these two words could well be transcribed ['dæɾi]. The second has the same sound, except that it is nasalized, so it could be transcribed ['mɛɾ̃i] in a narrow transcription. Nasalization is shown by the diacritic [˜] over a symbol. The following rule accounts for all these facts:

(13) Alveolar stops become voiced taps when they occur between two vowels, the second of which is unstressed.

Many speakers of American English require a similar rule to describe a sequence of an alveolar nasal followed by a stop. In words such as "painter, splinter," the / t / is lost and a nasal tap occurs. This has resulted in "winter" and

"winner" and "panting" and "panning" being pronounced in the same way. For these speakers we can restate (13) this way:

(13a) Alveolar stops and alveolar nasal plus stop sequences become voiced taps when they occur between two vowels, the second of which is unstressed.

There is a great deal of variation among speakers with respect to this rule. Some make taps in familiar words such as "auntie" but not in less common words such as "Dante." Some make them only in fast speech. Try to formulate a rule in a way that describes your own speech.

(14) Alveolar consonants become dentals before dental consonants, as in "eighth, tenth, wealth" [eɪt̪θ, tɛn̪θ, wɛl̪θ]. Note that this rule applies to all alveolar consonants, not just stops, and it often applies across word boundaries, as in "at this" [æt̪ ðɪs].

In a more rapid style of speech, some of these dental consonants tend to be omitted altogether. Say these words first slowly and then more rapidly, and see what you do yourself.

It is difficult to give rules for when consonants get deleted, because this depends so much on the style of speech being used. Most people say "most people" as [ˈmoʊs ˈpipl̩] with no audible [t], and they produce phrases such as "sand paper" with no audible [d]. We could write this rule as follows:

(15) Alveolar stops are reduced or omitted when between two consonants.

Rule (15) raises an interesting point of phonetic theory. Note that I said "there may be no audible [d]," and "alveolar stops are reduced or omitted." However, the tongue tip gesture for the alveolar stop in "most people" may not be omitted but is just not audible because it is completely overlapped by the following labial stop. More commonly, it is partially omitted; that is to say, the tongue tip moves up for the alveolar stop but does not get to make a complete closure. When we think in terms of phonetic symbols, we can write [ˈmoʊs ˈpipl̩] or [ˈmoʊst ˈpipl̩]. But it is not really a question of whether the [t] is there or not. Part of the tongue tip gesture may have been made; but that is a fact that we have no way of symbolizing. Check this rule for yourself by thinking how you say phrases such as "best game" and "grand master." Say these and similar phrases with and without the alveolar stop. You may find the rule needs rewording to take into account all the contexts where alveolar stops may not appear in your speech.

Another rule is needed to account for the shortening effects that occur when two identical consonants come next to each other, as in "big game, top post." It is usually not true to say that one of these consonants is dropped: even in casual speech, most people would distinguish between "stray tissue, straight issue, straight tissue." (Try saying these in sentences such as "That's a stray tissue" and see for yourself.) But there clearly is a shortening effect that we can state as follows:

(16) A consonant is shortened when it is before an identical consonant.

We need rules stating not only where consonants are dropped and shortened, but also where they are added. Words such as "something" and "youngster" often are pronounced as ['sʌmpθɪŋ] and ['jʌŋkstɚ]. In a similar way, many people do not distinguish between "prince" and "prints," or "tense" and "tents." All these words are pronounced with a short voiceless stop between the nasal and the voiceless fricative. The insertion of a sound into the middle of a word is known as **epenthesis.**

(17) A homorganic voiceless stop may be inserted after a nasal before a
voiceless fricative followed by an unstressed vowel in the same word.

Note that it is necessary to mention that the following vowel must be unstressed. Speakers who have an epenthetic stop in the noun "concert" do not usually have one in verbal derivatives such as "concerted" or in words such as "concern." Nothing need be said about the vowel before the nasal. Epenthesis may—like the [t] to [ɾ] change in rule (13)—occur between unstressed vowels. I have heard an inserted [t] in both "agency" and "grievances."

We can describe the more front articulation of / k / in "cap, kept, kit, key" [kæp, kɛpt, kɪt, ki] and of / g / in "gap, get, give, geese" [gæp, gɛt, gɪv, gis]. You should be able to feel the more front position of your tongue contact in the latter words of these series. The rule is as follows:

(18) Velar stops become more front as the following vowel in the same syl-
lable becomes more front.

Finally, we need to note the difference in the quality of / l / in "life" [laɪf] and "file" [faɪɫ], or "clap" [klæp] and "talc" [tæɫk], or "feeling" [fiːlɪŋ] and "feel" [fiːɫ].

(19) The lateral / l / is velarized when after a vowel or before a consonant at
the end of a word.

DIACRITICS

In this and the previous chapter we have seen how the transcription of English can be made more detailed by the use of diacritics, small marks added to a symbol to narrow its meaning. The six diacritics we have introduced so far are shown in Table 3.2. You should learn the use of these diacritics before you attempt any further detailed transcription exercises. Note that the nasalization diacritic is a small wavy line above a symbol, and the velarization diacritic is that same wavy line through the middle of a symbol. Nasalization is more common among vowels, which will be discussed in the next chapter.
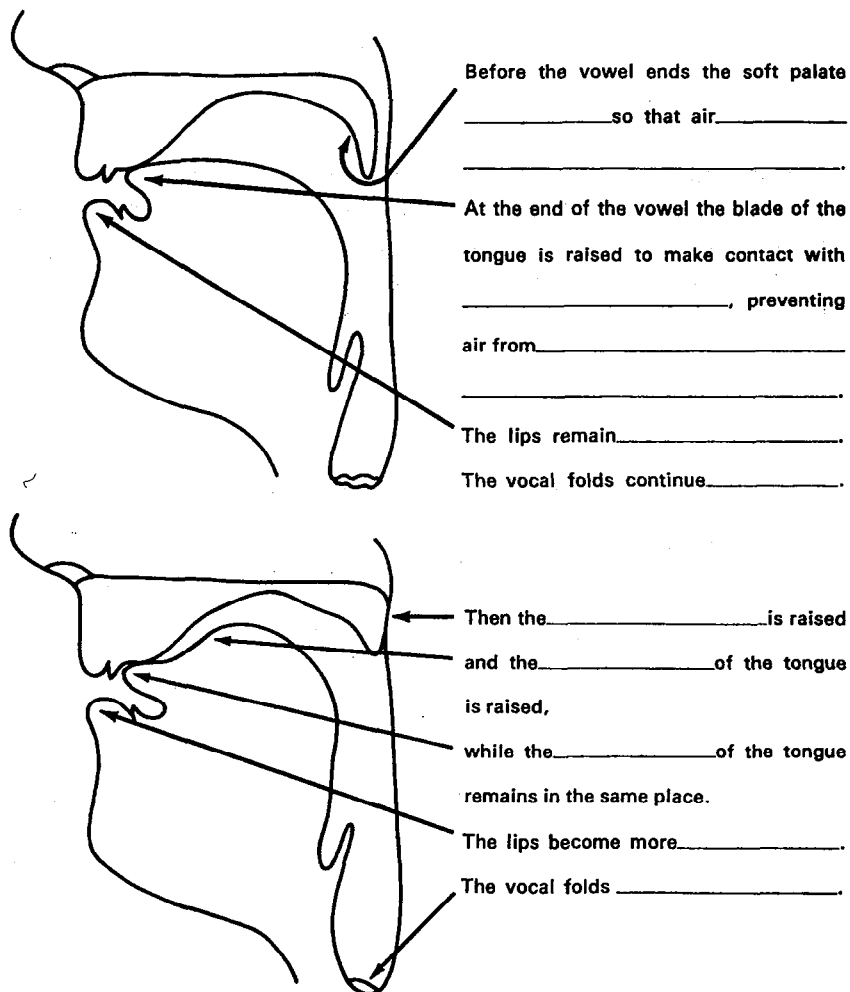
| TABLE 3.2 | | | Some diacritics that modify the value of a symbol. | |

	Voiceless	w̥	l̥	kwɪk, pleɪs	"quick, place"
h	Aspirated	tʰ	kʰ	tʰæp, kʰɪs	"tap, kiss"
̪	Dental	t̪	d̪	æt̪ ðə, hel̪θ	"at the, health"
~	Nasalized	r̃	æ̃	mæ̃n	"man"
~	Velarized	ɫ		pʰɪɫ	"pill"
ˌ	Syllabic	n̩	l̩	'mɪʔn̩	"mitten"

EXERCISES

A The sequence of annotated diagrams that follows illustrates the actions that take place during the consonants at the end of the word "branch." Fill in the blanks.

Before the vowel ends the soft palate _____ so that air _____

_____.

At the end of the vowel the blade of the tongue is raised to make contact with

_____, preventing

air from _____

_____.

The lips remain _____.
The vocal folds continue _____.

Then the _____ is raised

and the _____ of the tongue is raised,

while the _____ of the tongue remains in the same place.

The lips become more _____.

The vocal folds _____.

The_____remains_____

throughout the end of the word.

After a short period the_____

_____moves downward, but the

_____remains close

enough to the alveolar ridge to_____

_____.

The_____are_____.

The vocal folds _____.

B Annotate the following diagrams so as to describe the actions required for the consonants in the middle of the word "implant." Make sure that your annotations mention the action of the lips, the different parts of the tongue, the soft palate, and the vocal folds in each diagram. Try to make clear which of the vocal organs moves first in going from one consonant to another. The pronunciation illustrated is that of a normal conversational utterance; note the position of the tongue during the bilabial stop.
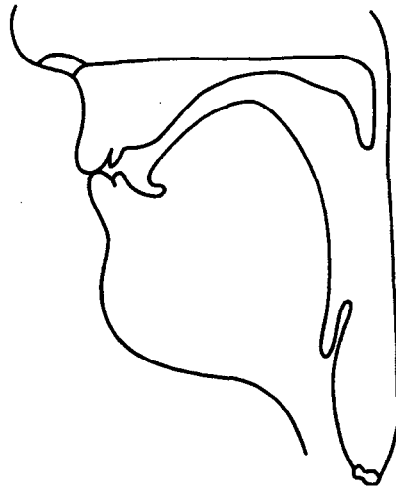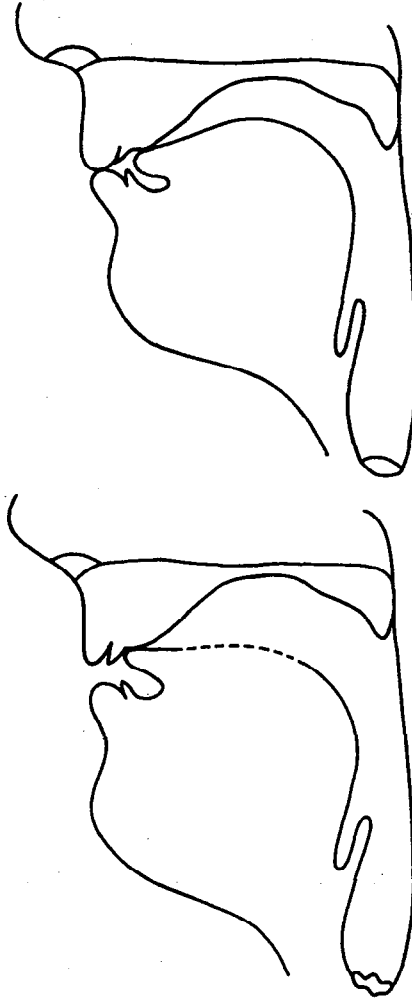
C Draw and annotate diagrams similar to those in the previous exercises, but this time illustrate the actions that occur in pronouncing the consonants in the middle of the phrase "thick snow." Make sure that you show clearly the sequence of events, noting what the lips, tongue, soft palate, and vocal folds do at each moment. Say the phrase over to yourself several times at a normal speed before you begin. Note especially whether the back of your tongue lowers before or after the tip of the tongue forms the articulation for subsequent consonants.

D As a transcription exercise, give a number of examples for each of rules (2) through (19) by making a narrow transcription of some additional words that fit the rules. Your examples should not include any words that have been transcribed in this book so far. Remember to mark the stress on words of more than one syllable.

