Chapter Two Phonetics

As noted in Chapter One, the core areas of linguistics include phonetics, phonology, morphology and syntax. This chapter is about phonetics, which deals the inventory and physical structure of speech sounds.¹

If we were to survey the sounds employed in human languages, we would discover a great many different sounds, many of which are not used in any individual language. Thus, for example, while American English displays a good number of speech sounds, this number represents only a subset of the sounds that represent the full range of human-language speech sounds. Interestingly, however, the full set of sounds employed across all human languages does not represent the full range of sounds that humans are capable of producing.²

1. Vocal Organs

So how are these sounds produced? The central organs involved in the production of human-language speech sounds are shown in the figures on the following pages. These organs include the *vocal tract* (the oral cavity, the nasal cavity, and the throat or *pharynx*) along with the lungs and the larynx.

¹Following tradition, one can further differentiate between *acoustic phonetics* and *articulatory phonetics*, where the former refers to the physical properties of speech sounds in terms of how