Rule (2)   three examples (one for each voiceless stop)

_____   _____   _____

Rule (3)   seven examples (one for each voiced obstruent)

_____   _____   _____

_____   _____   _____

_____

Rule (4)   eight examples (two for each voiced stop or affricate)

_____   _____

_____   _____

_____   _____

_____   _____

Rule (5)   four examples (one for each approximant)

_____   _____

_____   _____

Rule (6)   three examples (one for each voiceless stop)

_____   _____   _____

Rule (7)   four contrasting pairs (one for each place of articulation)

_____   _____

_____   _____

_____   _____

_____   _____

Rule (8)   six examples (one for each voiced and voiceless stop)

_____   _____   _____

_____   _____   _____

Rule (9)   three examples (not necessarily from your own speech)

_____   _____   _____

Rule (10)   three examples (use three different vowels)

_____   _____   _____

Rule (11)   three examples (use at least two different nasals)

_____   _____   _____

Rule (12a)   six examples (three each with / l / and / r /)

_____   _____   _____

_____   _____   _____

Rule (13a)   six examples (two each with / t, d, n /, one being after an unstressed vowel)

_____   _____   _____

_____   _____   _____

Rule (14)   three examples (one each for / t, d, n /)

_____   _____   _____

Rule (15)   three examples (any kind)

_____   _____   _____

Rule (16)   three examples (any kind)

_____   _____   _____

Rule (17)   two examples (use two different nasals)

_____   _____   _____

Rule (18)   four examples (use four different vowels)

_____   _____

_____   _____

Rule (19)   two contrasting pairs (try to make them reversible words)

_____   _____

_____   _____

E As a more challenging exercise, try to list two exceptions to some of these rules.

Rule (   ) _____

Rule (   ) _____

F Write a rule that describes the allophones of / h /.

_____

_____

_____

_____

PERFORMANCE EXERCISES

A Learn to produce some non-English sounds. First, in order to recall the sensation of adding and subtracting voicing while maintaining a constant articulation, repeat the exercise saying [ssszzzssszzz]. Now try a similar exercise, saying [mmmm̥m̥m̥mmmm̥m̥m̥]. Make sure that your lips remain together all the time. During [m̥] you should be producing exactly the same action as when breathing out through the nose. Now say [m̥] in between vowels, producing sequences such as [am̥a, im̥i], and so on. Try not to have any gap between the consonant and the vowels.

B Repeat this exercise with [n, ŋ, l, r, w, j] learning to produce [an̥a, aŋ̥a, al̥a, ar̥a, aʍa, aj̥a] and similar sequences with other vowels.

C Make sure that you can differentiate between the English words "whether, weather; which, witch," even if you do not normally do so. Say the following:

[hweðər]   "whether"

[weðər]   "weather"

[hwɪtʃ]   "which"

[wɪtʃ]   "witch"

D Learn to produce the following Burmese words. (You may for the moment neglect the tones, indicated by accents above the vowels.)

Voiced Nasals	Voiceless Nasals	Voiced Lateral	Voiceless Lateral
mâ "lift up"	m̥â "from"	la "moon"	l̥á "beautiful"
nă "pain"	n̥ă "nose"		
ŋâ "fish"	ŋ̥a "borrow"		

E Working with a partner, produce and transcribe _several_ sets of nonsense words. You should use slightly more complicated sets than previously. Make up your own sets on the basis of the illustrative set given here, including glottal stops, nasal and lateral plosion, and some combinations of English sounds that could not occur in English. Remember to mark the stress.

'kl̥ɑːntʃʊps'kweɪdʒ

'ʒiːʒm̩'spobm̩

'tsɪʔɾ'bɛʔɪdl̩

mbuː'tɾɪgŋ

'twaɪbreʔɪp

F In order to increase your memory span in perceiving sounds, include some simpler but longer words in your production–perception exercises. A set of possible words is given here. You may find words such as the last two, which have eight syllables each, too difficult at the moment. But try to push your hearing ability to its limit. When you are listening to your partner dictating words, remember to try to (1) look at the articulatory movements; (2) repeat, to yourself, as much as you can immediately afterward; and (3) write down as much as you can, including the stress, as soon as possible.

'kiːpuːtuɾ'piːkiːtuː

'begɾ'gɪdɛ'dɛdɪ

tɾiː'tʃɾʔiːtʃuɾ'druːdʒi

'riːle'tole'mɑnu'dʊli

'faɪθiðiɾ'vɔɪðuvu'θiːfi

4

ENGLISH VOWELS

ENGLISH PHONETIC DICTIONARIES

As we saw in Chapter 2, the vowels of English can be transcribed in many different ways, partly because accents of English differ greatly in the vowels they use and partly because there is no one right way of transcribing even a single accent of English. The set of symbols used depends on the reason for making the transcription. If one is aiming to reduce English to the smallest possible set of symbols, then "sheep" and "ship" could be transcribed as [ʃiːp, ʃip]. If one were working on problems of computer speech recognition, one might choose to emphasize differences in vowel quality and write [ʃip, ʃɪp]. In a book such as this, in which we want to note that both length and quality differences occur, then [ʃiːp, ʃɪp] is the preferable transcription, despite the fact that vowel quality and vowel length are linked, and there is no need to mark both.

There is another reason why this style of transcription, introduced in Chapter 2, is the preferable one for this book. In previous editions it was hard to find anything that one could point to as a widely accepted, authoritative publication that transcribed both British and American English. Now there are some good reference books that specify pronunciations in both languages. One is an updated version of the dictionary produced by the English phonetician Daniel Jones, whose acute observations of English dominated British phonetics in the first half of the twentieth century. The current edition, *English Pronouncing Dictionary*, 15th ed. (Cambridge: Cambridge University Press) is familiarly known as EPD 15. It still bears Daniel Jones's name, but it has been completely revised by the new editors, Peter Roach and James Hartman, so that it now shows both British and American pronunciations. Another authoritative work is the *Longman Pronunciation Dictionary* (Harlow, U.K.: 2nd. ed. Pearson Education, 2000) by John Wells. This dictionary, which is familiarly known as LPD2, also gives the British and American pronunciations of more than 75,000 words. Professor John Wells holds the chair in phonetics at University College, London, previously held by Daniel Jones. He is clearly the leading authority on contemporary English pronunciation in all its forms—British, American, and other variants of the worldwide language. A third dictionary, *The Oxford Concise Dictionary of*

Pronunciation by Clive Upton, William Kretzschmar, and Rafal Konopka (Oxford: Oxford University Press, 2000), is due out at the same time as this book. It is slightly different from the other two dictionaries in that it gives a wider range of both British and American pronunciations and also uses a larger set of symbols, aiming to show more detail, using a more allophonic transcription than either of the other two dictionaries.

Everyone seriously interested in English pronunciation should be using one of these dictionaries. Each of them shows the pronunciations typically used by national newscasters—what we may regard as Standard American Newscaster English and Standard BBC English (often shortened to just American English and British English in this book). Of course, in neither country is there really a standard accent. In both countries there are some newscasters who have notable local accents. The dictionaries give what would be accepted as reasonable pronunciations for communicating in the two countries. They allow one to compare British and American pronunciations in great detail, noting, for example, that most British speakers pronounce "Caribbean" as [kærɪ'biən], with the stress on the third syllable, whereas Americans typically say [kə'rɪbiən], with the stress on the second syllable.

Ordinary American college dictionaries also provide pronunciations, but the symbols they use are not in accordance with the principles of the International Phonetic Association (IPA) and are of little use for comparative phonetic purposes. American dictionary makers sometimes say that they deliberately do not use IPA symbols because their dictionaries are used by speakers with different regional accents, and they want readers to be able to learn how to pronounce an unfamiliar word correctly in their own accents. But, as we have been observing, IPA symbols are often used to represent broad regions of sounds, and there is no reason why dictionary makers should not assign them values in terms of key words, just as they do for their ad hoc symbols.

Two of the three dictionaries we have been discussing, LPD and EPD 15, use virtually the same set of symbols, differing only in the way they transcribe the vowel in American English "bird"; LPD has[ɚ], whereas EPD 15 has [ɜr]. *The Oxford Concise Dictionary of Pronunciation* uses a slightly different set of symbols, but those symbols are readily interpretable within the IPA tradition. In this book I have kept to the style of transcription used in Wells's LPD with one exception, which is simply a typographical change: I have used [ɛ] in words such as "head, bed" instead of [e]. In later chapters we will be comparing vowels in other languages such as French and German, so we will need to use both [e] and [ɛ].

VOWEL QUALITY

In the discussion so far I have deliberately avoided making precise remarks about the quality of the different vowels. This is because, as I said in Chapter 1, the traditional articulatory descriptions of vowels are not very satisfactory. Try

asking people who know as much about phonetics as you do to describe where the tongue is at the beginning of the vowel in "boy," and you will get a variety of responses. Can you describe where your own tongue is in a set of vowels?

It is difficult to give a meaningful answer to requests to describe the tongue position of a vowel in one's own speech. Very often people can only repeat what the books have told them, because they cannot find out for themselves where their tongue is. It is quite easy for a book to build up a set of terms that are not really descriptive but are in fact only labels. I started introducing terms of this kind for vowel qualities in Chapters 1 and 2 and will continue with this procedure here. But it is important for you to remember that the terms we are using are simply labels that describe how vowels sound in relation to one another. They are not absolute descriptions of the position of the body of the tongue.

Part of the problem in describing vowels is that there are no distinct boundaries between one type of vowel and another. When talking about consonants, the categories are much more distinct. A sound may be a stop or a fricative, but it cannot be halfway between the two. Vowels are different. It is perfectly possible to make a vowel that is halfway between a high vowel and a mid vowel. In theory (as opposed to what a particular individual can do in practice), it is possible to make a vowel at any specified distance between any two other vowels.

In order to appreciate the fact that vowel sounds form a continuum, try gliding from one vowel to another. Say [æ] as in "had" and then try to move gradually to [i] as in "he." Do not say just [æi], but try to spend as long as possible on the sounds in between them. If you do this correctly, you should pass through sounds that are something like [ɛ] as in "head" and [eɪ] as in "hay." If you have not achieved this effect already, try saying [æ-ɛ-eɪ-i] again, slurring slowly from one vowel to another.

Now do the same in the reverse direction, going slowly and smoothly from [i] as in "he" to [æ] as in "had." Take as long as possible over the in-between sounds. You should learn to stop at any point in this continuum so that you can make, for example, a vowel like [ɛ] as in "head," but slightly closer to [æ] as in "had."

Next, try going from [æ] as in "had" slowly toward [ɑ] as in "father." When you say [æ-ɑ], you probably will not pass through any other vowel of your own speech. But there is a continuum of possible vowel sounds between these two vowels. You may be able to hear sounds between [æ] and [ɑ] that are more like those used by other accents in "had" and "father." Some forms of Scottish English, for example, do not distinguish between the vowels in these words (or between "cam" and "calm"). Speakers with these accents pronounce both "had" and "father" with a vowel about halfway between the usual Midwestern American pronunciation of these two vowels. Some speakers of American English in the Boston area pronounce words such as "car, park" with a vowel between the more usual American vowels in "cam" and "calm." They do, however, distinguish the latter two words.

Last, in order to appreciate the notion of a continuum of vowel sounds, glide from [ɑ] as in "father" to [u] as in "who." In this case, it is difficult to be specific as to the vowels that you will go through on the way, because English accents differ considerably in this respect. But you should be able to hear that the movement from one of these sounds to the other covers a range of vowel qualities that have not been discussed so far in this section.

When you move from one vowel to another, you are changing the auditory quality of the vowel. You are, of course, doing this by moving your tongue and your lips; but, as we have noted, it is very difficult to say exactly how your tongue is moving. Consequently, because phoneticians cannot be very precise about the positions of the vocal organs in the vowels they are describing, they often simply use labels for the auditory qualities of the different vowels. The vowel [i] as in "heed" is called high front, meaning that it has the auditory quality high and the auditory quality front. Similarly, the vowel [æ] as in "had" has an auditory quality that may be called low front; and [ɛ] as in "head" sounds somewhere between [i] and [æ], but a little nearer to [æ], so we call it mid-low front. (Say the series [i, ɛ, æ] and check for yourself that this is true.) The vowel [ɑ] as in "father" is a low back vowel, and the vowel [u] in "who" is a high, fairly back vowel. The four vowels [i, æ, ɑ, u], therefore, give us something like the four corners of a space showing the auditory qualities of vowels, which may be drawn as in Figure 4.1.

None of the vowels has been put in an extreme corner of the space in Figure 4.1. It is possible to make a vowel that sounds more back than the vowel [u] that most people use in "who." You should be able to find this fully back vowel for yourself. Start by making a long [u], then round and protrude your lips a bit more. Now try to move your tongue back in your mouth while still keeping it

---

FIGURE 4.1   The vowel space.
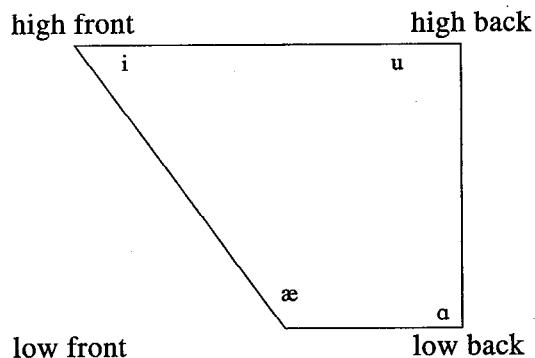
raised toward the soft palate. The result should be a fully back [u]. Another way of making this sound is to whistle the lowest note that you can, and then while retaining the same tongue and lip position, voice this sound. Again, the result will be an [u] sound that is farther back than the vowel in "who." Try saying [i] as in "heed," [u] as in "who," and then this new sound, which we may symbolize with an added underline [u̲]. If you say the series [i, u, u̲], you should be able to hear that [u] is intermediate between [i] and [u̲], but—for most speakers—much nearer [u̲].

Similarly, it is possible to make vowels with a more extreme quality than the usual English vowels [i, æ, ɑ]. If, for example, while saying [æ] as in "had," you lower your tongue or open your jaw slightly farther, you will produce a vowel that sounds relatively farther from [i] as in "heed." It will probably also sound a little more like [ɑ] as in "father."

Given a notion of an auditory vowel space of this kind, we can plot the relative quality of the different vowels. Remember that the labels high–low and front–back should not be taken as descriptions of tongue positions. They are simply indicators of the way one vowel *sounds* relative to another. The labels describe the relative auditory qualities, not the articulations.

Students of phonetics often ask why we use terms like high, low, back, and front if we are simply labeling auditory qualities and not describing tongue positions. The answer is that it is largely a matter of tradition. For many years phoneticians thought that they were describing tongue positions when they used these terms for specifying vowel quality. But there is only a rough correspondence between the traditional descriptions in terms of tongue positions and the actual auditory qualities of vowels. If you could take x-ray pictures that showed the position of your tongue while you were saying the vowels [i, æ, ɑ, u], you would find that the relative positions were not as indicated in Figure 4.1. But, as we will see in Chapter 8, if you use acoustic phonetic techniques to establish the auditory qualities, you will find that these vowels do have the relationships indicated in this figure.

At one time I thought of inventing new names to describe the auditory qualities. Other linguists have, in fact, used terms such as grave and acute instead of back and front in the description of vowels. But, for a variety of reasons, these terms did not become widely used, and it seems preferable to stick with the old terms high, low, back, and front, even though they are being used to describe auditory qualities rather than tongue positions.

Most of the vowels of a form of Standard American Newscaster English typical of many Midwestern speakers are shown in the upper part Figure 4.2. A comparable diagram of the vowels of British English as spoken by BBC newscasters is shown in the lower part of Figure 4.2. In both diagrams, the solid points represent the vowels that we are treating as monophthongs, and the lines represent the movements involved in the diphthongs. The symbols labeling the diphthongs are placed near their origins. There is a good scientific basis for placing the vowels as

FIGURE 4.2   The relative auditory qualities of some of the vowels of Standard American Newscaster English and of British (BBC newscaster) English.

high

American
(national newscaster)

low

high

British
(BBC newscaster)

low
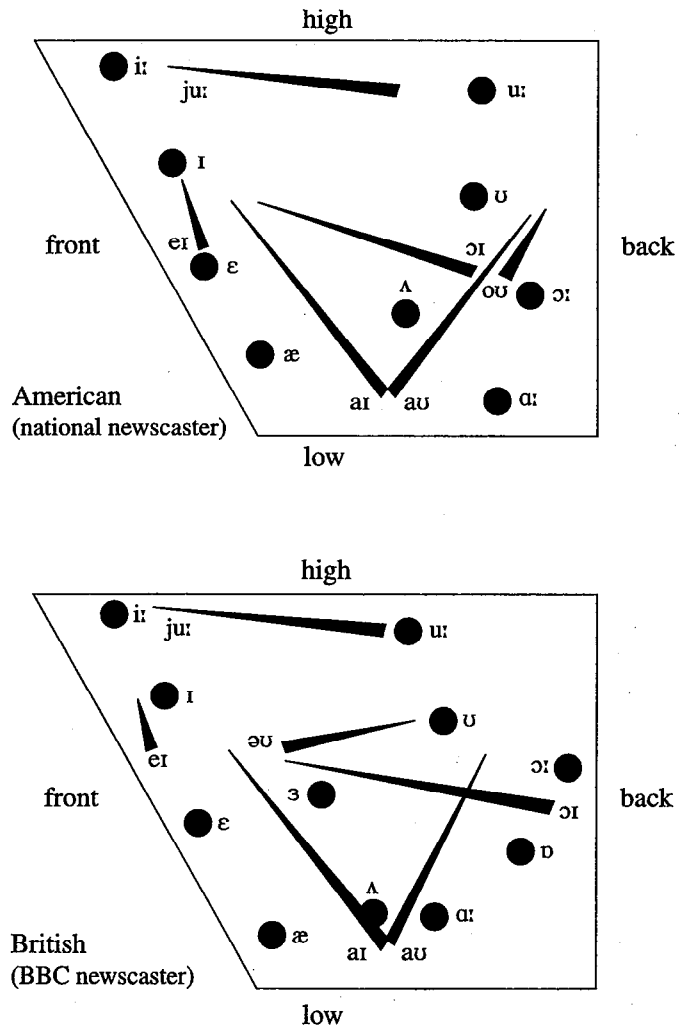
shown here. The positions of both monophthongs and diphthongs are not just the result of my own listening. The data are taken from the acoustic analyses of a number of authorities. We will return to this point in Chapters 8 and 9 when we discuss acoustic phonetics. Meanwhile, if you are able to listen to a speaker of Midwestern American English or BBC English, you should be able to hear that

the relative vowel qualities are as indicated. Other varieties of English will differ in some respects, but you should find that in most accents the majority of the relationships are the same. We will note the cases in which there are substantial differences as we discuss the individual vowels.

Listen first of all to the vowels [iː, ɪ, ɛ, æ] as in "heed, hid, head, had." Do they sound as if they differ by a series of equal steps? Make each vowel about the same length (although in actual words they differ considerably), saying just [i, ɪ, ɛ, æ]. Now say them in pairs, first [i, ɪ], then [ɪ, ɛ], then [ɛ, æ]. In many forms of English, [i] sounds about the same distance from [ɪ] as [ɪ] is from [ɛ], and as [ɛ] is from [æ]. Some Eastern American speakers make a distinct diphthong in "heed" so that their [i] is really a glide starting from almost the same vowel as that in "hid." Other forms of English, for example, as spoken in the Midlands and the North of England, make a lower and more back vowel in "had," making it sound a little more like the [ɑ] in "father." This may result in the distance between [ɛ] and [æ] being greater than that between [ɛ] and [ɪ]. But speakers who have a lower [æ] may also have a slightly lower [ɛ], thus keeping the distances between the four vowels [i, ɪ, ɛ, æ] approximately the same.

The remaining front vowel in English is [eɪ] as in "hay." We will discuss this vowel after we have discussed some of the back vowels. The back vowels vary considerably in different forms of English, but no form of English has them evenly spaced like the front vowels. Say for yourself [ɑː, ɔː, ʊ, u] as in "father, author, good, food." As before, make each vowel about the same length, and say just [ɑ, ɔ, ʊ, u]. Consider pairs of vowels as you did the front vowels. Estimate the distances between each of these vowels, and compare them with those shown in Figure 4.2.

We have already noted that many Midwestern and Californian speakers do not distinguish [ɑː] and [ɔː] as in "cot" and "caught." They usually have a vowel intermediate in quality between the two points shown on the chart. On the other hand, most speakers of British English have an additional vowel in this area. They distinguish between the vowels [ɑː, ɒ, ɔː] as in "balm, bomb, bought." This results in a different distribution of the vowel qualities, as shown in the lower diagram in Figure 4.2. The additional vowel [ɒ] is more back and slightly more rounded than [ɑ].

The vowels [ʊ, u] as in "good, food" also vary considerably. Many speakers have a very unrounded vowel in "good" and a rounded but central vowel in "food." Look in a mirror and observe your own lip positions in these two vowels.

Both British and American English speakers have a mid-low central vowel [ʌ] as in "bud." In many forms of British English, this vowel may be a little lower than in American English. In this way it is distinct from the British English central vowel [ɜ] in "bird." The vowel in American English "bird" is not shown in the upper part of Figure 4.2 as it is distinguished from the vowel in "bud" by having r-coloring, which we will discuss later.

We must now consider the diphthongs shown in Figure 4.2. Each of these sounds involves a change in quality within the one vowel. As a matter of convenience, they can be described as movements from one vowel to another. The first part of the diphthong is usually more prominent than the last. In fact, the last part is often so brief and transitory that it is difficult to determine its exact quality. Furthermore, the diphthongs often do not begin and end with any of the sounds that occur in simple vowels.

As you can see from Figure 4.2, both of the diphthongs [aɪ, aʊ], as in "high, how," start from more or less the same low central vowel position, midway between [æ] and [ɑ] and, in BBC English, closer to [ʌ] than to any of the other vowels. (*The Oxford Concise Dictionary of Pronunciation* transcribes our [aɪ] as [ʌɪ] in British English.) Say the word "eye" very slowly and try to isolate the first part of it. Compare this sound with the vowels [æ, ʌ, ɑː] as in "bad, bud, father." Now make a long [ɑː] as in "father," and then say the word "eye" as if it began with this sound. The result should be something like some forms of New York or London Cockney English pronunciation of "eye." Try some other pronunciations, starting, for example, with the vowel [æ] as in "bad." In this case, the result is a somewhat affected pronunciation.

The diphthong [aɪ], as in "high, buy," moves toward a high front vowel, but in most forms of English it does not go much beyond a mid front vowel. Say a word such as "buy," making it end with the vowel [ɛ] as in "bed" (as if you were saying [baɛ]). A diphthong of this kind probably has a smaller change in quality than occurs in your normal pronunciation (unless you are one of the speakers from Texas or elsewhere in the South and Southwest who make such words as "by, die" into long monophthongs—[baː, daː]). Then say "buy," deliberately making it end with the vowel [ɪ] as in "bid." This vowel is usually slightly higher than the ending of this diphthong for many speakers of English. Finally, say "buy" with the vowel [i] as in "heed" at the end. This is a much larger change in quality than normally occurs in this word. But some speakers of Scottish English and Canadian English have a diphthong of this kind in words such as "sight," which is different from the diphthong that they have in "side."

The diphthong [aʊ] in "how" usually starts with a very similar quality to that at the beginning of "high." Try to say "owl" as if it started with [æ] as in "had," and note the difference from your usual pronunciation. Some speakers of the type of English spoken around London and the Thames estuary (often called Estuary English) have a complicated movement in this diphthong, making a sequence of qualities like those of [ɛ] as in "bed," [ʌ] as in "bud," and [u] as in "food." Say [ɛ-ʌ-u] in quick succession. Now say the phrase "how now brown cow" using a diphthong of this type.

The diphthong [eɪ] as in "hay" varies considerably in different forms of English. Some American English speakers have a diphthong starting with a vowel very like [ɛ] in "head" (as shown in the upper part of Figure 4.2). Most BBC English speakers and many Midwestern Americans have a smaller diphthong,

starting closer to [ɪ] as in "hid." London and Estuary English has a larger diph-
thong, so that words such as "mate, take" sound somewhat like "might, tyke."
Conversely, others (including many Scots) have a higher vowel, a monophthong
that can be written [e]. Check your own pronunciation of "hay" and try to
decide how it should be represented on a chart as in Figure 4.2.

The diphthong [oʊ] as in "hoe" may be regarded as the back counterpart of
[eɪ]. In American English it is usually a movement in the high–low dimension,
but in most forms of British English the movement is more in the front–back
dimension, as you can see in Figure 4.2. Some British English speakers make
this vowel start near [ɛ] and end a little higher than [ʊ]. Say each part of this
diphthong and compare it with other vowels.

The remaining diphthong moving in the upward direction is [ɔɪ] as in "boy."
Again, this diphthong does not end in a very high vowel. It often ends with a
vowel similar to that in "bed." I might well have transcribed "boy" as [bɔɛ], if
I had not been trying to keep the style of transcription used in this book as simi-
lar as possible to other widely used transcriptions.

The last diphthong, [ju] as in "cue," differs from all the other diphthongs
in that the more prominent part occurs at the end. Because it is the only vowel
of this kind, many books on English phonetics do not even consider it as a
diphthong; they treat it as a sequence of a consonant followed by a vowel and
symbolize it by [ju] (or [yu], in the case of books not using the IPA system of
transcription). I have considered it to be a diphthong because of the way it pat-
terns in English. Historically, it is a vowel, just like the other vowels we have
been considering. Furthermore, if it is not a vowel, then we have to say that
there is a whole series of consonant clusters in English that can occur before
only one vowel. The sounds at the beginning of "pew, beauty, cue, spew, skew"
and (for most speakers of British English) "tune, dune, sue, Zeus, new, lieu,
stew" occur only before / u /. There are no English words beginning with / pje /
or / kjæ /, for example. In stating the distributional properties of English
sounds, it seems much simpler to recognize / ju / as a diphthong and thus reduce
the complexity of the statements one has to make about the English consonant
clusters.

The only common stressed vowel of American English not shown in Figure
4.2 is [ɝ] as in "sir, herd, fur." This vowel does not fit on the chart because it
cannot be described simply in terms of the features high–low, front–back, and
rounded–unrounded. The vowel [ɝ] can be said to be *r-colored.* It involves an
additional feature called **rhotacization.** Just like high–low and front–back, the
feature rhotacization describes an auditory property, the r-coloring, of a vowel.
When we describe the height of a vowel, we are saying something about how it
sounds rather than something about the tongue position necessary to produce it.
Similarly, when we describe a sound as a rhotacized vowel, we are saying some-
thing about how it sounds. In most forms of American English, there are both
stressed and unstressed rhotacized vowels. The transcription for the phrase "my
sister's bird" in most forms of American English would be [maɪ ˈsɪstɚz ˈbɝd].

Rhotacized vowels are often called retroflex vowels, but there are at least two distinct ways in which the r-coloring can be produced. Some speakers have the tip of the tongue raised, as in a retroflex consonant, but others keep the tip down and produce a high bunched tongue position. These two gestures produce a very similar auditory effect. X-ray studies of speech have shown that in both these ways of producing a rhotacized quality there is usually a constriction in the pharynx caused by retraction of the part of the tongue below the epiglottis.

The most noticeable difference among accents of English is in whether they have r-colored vowels or not. In many forms of American English, rhotacization occurs when vowels are followed by [r], as in "beard, bared, bard, board, poor, tire, hour." Accents that permit some form of [r] after a vowel are said to be **rhotic.** The rhotacization of the vowel is often not so evident at the beginning of the vowel, and something of the quality of the individual vowel remains. But in "sir, herd, fur" the whole vowel is rhotacized (which is why I prefer LPD2 ['bɝd], rather than EPD 15 ['bɜrd]). Insofar as the quality of this vowel can be described in terms of the features high–low and front–back, it appears to be a mid-central vowel such as [ɜ] with added rhotacization.

Rhotic accents are the norm in most parts of North America. They were prevalent throughout Britain in Shakespeare's time and still occur in the West Country, Scotland, and other regions distant from London. Shortly after it became fashionable in the Southeast of England to drop post-vocalic / r /, this habit spread to areas of the United States in New England and parts of the South. These regions are now nonrhotic to various degrees. Try to find a speaker of English with an accent that is the opposite of your own—rhotic or nonrhotic as the case may be. Listen to the speaker's vowels in words such as "mirror, fairer, surer, poorer, purer" and compare them with your own.

Standard BBC English is not rhotic and has diphthongs (not shown in Figure 4.2) going from a vowel near the outside of the vowel space toward the central vowel [ə]. In words such as "here" and "there," these are transcribed [ɪə] and [ɛə]. Some speakers have a long [ɛː] instead of [ɛə], particularly before [r] as in "fairy, bearing." Some people have a centering diphthong [ʊə] in words such as "poor," but this is probably being replaced in most nonrhotic accents of British English by [ɔː]. We also noticed in Chapter 2 that some speakers have a centering diphthong (though we did not call it that at the time) in "hire, fire," which are [haə, faə].

UNSTRESSED SYLLABLES

In all forms of English, the symbol [ə], which is not shown in Figure 2.2, may be used to specify a range of mid-central vowel qualities. As we saw in Chapter 2, vowels of this sort occur in grammatical function words, such as "to, the, at" [tə, ðə, ət]. They also occur at the ends of the words "sofa, China"

['soʊfə, 'tʃamə] and, for most British speakers, "better, farmer" ['bɛtə, 'faːmə]. In American English the vowel at the end of words with the "-er" spelling is usu-ally [ɚ], a very similar quality, but with added r-coloring. As the vowel chart in Figure 4.2 represents a kind of auditory space, vowels near the outside of the chart are more distinct from one another than vowels in the middle, and differ-ences in vowel quality become progressively reduced among vowels nearer the center. The symbol [ə] may be used to designate many vowels that have a **reduced vowel** quality.

We will be considering the nature of stress in English in the next chapter, but we can note here that vowels in unstressed syllables do not necessarily have a completely reduced quality. All the English vowels can occur in unstressed syl-lables in their full (unreduced) forms. Many of them can occur in three forms, as shown in Table 4.1. In this table, the vowel to be considered is in boldface. The words in the first column illustrate the full forms of the vowels. The second col-umn gives an example of the same unreduced vowel in an unstressed syllable. The last column illustrates the same underlying vowel as a reduced vowel. For many people, the reduced vowels in this last column are all very similar. Some accents have slightly different qualities in some of these words, but still within the range of a mid-central vowel that can be symbolized by [ə]. Others have [ɪ] in some of these words, such as "recitation," or a high-central vowel, which

	TABLE 4.1	Examples of vowels in stressed and unstressed syllables and in reduced syllables. The boldface type shows the vowel under consideration.

	Stressed Syllable	**Unstressed Syllable**	**Reduced Syllable**
iː	depreciate	create	deprecate
ɪ	implicit	simplistic	implication
eɪ	explain	chaotic	explanation
ɛ	allege	tempestuous	allegation
æ	emphatic	fantastic	emphasis
ɑː, ɒ	demonstrable	prognosis	demonstration
ɔː	cause	causality	
oʊ, əʊ	invoke	vocation	invocation
ʊ	hoodwink	neighborhood	
uː	acoustic	acoustician	
ʌ	confront	umbrella	confrontation
ɝ, ɜː	confirm	verbose	confirmation
aɪ	recite	citation	recitation
aʊ	devout	outsider	
ɔɪ	exploit	exploitation	
juː	compute	computation	circular

may be symbolized by [i]—a symbol that is sometimes called "barred i." Say all these words yourself and find out which vowels you have.

There are some widely applicable rules of English relating the pronunciation of the words in the first column to that of the words in the third column. Consequently, we are able to say that the same underlying vowels occur in the words in the first and third columns. If we were making a high level phonological transcription, we could transcribe the vowels in the different columns with the same symbols and allow the rules to make it clear that different allophones occurred. Thus we could transcribe "emphatic" as / ɛmfætɪk / and "emphasis" as / ɛmfæsɪs /, as long as we also have a rule that assigns the stress and makes / æ / into [ə] in the second word.

The rules accounting for the allophones are very general in the sense that they account for thousands of similar alternations among English words. But they are also very complicated. They have to account for the blanks in the third column, which show that some vowels can be completely reduced but others cannot. There is, for example, a completely reduced vowel in "explanation, demonstration, recitation," but not, for most people, in the very similar words "exploitation, computation." As you can also see from an examination of Table 4.1, some vowels, such as [ʊ, uː], do not fit into this scheme of alternations in the same way as the other vowels. Because the rules are so complicated, we will not use transcriptions showing the underlying forms of English in this elementary textbook. Instead, we will continue to use [ə] or [ɪ] in reduced syllables.

Most British and some American English speakers have a vowel more like [ɪ] in suffixes such as "-ed, -(e)s" at the ends of words with alveolar consonants such as "hunted, houses" [ˈhʌntɪd, ˈhaʊzɪz]. For these speakers, both vowels in "pitted" [ˈpɪtɪd] have much the same quality. A reduced vowel more like [ʊ] may occur in the suffix "-ful" as in "dreadful" [ˈdredfʊl], but for many people this is just a syllabic [l̩], [ˈdrɛdfl̩].

TENSE AND LAX VOWELS

The vowels of English can be divided into what may be called **tense** and **lax** sets. These terms are really just labels that are used to designate two groups of vowels that behave differently in English words. There are phonetic differences between the two groups, but they are not simply a matter of "tension." To some extent the differences between the two sets are due to facts in the history of English that are still represented in the spelling. The tense vowels occur in the words with a final, so-called silent *e* in the spelling (e.g., "mate, mete, kite, cute"). The lax vowels occur in the corresponding words without a silent *e*: "mat, met, kit, cut." In addition, the vowel in "good," which, for reasons connected with the history of English, has no silent *e* partner, is also a member of the lax set. This spelling based distinction is, however, only a rough indication of the difference between the two sets. It is better exemplified by the data in Table 4.2.

| TABLE 4.2 | | The distribution of tense and lax vowels in stressed syllables in American English. | | | | |

Tense Vowels	Lax Vowels	Most Closed Syllables	Open Syllables	Syllables Closed by [r]	Syllables Closed by [ŋ]	Syllables Closed by [ʃ]
iː		beat	bee	beer		(leash)
	ɪ	bit			sing	wish
eɪ		bait	bay			
	ɛ	bet		bare	length	fresh
	æ	bat			hang	crash
ɑː		hot	pa	bar		slosh
ɔː		bought	saw	bore	long	(wash)
oʊ		boat	low	(boar)		
	ʊ	good				push
uː ·		boot	boo	poor		
	ʌ	but			hung	crush
aɪ		bite	buy	fire		
aʊ		bout	bough	hour		
ɔɪ		void	boy	(coir)		
ju		cute	cue	pure		

The difference between the two sets can be discussed in terms of the different kinds of syllables in which they can occur. Table 4.2 shows some of the restrictions for one form of American English. The first column of words illustrates a set of **closed syllables**—those that have a consonant at the end. All of the vowels can occur in these circumstances. The next column shows that in **open syllables**—those without a consonant at the end—only a restricted set of vowels can occur.

None of the vowels [ɪ, ɛ, æ, ʊ, ʌ] as in "bid, bed, bad, good, bud" can appear in stressed open syllables. This is the set of vowels that may be called lax vowels as opposed to the tense vowels in the other words. In order to characterize the differences between tense and lax vowels, we can consider some of them in pairs, each pair consisting of a tense vowel and the lax vowel that is nearest to it in quality. Three pairs of this kind are [iː, ɪ] as in "beat, bit"; [eɪ, ɛ] as in "bait, bet"; and [uː, ʊ] as in "boot, foot." In each of these pairs, the lax vowel is shorter, lower, and slightly more centralized than the corresponding tense vowel. There are no vowels that are very similar in quality to the remaining two lax vowels in most forms of American English, [æ] as in "hat, cam" and [ʌ] as in "hut, come." But both of these low lax vowels are shorter than the low tense vowel [ɑː] as in "spa." Speakers of most forms of British English have an additional lax vowel. They have the tense vowel [ɑː] as in "calm, car, card" in both open and closed syllables; and they also have a lax vowel [ɒ] as in "cod, common, con" [kɒd, 'kɒmən, kɒn], which occurs only in closed syllables.

The fifth column in Table 4.2 shows the vowels that can occur in syllables closed by / r / in American English. There is no contrast between a tense vowel and the lax vowel nearest to it in quality in a syllable closed by / r /. Consequently, as often happens in contexts in which there is no opposition between two sounds, the actual sound produced is somewhere between the two. (We have already observed another example of this tendency. We saw that after / s / at the beginning of a word, there is no contrast between / p / and / b /, or / t / and / d /, or / k / and / g /. Consequently, the stops that occur in words such as "spy, sty, sky" are in between the corresponding voiced and voiceless stops; they are unaspirated, but they are never voiced.)

I put the words "boar" and "coir" in parentheses in this column because for many people [oʊ] and [ɔɪ] do not occur before / r /. The word "coir" [kɔɪr], which is the only word I have heard pronounced with [ɔɪr], is not in many people's vocabulary; and many people make no difference between "bore" and "boar." But some speakers do contrast [ɔː] and [oʊ] in these two words, or in other pairs such as "horse" and "hoarse."

The next column shows the vowels that occur before [ŋ]. In these circumstances, again, there is no possible contrast between tense and lax vowels. But, generally speaking, it is the lax vowels that occur. However, many younger Americans pronounce "sing" with a vowel closer to that in "beat" rather than to that in "bit." And in some accents, "length" is regularly pronounced with virtually the same vowel as that in "bait" rather than that in "bet"; in others, it is pronounced with the vowel in "bit." The pronunciation of "long" varies. It is [lɑːŋ] or [lɔːŋ] in most forms of American English and [lɒŋ] in most forms of British English. Several other changes are true of vowels before all nasals in many forms of American English. For example, [æ] may be considerably raised in "ban, lamb" as compared with "bad, lab." In many accents, "pin, pen" and "gym, gem" are not distinguished.

The last column shows that there are similar restrictions in the vowels that can occur before [ʃ]. By far the majority of words ending in / ʃ / have lax vowels for most speakers, although some accents (e.g., that used in parts of Indiana) have [iː] in "fish" (making it like "fiche") and [uː] in "push" and "bush." In my own speech, the only words containing the tense vowel / iː / before / ʃ / are "leash, fiche, quiche." Some speakers have tense vowels in a few new or unusual words such as "creche, gauche," which may be [kreɪʃ, goʊʃ]. The pronunciation of "wash" varies in much the same way as that of "long." Both [wɑːʃ] and [wɔːʃ] occur in American English.

RULES FOR ENGLISH VOWEL ALLOPHONES

We can conclude this chapter as we did the previous one—by considering some rules that apply to vowels. The first concerns vowel length:

(1) A given vowel is longest in an open syllable, next longest in a syllable closed by a voiced consonant, and shortest in a syllable closed by a voiceless consonant.

If you compare words such as "sea, seed, seat" or "sigh, side, site," you will hear that the vowel is longest in the first word in each set, next longest in the second, and shortest in the last. You can see an example of part of this rule in Figure 3.3, which showed the waveforms of the words "mat" and "mad." Because some vowels are inherently longer than others, we have to restrict rule (1) to a vowel of a given quality. On many occasions the so-called short vowel in "bid" may be longer than the so-called long vowel in "beat."

Even when we are considering the same vowel in syllables with the same consonants, there may be a difference in vowel length. Stressed syllables are longer than the corresponding unstressed syllables. Compare words such as "below" and "billow." You will find that the vowel [oʊ] in the stressed syllable in the first word is longer than the same vowel in the second word, where it occurs in an unstressed syllable. We will therefore have a rule:

(2) Other things being equal, vowels are longer in stressed syllables.

We still had to hedge this rule with the phrase "other things being equal," as there are other causes of variation in vowel length. Another kind of length variation is exemplified by sets of words such as "speed, speedy, speedily." Here the vowel in the stressed syllable gets progressively shorter as a result of adding extra syllables in the same word. The reasons for this phenomenon will be dealt with in the next chapter. Here we will simply state the following:

(3) Other things being equal, vowels are longest in monosyllabic words, next longest in words with two syllables, and shortest in words with more than two syllables.

We should also add a statement about unstressed vowels, which may become voiceless in words such as "potato, catastrophe." For some people, this happens only if the following syllable begins with a voiceless stop; but for many (including myself), it also happens in a normal conversational style in words such as "permission, tomato, compare." One wording of this rule would be as follows:

(4) A reduced vowel may be voiceless when it occurs after a voiceless stop (and before a voiceless stop).

The parenthesized phrase could be omitted for many people.

A further extension of this rule would allow for completely voiceless syllables when the unstressed vowel follows a voiceless stop cluster with / r /, as in "*pre*paratory, spec*tro*graph, in*tro*duction." What about clusters with / l /, such as "re*pli*cate, com*pli*cate"? Some people make some of these syllables voiceless as well.

Other rules can be used to specify the changes due to the vowel being influenced by the following consonant. The most obvious is that vowels tend to become nasalized before nasal consonants. In a word such as "ban," the soft palate often lowers for the nasal considerably before the tongue tip rises to make the articulatory contact. As a result, much of the vowel is nasalized. As we saw in the previous chapter, nasalization is shown by the diacritic [˜] over a symbol. In a narrow transcription; "ban" might be transcribed [bæ̃n]. We can make the following general statement:

(5)   Vowels are nasalized in syllables closed by a nasal consonant.

The degree of nasalization in a vowel varies extensively. Many people will have the velum lowered throughout a syllable beginning and ending with a nasal, such as "man," making the vowel fully nasalized.

Finally, we must note the allophones produced when vowels occur in syllables closed by / l /. Compare your pronunciation of / iː / in "heed" and "heel," of / eɪ / in "paid" and "pail," and [æ] in "pad" and "pal." In each case, you should be able to hear a noticeably different vowel quality before the velarized [ɫ]. All the front vowels become considerably retracted in these circumstances. It is almost as if they became diphthongs with an unrounded form of [ʊ] as the last element. In a narrow transcription, we could transcribe this element so that "peel, pail, pal" would be [pʰiʊɫ, pʰeʊɫ, pʰæʊɫ]. Note that I omitted the usual second element of the diphthong [eɪ] in order to show that in these circumstances the vowel moved from a mid-front to a mid-central rather than to a high-front quality.

Back vowels, as in "haul, pull, pool," are usually less affected by the final [ɫ] because they already have a tongue position similar to that of [ɫ]. But there is often a great difference in quality in the vowels in "hoe" and "hole." As we have seen, many speakers of British English have a fairly front vowel as the first element in the diphthong [əʊ]. This vowel becomes considerably retracted before / ɫ / at the end of the syllable. You can observe the change by comparing words such as "holy," where there is no syllable final [ɫ], and "wholly," where the first syllable is closed by [ɫ].

The exact form of the rule for specifying vowel allophones before [ɫ] will vary from speaker to speaker. But, so that we can include a rule in our set summarizing some of the main allophones of vowels in English, we may say the following:

(6)   Vowels are retracted before syllable final [ɫ].

Some speakers have a similar rule that applies to vowels before / r /, as in "hear, there," which might be [hiᵊr, ðeᵊr]. Note again how / l, r / may act together in rules, as we found when discussing consonants in rule (12a) in the preceding chapter.

Again, let me emphasize that these rules roughly specify only some of the major aspects of the pronunciation of English. They do not state everything

about English vowels that is rule governed, nor are they completely accurately formulated. There are problems, for example, in saying exactly what is meant by a word or a syllable, and it is possible to find both exceptions to these rules and additional generalizations that can be made.

EXERCISES

A Put your own vowels in this chart, using a set of words such as that given in Table 2.2. Listen to each vowel carefully and try to judge how it sounds relative to the other vowels. You will probably find it best to say each vowel as the middle vowel of a three-member series, with the vowels in the words above and below forming the first and last vowels in the series. In the case of the diphthongs, you should do this with both the beginning and the ending points.
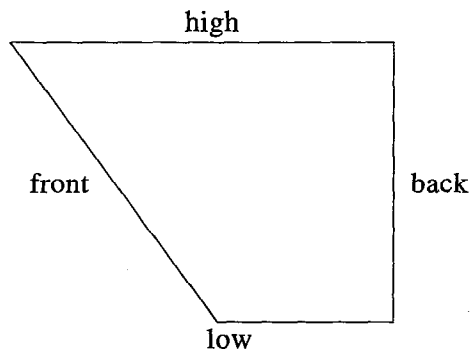
B Try to find a speaker with an accent different from your own (or perhaps a foreigner who speaks English with an accent) and repeat Exercise A, using this blank chart.
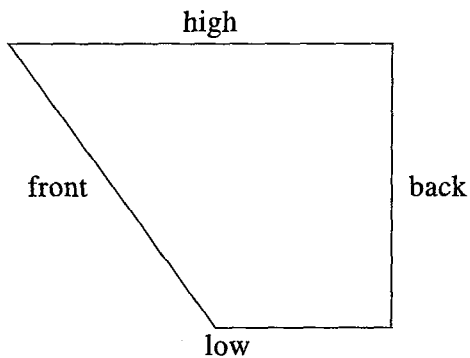
C List words illustrating the occurrence of vowels in syllables closed by / p /. Do not include names of foreign origin. You will find that some vowels cannot occur in these circumstances.

iː

ɪ

eɪ

ɛ

æ

ɑː

ɔː

oʊ

ʊ

uː

ʌ

aɪ

aʊ

ɔɪ

D Considering only the vowels that *cannot* occur in syllables closed by / p / as in Exercise C, give words, if possible, illustrating their occurrence in syllables closed by the following consonants.

/ b /

/ m /

/ f /

/ t /

/ n /

/ l /

/ s /

/ z /

/ k /

/ g /

E Which vowel occurs before the smallest number of consonants? Also, which class of consonants occurs after the largest number of vowels? (Define the class in terms of the place of articulation at which these consonants are made.)

_____

_____

_____

_____

F Look at Table 4.1. Find additional examples illustrating the relationship between the words in the first and third columns. Transcribe each pair of words as shown here for the vowel / iː /.

vowel	stressed syllable	reduced syllable
iː	secrete [sə'kriːt]	secretive ['siːkrətɪv]
ɪ		
eɪ		
ɛ		
æ		
ɑː or ɒ		
oʊ		
aɪ		

G Make up and transcribe a sentence containing at least eight different vowels.

_____

_____

H Give a number of examples for each of rules (1) through (6) by making a transcription of some additional words that fit the rules. Your examples should not include any words that have been transcribed in this book so far. Remember to mark the stress on words of more than one syllable.

Rule (1)   three examples (one for each syllable)

_____   _____   _____

Rule (2)   two pairs of examples (each showing words differing principally in stress)

_____   _____

_____   _____

Rule (3)   two sets of examples (each containing a one-syllable, a two-syllable and a three-syllable word, with the first, stressed syllable remaining constant)

———————   ———————   ———————

———————   ———————   ———————

Rule (4)   four examples

———————   ———————   ———————

———————

Rule (5)   four examples (use different vowels and different nasals)

———————   ———————

———————   ———————

Rule (6)   two sets of examples, each containing a contrasting pair of words

———————   ———————

———————   ———————

PERFORMANCE EXERCISES

A Learn to produce only the first part of the vowel [eɪ] as in "hay." Try saying this sound in place of your normal diphthong in words such as "they came late." Similarly, learn to produce a mid-high back vowel [o], and say it in words that you have been transcribing with the diphthong [oʊ], such as "Don't go home."

B Incorporate [e] and [o] in nonsense words for production and perception exercises. These words might also now include the voiceless sounds [m̥, n̥, ŋ̥, w̥, j̥]. Remember to practice saying the words by yourself, so that you can say them fluently to your partner. Start with easy words such as the following:

mɑˈŋ̥ɑ

n̥eme

ˈŋ̥ɑle

ˈmoʔi

ˈl̥ele

Then go on to more difficult words like these:

heˈm̥ɑn̥e

ˈŋambm̥bel̥

ˈspoʔetn̩ʔɔɪ

ˈwo̥θʃoˈɾesfi

ˈtlepɾidʒiˈkuʒ

C Again working with a partner, write the numbers 1 through 5 somewhere on a vowel chart as, for example, shown in the chart below. Now say vowels corresponding to these numbered positions in nonsense monosyllables, saying, for example, something like [dub]. Your partner should try to plot these vowels on a blank chart. When you have pronounced five words, compare notes, and then discuss the reasons for any discrepancies between the two charts. Then reverse roles and repeat the exercise.
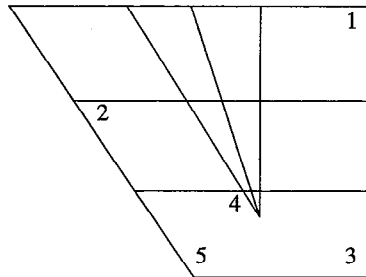
D Repeat Performance Exercise C with as many different partners as you can. It is difficult to make perceptual judgments of the differences among vowels, but you should be able to find a rough consensus.

E In addition to nonsense words of the kind given in Performance Exercise B, continue practicing with words to increase your auditory memory span. Say each word only two or three times. Remember that you should be spending at least one hour a week on production and perception exercises.

θeˈmifeˈðim̥e

ˈserapoˈsɑpofiˈpos

moˈpretepleteˈki

n̥ɑˈkotoˈtɑkpoto

lɑˈkimitiˈnoneʔe

5

ENGLISH WORDS AND SENTENCES

STRONG AND WEAK FORMS

The form in which a word is pronounced when it is considered in isolation is called its **citation form.** At least one syllable is fully stressed and has no reduction of the vowel quality. But in connected speech, many changes may take place. Some smaller words such as "and, to, him" may be considerably altered. They will usually be completely unstressed, the vowel may be reduced to [ə] or may disappear altogether, and one or more of the consonants may be dropped or altered. Thus "and" in its reduced form may be pronounced as [ənd] or [ən] or [n]. Try to pronounce it in these three different ways in a phrase such as "bread and butter."

Many words are like "and" in that they seldom maintain their citation form in conversational speech. These words may be said to have two different forms of pronunciation. There is a **strong form,** which occurs when the word is stressed, as in sentences such as "I want money *and* happiness, not money *or* happiness." There is also a **weak form,** which occurs when the word is in an unstressed position. Table 5.1 lists a number of common English words that have strong and weak forms.

Several of the words in Table 5.1 have more than one weak form. Sometimes, as in the case of "and," there are no clear rules as to when one as opposed to another of these forms is likely to occur. After a word ending with an alveolar consonant, I have a tendency to drop the vowel and say [n̩] or [n̩d] in phrases such as "cat and dog" or "his and hers." But this is far from invariable.

For some words, however, there are rules that are nearly always applicable. The alternation between "a" [ə] before a consonant and "an" [ən] before a vowel is even recognized in the spelling. Similar alternations occur with the words "the, to," which are [ðə, tə] before consonants and are often [ðiː, tuː] or [ðɪ, tʊ] before vowels. Listen to your own pronunciation of these words in the sentence, "The [ðə] man and the [ðɪ] old woman went to [tə] Britain and to [tʊ] America." The two examples of "the" will often be pronounced differently. It should be noted, however, that there is a growing tendency for younger American English speakers to use the form [ðə] in all circumstances, even before a

TABLE 5.1	Strong and weak forms of some common English words. More than five times as many could easily have been listed.		
Word	**Strong Form**	**Weak Form**	**Example of a Weak Form**
a	eɪ	ə	a cup [ə 'kʌp]
and	ænd	ənd, nd, ən, n̩	you and me ['juː ən 'miː]
as	æz	əz	as good as [əz 'gʊd əz]
at	æt	ət	at home [ət 'hoʊm]
can	kæn	kən, kŋ	I can go [aɪ kŋ 'goʊ]
has	hæz	həz, əz, z, s	he's left [hɪz 'left]
he	hi	i, hɪ, ɪ	will he go? [wɪl ɪ 'goʊ]
must	mʌst	məst, məs, m̩s	I must sell [aɪ m̩s 'sɛl]
she	ʃi	ʃɪ	did she go? ['dɪd ʃɪ 'goʊ]
that	ðæt	ðət	he said that it did [hɪ 'sɛd ðət ɪt 'dɪd]
to	tuː	tʊ, tə	to Mexico [tə 'meksɪkoʊ]
would	wʊd	wəd, əd, d	it would do ['ɪt əd 'duː]

vowel. If a glottal stop is inserted before words beginning with a vowel (another growing tendency in American English), then the form [ðə] is even more likely to be used.

Some of the words in Table 5.1 are confusing in that the same spelling represents two words with different meanings (two homonyms). Thus the spelling "that" represents a demonstrative pronoun in a phrase such as "that boy and the man," but it represents a relative pronoun in "he said that men were better." Only the relative pronoun has a weak form. The demonstrative "that" is always pronounced [ðæt]. Similarly, when "has" indicates past tense, it may be [z], as in "she's gone," but it is [həz] or [əz] when it indicates possession, as in "she has nice eyes."

There is another way in which words can be affected when they occur in connected speech. As you already know, sounds are often affected by adjacent sounds—for example, the [n] in "tenth" is articulated on the teeth because of the following dental fricative [θ]. Similar effects commonly occur across word boundaries, so that in phrases such as "in the" and "on the" the [n] is realized as a dental [n̪] because of the following [ð].

When one sound is changed into another because of the influence of a neighboring sound, there is said to be a process of **assimilation.** There is an assimilation of [n] to [n̪] because of the [ð] in the phrase "in the." Anticipatory coarticulation is by far the most common cause of assimilations in English. But perseverative assimilations do occur, for example, in the pronunciation of the phrase "it is" [ɪt ɪz] as "it's" [ɪts] as a result of the perseveration of the voicelessness of [t].

There is, of course, nothing slovenly or lazy about using weak forms and assimilations. Only people with artificial notions about what constitutes so-

called good speech could use adjectives such as these to label the kind of speech I have been describing. Weak forms and assimilations are common in the speech of every sort of speaker in both Britain and America. Foreigners who make insufficient use of them sound stilted.

STRESS

A stressed syllable is usually produced by pushing more air out of the lungs in one syllable relative to others. A stressed syllable thus has greater respiratory energy than neighboring unstressed syllables. It may also have an increase in laryngeal activity. Stress can always be defined in terms of something a speaker does in one part of an utterance relative to another.

It is difficult to define stress from a listener's point of view. A stressed syllable is often, but not always, louder than an unstressed syllable. It is usually, but not always, on a higher pitch. The most reliable thing for a listener to detect is that a stressed syllable frequently has a longer vowel than that same vowel would be if it were unstressed. But this does not mean that all long vowels are necessarily stressed. The second and third vowels in "radio," for example, are comparatively long, but they do not have the extra push of air from the lungs that occurs on the first vowel. Conversely, the vowels in the first syllables of "Russia" and "hit man" are comparatively short, but they have extra respiratory energy and so are felt to be stressed.

Stress can always be correlated with something a speaker does rather than with some particular acoustic attribute of the sounds, Consequently, you will find that the best way to decide whether a syllable is stressed is to try to tap out the beat as a word is said. This is because it is always easier to produce one increase in muscular activity—a tap—exactly in time with an existing increase in activity. When as listeners we perceive the stresses that other people are making, we are probably putting together all the cues available in a particular utterance in order to deduce the motor activity (the articulations) we would use to produce those same stresses. It seems as if listeners sometimes perceive an utterance by reference to their own motor activities. When we listen to speech, we may be considering, in some way, what we would have to do in order to make similar sounds.

Stress has several different functions in English. In the first place, it can be used simply to give special emphasis to a word or to contrast one word with another. As we have seen, even a word such as "and" can be given a contrastive stress. The contrast can be implicit rather than explicit. For example, if someone else says, or if I had even thought that someone else might possibly say

'John or 'Mary should 'go

I might, without any prior context actually spoken, say

'I think 'John '*and* 'Mary should 'go.

Another major function of stress in English is to indicate the syntactic relationships between words or parts of words. There are many noun–verb oppositions, such as "an 'insult, to in'sult; an 'overflow, to over'flow; an 'increase, to in'crease." In all these pairs of words, the noun has the stress on the first syllable and the verb has it on the last. The placement of the stress indicates the syntactic function of the word.

Similar oppositions occur in cases where two word phrases form compounds, such as "a 'walkout, to 'walk 'out; a 'put-on, to 'put 'on; a 'pushover, to 'push 'over." In these cases, there is a stress only on the first element of the compound for the nouns but on both elements for the verbs. Stress also has a syntactic function in distinguishing between a compound noun, such as "a 'hot dog" (a form of food), and an adjective followed by a noun, as in the phrase "a 'hot 'dog" (an overheated animal). Compound nouns have a single stress on the first element, and the adjective plus noun phrases have stresses on both elements.

Many other variations in stress can be associated with the grammatical structure of the words. Table 5.2 exemplifies the kind of alternations that can occur. All the words in the first column have the main stress on the first syllable. When the noun-forming suffix "-y" occurs, the stress in these words shifts to the second syllable. But as you can see in the third column, the adjectival suffix "-ic" moves the stress to the syllable immediately preceding it, which in these words is the third syllable. If you make a sufficiently complex set of rules, it is possible to predict the location of the stress in the majority of English words. There are very few examples of lexical items such as "differ" and "defer" that have the same syntactic function (they are both verbs) but different stress patterns. "Billow" and "below" is another pair of words that illustrates that differences in stress are not always differences between nouns and verbs.

DEGREES OF STRESS

In some longer words, it might seem as if there is more than one degree of stress. For example, say the word "multiplication" and try to tap on the stressed syllables. You will find that you can tap on the first and the fourth syllables of "multipli'cation." The fourth syllable seems to have a higher degree of stress.

TABLE 5.2	English word stress alternations.

'‿ _ _	‿ ' _ _	_ _ ' _
diplomat	diplomacy	diplomatic
photograph	photography	photographic
monotone	monotony	monotonic

The same is true of other long words such as "'magnifi'cation'" and "'psycholin'guistics." But this apparently higher degree of stress on the later syllable only occurs when the word is said in isolation or at the end of a phrase. Try saying a sentence such as "The 'psycholin'guistics 'course was 'fun." If you tap on each stressed syllable, you will find that there is no difference between the first and fourth syllables of "psycholinguistics." If you have a higher degree of stress on the fourth syllable in "psycholinguistics," this word will be given a special emphasis, as though you were contrasting some other psychology course with a psycholinguistics course. The same is true of the word "magnification" in a sentence such as "The de'gree of 'magnifi'cation de'pends on the 'power of the 'lens." The word "magnification" will not have a larger stress on the fourth syllable as long as you do not break the sentence into two parts and leave this word at the end of the first phrase.

Why does it seem as if there are two degrees of stress in a word when it occurs at the end of a phrase or when it is said alone—which is, of course, at the end of a phrase? The answer is that in these circumstances another factor is present. As we will see in the next section, the last stressed syllable in a phrase often accompanies a peak in the intonation. In longer words containing two stresses, the apparent difference in the levels of the first and the second stress is really due to the superimposition of an intonation pattern. When these words occur within a sentence in a position where there are no intonation effects, then there are no differences in the stress levels.

A lower level of stress may also seem to occur in some English words. Compare the words in the two columns in Table 5.3. The words in both columns have the stress on the first syllable. The words in the first column might seem to have a second, weaker, stress on the last syllable as well, but this is not so. The words in the first column differ from those in the second by having a full vowel in the final syllable. This vowel is always longer than the reduced vowel—usually [ə]—in the final syllable of the words in the second column. The result is that there is a difference in the rhythm of the two sets of words. This is due to a difference in the vowels that are present; it is not a difference in stress. There is not

TABLE 5.3	Three-syllable words exemplifying the difference between an unreduced vowel in the final syllable (first column) and a reduced vowel in the final syllable (second column).

'multiply	'multiple
'regulate	'regular
'copulate	'copula
'circulate	'circular
'criticize	'critical
'minimize	'minimal

a strong increase in respiratory activity on the last syllable of the words in the first column. Both sets of words have increases in respiratory activity only on the first syllable.

In summary, we can note that the syllables in an utterance vary in their degrees of prominence, but these variations are not all associated with what we want to call stress. A syllable may be especially prominent because it accompanies a peak in the intonation. We will say that syllables of this kind have a tonic stress. Given this, we can note that English syllables are either stressed or unstressed. If they are stressed, they may or may not be the tonic stress syllables that carry the major pitch change in the tone group. If they are unstressed, they may or may not have a reduced vowel. These relationships are shown in Figure 5.1.

As an aid to understanding the difference between these processes, consider the set of words "explain, explanation, exploit, exploitation." If each of these words is said in its citation form, as a separate tone group, the set will be pronounced as shown here (using a schematic representation of the intonation peak):

Intonation Peak	↑	↑	↑	↑
Stress	ex'plain	'expla'nation	ex'ploit	'exploi'tation
Segments	[ɪkspleɪn	ekspləneɪʃən	ɪksplɔit	eksplɔiteɪʃən]

Another way of representing some of these same facts is shown in Table 5.4. This table shows just the presence (+) or absence (–) of an intonation peak (a tonic accent), a stress, and a full vowel in each syllable in these four words. Considering first the stress (in the middle row), note that the two-syllable words are marked [+ stress] on the second syllable, and the four-syllable words are marked [+ stress] on both the first and third syllables.

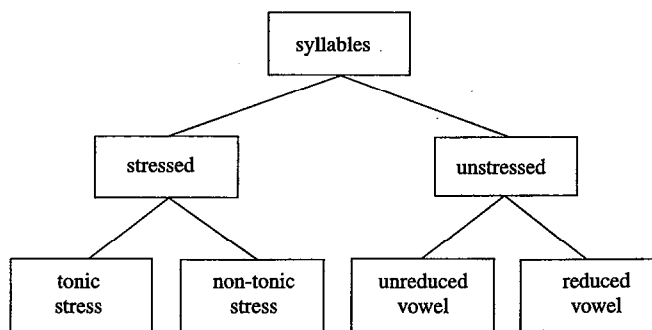**FIGURE 5.1**   Degrees of prominence of different syllables in a sentence.

TABLE 5.4	The combination of stress, intonation, and vowel reduction in a number of words.		

	explain	**explanation**	**exploit**	**exploitation**
tonic accent	− +	− − + −	− +	− − + −
stress	− +	+ − + −	− +	+ − + −
full vowel	− +	+ − + −	− +	+ + + −

As you can see by comparing the middle row with the top row, the last [+ stress] syllable in each word has been marked [+ tonic accent]. There is a [+] in the third row if the vowel is not reduced. Note that the difference between "explanation" and "exploitation" is that the second syllable of "explanation" has a reduced vowel but this syllable in "exploitation" has a full vowel. As we saw in the previous chapter, there are a number of vowels that do not occur in reduced syllables. Furthermore, the actual phonetic quality of the vowel in a reduced syllable varies considerably from dialect to dialect. I have transcribed the first vowel in "explain" as [ɪ] because that is the form I use. But other dialects may have [ə] or some other quality.

Some other books do not make the distinctions described here, maintaining instead that there are several levels of stress in English. The greatest degree of stress is called stress level one, the next is level two, the next level is three, a lower level still is level four, and so on. Note that in this system a smaller degree of stress has a larger number.

You can easily convert our system into a multilevel stress system by adding the number of [+] marks on a syllable in a table of the sort just used and subtracting this number from four. If there are three [+] marks, it is stress level one; if two, stress level two; if one, stress level three; and if none, it is stress level four. Try this for yourself with the data in Table 5.4. Writing the stress levels as superscripts after the vowels, you will find that "explanation" and "exploitation" are "e^{2}xpla^{4}na^{1}tio^{4}n" (a pattern of 2414) and "e^{2}xploi^{3}ta^{1}tio^{4}n" (a pattern of 2314).

I personally do not consider it useful to think of stress in terms of a multilevel system. Descriptions of this sort are not in accord with the phonological facts. But as it is so commonly said that there are many levels of stress in English, I thought I should explain how these terms are used. In this book, however, we will continue to regard stress as something that either does or does not occur on a syllable in English, and we will view vowel reduction and intonation as separate processes.

We can sometimes predict by rules whether a vowel will be reduced to [ə] or not. For example, we can formalize a rule stating that [ɔɪ] never reduces. But other cases are usually a matter of how recently the word came into common

use. Factors of this sort seem to be the reason why there are reduced vowels at the end of "postman," "bacon," and "gentleman," but not at the end of "mailman," "moron," and "superman."

SENTENCE STRESS

The stresses that can occur on words sometimes become modified when the words are part of sentences. The most frequent modification is the dropping of some of the stresses. There is a stress on the first syllable of each of the words "Mary, younger, brother, wanted, fifty, chocolate, peanuts" when these words are said in isolation. But there are normally fewer stresses when they occur in a sentence such as "Mary's younger brother wanted fifty chocolate peanuts." Tap with your finger at each stressed syllable while you say this phrase in a normal conversational style. You will probably find it quite natural to tap on the first syllables marked with a preceding stress mark in "'Mary's younger 'brother wanted 'fifty chocolate 'peanuts." Thus the first syllables of "younger," "wanted," and "chocolate" are pronounced without stresses (but with their full vowel qualities).

The same kind of phenomenon can be demonstrated with monosyllabic words. Say the sentence "The big brown bear ate ten white mice." It sounds unnatural if you put a stress on every word. Most people will say "The 'big brown 'bear ate 'ten white 'mice." As a general rule, English tries to avoid having stresses too close together. Very often, stresses on alternate words are dropped in sentences where they would otherwise come too near one another.

The tendency to avoid having stresses too close together may cause the stress on a polysyllabic word to be on one syllable in one sentence and on another in another sentence. Consider the word "clarinet" in "He had a 'clarinet 'solo" and in "He 'plays the clari'net." The stress is on the first or the third syllable, depending on the position of the other stresses in the sentence. Similar shifts occur in phrases such as "Vice-president 'Jones" versus "Jones, the vice-'president." Numbers such as "14, 15, 16" are stressed on the first syllable when counting, but not in phrases such as "She's 'only six'teen." Read all these phrases with the stresses as indicated and check that it is natural to tap on the stressed syllables. Then try tapping on the indicated syllables while you read the next paragraph.

'Stresses in 'English 'tend to re'cur at 'regular 'intervals of 'time. (') It's 'often 'perfectly 'possible to 'tap on the 'stresses in 'time with a 'metronome. (') The 'rhythm can 'even be 'said to de'termine the 'length of the 'pause between 'phrases. (') An 'extra 'tap can be 'put in the 'silence, (') as 'shown by the 'marks with'in the pa'rentheses. (')

Of course, not all sentences are as regular as those in the preceding paragraph. I said that stresses *tend* to recur at regular intervals. It would be quite untrue to say that there is always an equal interval between stresses in English. It is just that English has a number of processes that act together to maintain the rhythm.

I have already mentioned two of these processes. First, we saw that some words that might have been stressed are nevertheless often unstressed so as to prevent too many stresses coming together. Thus, to give another example, both "wanted" and "pretty" are stressed in "She 'wanted a 'pretty 'parrot," but they may not be in "My 'aunt wanted 'ten pretty 'parrots." Second, we saw that some words have variable stress; compare "the 'unknown 'man" with "the 'man is un'known."

We can also consider some of the facts mentioned in the previous chapter as part of this same tendency to reduce the variation in the interval between stresses. We saw that the vowel in "speed" is longer than that in "speedy," and this in turn is longer than that in "speedily." This can be interpreted as a tendency to minimize the variation in the length of words containing only a single stress, so that adjacent stresses remain much the same distance apart.
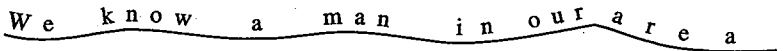
Taking all these facts together, along with others that will not be dealt with here, it is as if there were a conspiracy in English to maintain a regular rhythm. However, this conspiracy is not strong enough to completely override the irregularities caused by variations in the number and type of unstressed syllables. In a sentence such as "The 'red 'bird flew 'speedily 'home," the interval between the first and second stresses will be far less than that between the third and fourth. Stresses tend to recur at regular intervals. But the sound pattern of English does not make this an overriding necessity, adjusting the lengths of syllables so as to enforce complete regularity. The interval between stresses is affected by the number of syllables within the stress group, by the number and type of vowels and consonants within each syllable, and by other factors such as the variations in emphasis that are given to each word.
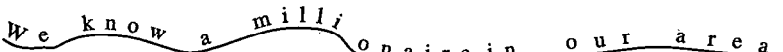
INTONATION

Listen to the pitch of the voice while someone says a sentence. You will find that it is changing continuously. The difference between speaking and singing is that in singing you hold a given note for a noticeable length of time and then jump to the pitch of the next note. But when one is speaking, there are no steady-state pitches. Throughout every syllable in a normal conversational utterance, the pitch is going up or down. (Try talking with steady-state pitches and notice how odd it sounds.)

The intonation of a sentence is the pattern of pitch changes that occurs. The part of a sentence over which a particular pattern extends is called an **intonational phrase.** A short sentence forming a single intonational phrase is shown in (1). The line above the sentence shows the pitch changes that occurred when this sentence was produced by a speaker of American English (see Sources at the end of the book for details). The positioning of the individual words above this line gives an indication of their relative timing. The sentence is shown below the line

in ordinary spelling, but with IPA stress marks added and one syllable preceded by an asterisk. Within the intonational phrase, each stressed syllable has a minor pitch increase; but there is usually a single syllable that stands out because it carries the major pitch change. A syllable of this kind is called the **tonic syllable** and will be marked in this section by an asterisk. In sentence (1), the first syllable of "area" is the tonic syllable, and, as you can see, it has the greatest pitch change. Each of the stress syllables is accompanied by a small increase in pitch, but the major pitch movement starts on the first syllable of the last word.
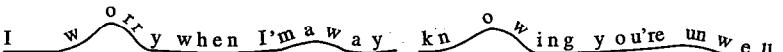
(1)

| We | 'know | a | 'man | in | our | *area |

The tonic accent usually occurs on the last stressed syllable in a tone group. But it may occur earlier, if some word requires emphasis.

(2)

| We | 'know | a | *millionaire | in | our | 'area |

The pitch changes that start on the tonic syllable may be continued on the following syllables. In the previous examples, the fall in pitch continues (but at a slower rate and with a small increase on the stressed syllable) until the end of the sentence.

Sometimes there are two or more intonational phrases within an utterance. When this happens the first one ends in a small rise, which we may call a continuation rise. It indicates that there is more to come, the speaker has not yet completed the utterance.
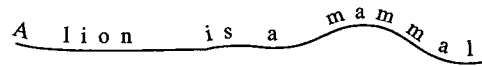
(3)

| I | *worry | when | I'm | a 'way, | ‖ *know | ing | you're | un 'well |

The beginning of a new intonational phrase may be marked, as in (3), by ‖. The pitch changes that begin on the tonic syllable in the first intonational phrase continue only until the beginning of the next phrase.
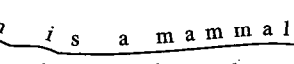
There is no syntactic unit exactly corresponding to an intonational phrase. When speaking slowly in a formal style, a speaker may choose to break a sentence up into a large number of phrases. The way in which a speaker breaks up a sentence depends largely on what that person considers to be the important information points in the sentence. An intonational phrase is a unit of information rather than a syntactically defined unit. It is only in rapid conversational style that there is likely to be one intonational phrase per sentence.

It is also usually impossible to predict which syllable will be the tonic syllable in an intonational phrase. Again, it depends on what the speaker considers to be important. In general, new information is more likely to receive a tonic accent

than material that has already been mentioned. The topic of a sentence is less likely to receive the tonic accent than the comment that is made on that topic. Thus, if I were telling someone about lions, I might say:

(4)

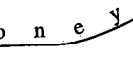A   'lion   is   a   * mammal

In this case, the topic of the sentence is a lion, and the comment on that topic is that it is a mammal. But if I were discussing mammals, and considering all the animals that fitted into that category, I might say:

(5)

A   *lion   is   a   'mammal

Various pitch changes are possible within the tonic accent. In sentences (1) through (5) the intonation may be simply described as having a falling contour, except for the continuation rise in the middle of (3). Another possibility is that the tonic syllable is the start of an upward glide of pitch. This kind of pitch change, which we will refer to as a rising contour, is typical in questions requiring the answer "yes" or "no," such as:

(6)

Will   you   'mail   me   my   *money?

As with falling contours, the syllable that starts the rising contour is not necessarily the last stressed syllable in an intonational phrase. It occurs earlier in:

(7)

Will   you   *mail   me   my   'money?

Now consider what you do in questions that cannot be answered "yes" or "no," such as:

(8)

'when   will   you   'mail   my   *money?

Of course, there are many possible ways of saying this sentence. But probably the most neutral is with a falling contour starting on the final stressed syllable. Questions that begin with *wh*–question words, such as "where, when, who, why, what," are usually pronounced with a falling intonation.

As we saw in (3), a small rising intonation occurs in the middle of sentences, a typical circumstance being at the end of a clause, as, to give another example, in:

(9) When you are *winning ‖ I will run a *way

'When you are *winning ‖ I will run a *way

A list of items also has a continuation rise:

(10) We knew 'Anna 'Lenny 'Mary and *Nora

We knew 'Anna 'Lenny 'Mary and *Nora

Note that yes–no questions can nearly always be reworded so that they fit into this pattern:

(11) Will you mail me my money, or not?

Will you mail me my money, or not?

It is useful to distinguish between two kinds of rising intonation. In the one, which typically occurs in yes–no questions, there is a large upward movement of pitch. In the other, the continuation rise that usually occurs in the middle of sentences, there is a smaller upward movement. These two intonations are often used contrastively. Thus, if there is a low rising intonation on an utterance, it means that there is something more to come. I might have a slightly rising intonation in:

(12) Yes   (13) Go on

These are the kinds of utterances one makes when listening to someone telling a story. They mean, "I hear you; please continue." If I have a larger rise in pitch and say:

(14) Yes   (15) Go on

it means, "Did you say 'Yes'?" or "Did you say 'Go on'?" It should be noted, however, that people are not entirely consistent in the way they use this difference in intonation.

Both rising and falling intonations can occur within the same tonic accent. If you tell me something that surprises me, I might have a distinct fall on the tonic syllable followed by a rise on the remainder of the intonational phrase:

(16) Your *mom will 'marry a 'law yer?

Your *mom will 'marry a 'law yer?

There are also distinct intonation patterns that one can use when addressing or calling someone. If I am answering a question such as "Who is that over there?" I will have a falling intonation over much of my pitch range, as shown in (17). However, if I am trying to attract someone's attention, I might use a falling intonation, but with only half the range of the full fall, as in (18). Calling to someone in this way can even be done as a chant, with comparatively steady pitches after the first rise, as in (19).

(17) Laura   **(18)** Laura   **(19)** Laura

We can sum up many differences in intonation by referring to the different ways in which a name can be said, particularly if the name is long enough to show the pitch curve reasonably fully. In curves (20) through (24), I have included dotted lines showing part of a pitch scale, so as to make comparisons easier. (It is actually an octave, 100 Hz to 200 Hz.) These curves show different pronunciations of the name "A'melia" (not written along the curve, as in the previous examples). (20) is a simple statement, equivalent to "Her name is Amelia." (21) is the question, equivalent to "Did you say Amelia?" (22) is the form with the continuation rise, which might be used when addressing Amelia, indicating that it is her turn to speak. (23) is a question expressing surprise, equivalent to "Was it really Amelia who did that?" Lastly, (24) is the form for a strong reaction, reprimanding Amelia.

(20)   **(21)**   **(22)**   **(23)**   **(24)**
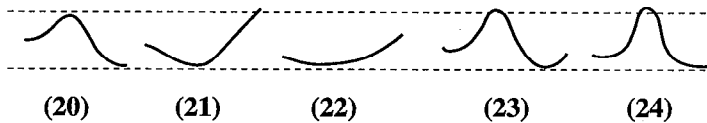
TARGET TONES

We have been considering intonation in terms of tunes that apply over whole sentences or phrases, but there are a number of other ways in which intonation can be described. Instead of considering the shape of the curve over a whole phrase, we could describe the intonation in terms of a sequence of high (H) and low (L) target pitches. When people talk they aim to make either a high or low pitch on a stressed syllable and to move upward or downward as they go into or come away from this target. One system for representing pitch changes of this kind is known as **ToBI,** standing for Tone and Break Indices. In this system, target tones H* and L* (called H star and L star) are typically written on a line (called a tier) above the segmental symbols, and put immediately above the stressed syllables. A high tone, H*, can be preceded by a closely attached low pitch, written L + H*, so that the listener hears a sharply rising pitch. Similarly

L* can be followed by a closely attached high pitch, L* + H, so that the listener hears a scoop upward in pitch after the low pitch at the beginning of the stressed syllable. Sometimes a stressed syllable can be high but nevertheless contain a small step down of the pitch. This is known as High plus downstepped High, and written H + !H*, with the exclamation mark indicating the small downstep in pitch. In special circumstances, to be discussed at the end of this section, a downstepped high syllable, !H*, can itself be a pitch accent. There are therefore six possibilities, shown in Table 5.5, that can be regarded as the possible pitch accents that occur in English.

The last pitch accent in a phrase is called the nuclear pitch accent. The ToBI system allows the phrase to be marked by an additional tone after the nuclear pitch accent. This tone, called the phrase accent, is written H– (H minus) or L– (L minus). Finally there is a boundary tone, which is marked H% or L%, depending on whether the phrase ends on a rising or falling pitch. In this framework all English intonations consist of a sequence of tones formed as shown in Table 5.5. As you can see in the first column, there are a number of pitch accents that may occur on stressed syllables before the nuclear pitch accent. The second colunm shows the nuclear pitch accents, one of which must always be present in a phrase. The part of the intonational phrase after the cnuclear pitch accent, the phrase accent, must be high or low; and there must be a high or low boundary tone.

The ToBI system also allows us to transcribe the strength of the boundary between words by means of a number. This is called a Break Index giving the name ToBI (To and Break Indices) to this transcription system. If there is no

| TABLE 5.5 | The ToBI system for characterizing English intonations. Each intonational phrase (tone group) must have one item from each of the last three columns, and it may also have additional pitch accents marked on other stressed syllables, as shown in the first column. The parenthesized accent, (!H*), will be explained at the end of this section. |

Optional Pre-Nuclear Pitch Accents on Stressed Syllables	Nuclear Pitch Accent	Phrase Accent	Boundary Tone
	H*		
H*			
L*	L*		
L + H*	L + H*	H–	H%
L* + H	L* + H		
H + !H*	H + !H*	L–	L%
(!H*)	(!H*)		

break, as, for example, in "you're" (which is usually identical with "your"), the Break Index can be marked as 0. This is a useful way of showing that a phrase such as "to Mexico" is usually pronounced as if it were a single word—there's no added break in "to Mexico" as compared with "tomorrow." Intervals between words are classified as having Break Index 1 (although, to my mind, there is usually nothing that can be called a break between words). Higher levels of break indices show, roughly speaking, greater pauses. A Break Index of 3 is usual between clauses that form intermediate intonational phrases, and a Break Index of 4 occurs between larger intonational phrases, such as whole sentences.

The last five intonation curves we considered would be transcribed in a ToBI transcription as follows (without indicating the Break Indices, which would always be a 4 at the end of each of these utterances):

(20) A'melia.   Tone tier   [H*   L–L%]
Simple statement in response to   Segmental tier   [ə m iː l iː ə]
"What is her name?"

(21) A'melia?   Tone tier   [L*   H–H%]
A question, equivalent to   Segmental tier   [ə m iː l iː ə]
"Did you say Amelia?"

(22) A'melia—   Tone tier   [L*   L–H%]
Addressing Amelia, indicating   Segmental tier   [ə m iː l iː ə]
that it is her turn to speak.

(23) A'melia!?   Tone tier   [L + H* L–H%]
A question indicating surprise.   Segmental tier   [ə m iː l iː ə]

(24) A'melia!!   Tone tier   [L + H* L–L%]
A strong reaction, reprimanding Amelia.   Segmental tier   [ə m iː l iː ə]

The ToBI transcription for (20), [H* L–L%], is typical of a simple statement with only one stressed syllable receiving a pitch accent. The part of the phrase after the nuclear accent is low, and the phrase ends with a low boundary tone. Similarly, the transcription for (21), [L* H– H%], is a typical tune for a question that can be answered by "yes" or "no," which ends with a fairly large pitch rise. At the end of the next phrase, (22), there is a smaller rise of the kind that occurs in an unfinished utterance, or in a list of words such as we exemplified in (10), ". . . Anna, Lenny, Mary and . . . ." The way in which ToBI separates the large pitch rise in (21) —the question rise—from the smaller rise in (22)—the continuation rise—is by making the phrase tone L–, so that (22) has the tune [L* L– H%] instead of [L* H– H%] as in (21). The low phrase tone prevents the final high boundary from being so high. The stressed syllables of the final two tunes begin with an L, ensuring that the H* indicates a sharp rise from a low pitch. Thus in (23) we have [L + H* L– H%], with a low phrase tone and a

small pitch rise at the end, much as in (22). Finally (24), [L + H* L– L%], is like (23) in beginning with a strong rise, but it ends with a low boundary tone.

The simple statements, questions, and other intonations that we discussed earlier can be transcribed in a similar way.

(1) We know a 'man in our 'area. Tone tier   [H*   H*L–L%]
Simple statement.   Segmental tier [wiː nou ə mæn m aːr eriə   ]

(6) Will you 'mail me my 'money? Tone tier   [H*   L*H–H%]
Simple yes–no question.   Segmental tier [wɪl juː meɪl miː maɪ mʌni   ]

(9) When you are 'winning, 'I will run a'way.
Two clauses with a Break Index 4. Break Index   [1 1 1   4 1 1   1   4]
                                   Tone tier   [H*L–H%H*   H*LL%]
                                   Segmental tier [wen juː aː wɪnɪŋ   aɪ wɪl rʌn əweɪ   ]

The Break Indices are shown in (9). At the end of the first intonational phrase there is a continuation rise represented by L– H%, and a break index 4. All the words are closely joined together, so in each case the Break Index is 1.

Finally we must consider how to transcribe another fact about English intonation (which also applies to many other languages). The pitch in most sentences has a tendency to drift down. Earlier, when discussing stress, we considered the sentence "Mary's younger 'brother wanted 'fifty chocolate 'peanuts," with stresses on alternate words, "Mary's," " brother," "fifty," and "peanuts." If you say this sentence with these stresses, you will find that there is an H* pitch accent on each of the stressed syllables, but each of these high pitches is usually a little lower than the preceding high pitch. This phenomenon is known as **downdrift.** We can represent this in the transcription by marking the H* pitch accents as being downstepped, a notion that was mentioned earlier in connection with a fall from a high pitch within a syllable. In Table 5.5 we used a small raised exclamation mark, !. We can indicate that each of the H* tones is a little lower than the preceding one by transcribing them as **downstepped** highs, ! H*, in the tone tier for this sentence:

[H*   !H*   !H*   !H* L–L%]
(25) Mary's younger brother wanted fifty chocolate peanuts.

Note that successive H* pitch accents do not have to be downstepped. If I had wanted to put a very slight emphasis on "brother," indicating that it was Mary's younger brother, not her younger sister, who had this peculiar desire, then I could have made the downstepping begin at "fifty" and said:

[H*   H*   !H*   !H* L–L%]
(26) Mary's younger brother wanted fifty chocolate peanuts.

The ToBI system is a way of characterizing English intonation in terms of a limited set of symbols—a set of six possible pitch accents including a downstep mark, two possible phrase accents, two possible boundary tones, and four possible Break Indices, going from 1 (close connection) to 4 (a boundary between intonation phrases). It was designed specifically for English intonations, but, with a few modifications, it may be appropriate for other languages as well.

EXERCISES

A List the strong and weak forms of ten words not mentioned in this chapter. For each word, transcribe a short utterance illustrating the weak form (as in Table 5.1).

Word	Strong Form	Weak Form	Example of Weak Form
————	————	————	————
————	————	————	————
————	————	————	————
————	————	————	————
————	————	————	————
————	————	————	————
————	————	————	————
————	————	————	————
————	————	————	————
————	————	————	————

B Give two new examples of each of the following kinds of assimilations, one of the examples involving a change within a word, the other involving a change across word boundaries. (Even if you yourself do not say assimilations of the kind illustrated, make up plausible examples. I myself have heard all the examples given.)

A change from an alveolar to a bilabial consonant.

input	[ɪmpʊt]	*Saint Paul's*	[sm̩ˈpɔːlz]
————	————	————	————

A change from an alveolar consonant to a dental consonant.

tenth	[ten̪θ]	*In this*	[ɪn̪ ðɪs]
————	————	————	————

A change from an alveolar consonant to a velar consonant.

synchronous   ['sɪŋkrənəs]   *within groups*   [wɪð'ɪŋ gruːps]

_____ _____ _____ _____

A change from a voiceless consonant to a voiced consonant.

catty   ['kædi]   *sit up*   [sɪˈdʌp]

_____ _____ _____ _____

C Give five more examples of assimilation. Choose examples as different as possible from any that have been given before.

_____   [_____]

_____   [_____]

_____   [_____]

_____   [_____]

_____   [_____]

D Make up pairs of phrases or sentences that show how each of the following words can have two different stress patterns.

Example: continental
It's a 'continental 'breakfast.
She's 'very conti'nental.

afternoon

_____

_____

artificial

_____

_____

diplomatic

_____

_____

absentminded

_____

_____

New York

_____

_____

E Fill in plus and minus signs so as to indicate which syllables in the following table have tonic accents, which have stress, and which have full vowels. You may find it useful to refer back to Table 5.4.

	computation	compute	inclination	incline (verb)
tonic accent				
stress				
full vowel				

F About 100 years ago, the following words had stress as shown. Some of them still do for some people. But many of them (in my speech, all of them) are stressed differently nowadays. Transcribe these words and show the stress on each of them in your own speech. Then state a general rule describing this tendency for the position of the stress to change to a particular syllable.

an'chovy _____

ab'domen _____

'applicable _____

'controversy _____

'nomenclature _____

tra'chea _____

eti'quette _____

re'plica _____

va'gary _____

blas'phemous _____

a'cumen _____

Rule: _____

_____

G List three more sets of words showing the stress alternations of the kind shown in Table 5.2.

'photograph	pho'tography	photo'graphic
_____	_____	_____
_____	_____	_____
_____	_____	_____

H Indicate the stress and intonation patterns that might occur in the situations described for the following utterances. Draw curves indicative of the pitch rather than using ToBI symbols.

1. "Can you pass me that book?" (said politely to a friend)

2. "Where were you last night?" (angry father to daughter)

3. "Must it be typewritten?" (polite question)

4. "Who is the one in the corner?" (excitedly, to a friend)

5. "He's very nice . . ." (but I don't like him)

I Make a segmental transcription and also show the Tone Tier with a ToBI transcription of the following utterances for which the pitch curves have been drawn in this chapter.

(2) We know a millionaire in our area.

(4) A lion is a mammal.

(5) A lion is a mammal.

(7) Will you mail me my money?

(8) When will you mail me my money?

(10) We knew Anna, Mary, Lenny, and Nora.

PERFORMANCE EXERCISES

A Pronounce the following phrases exactly as they have been transcribed, with all the assimilations and elisions. (Each of these transcriptions is a record of an utterance that I have actually heard in normal conversations between educated speakers.)

"What are you doing?"	['wɑdʒə'dum]
"I can inquire."	['aɪkŋ ŋ 'kwaɪə]
"Did you eat yet?"	['dʒitjɛʔ]
"I don't believe him."	[aɪ'doʊmbə'liːvɪm]
"We ought to have come."	[wiˈɔːtf'kʌm]

B Working with a partner, try to transcribe the intonation of a few sentences. You may find it difficult to repeat a sentence over and over again with the same intonation. If you do, try to work from a tape recording. In any case, write down the sentence and the intonation you intend to produce. Practice saying it in this way before you say it to your partner.

C Take turns saying nonsense words such as those shown below, transcribing them, and comparing transcriptions.

ʃkeɪʒdʒ'minʒe

ʔɑŋkliθuntθ

sfe'eʔemɑ

grɔɪpst'braɪgz

D Also make up lists of words for improving your memory span. These words are more difficult if the stress is varied and if the sounds are mainly of the same class (stops, front vowels, voiceless fricatives, etc.).

tipe'kiketi'pe

θɔɪ'saɪθauˈfɔɪʃauθau

'monɑŋuˈŋonɔmɑ

woʔɔɪlauraˈrɔlojɔ

bəbdɪg'bedgɪbdɛd'bɛbdəd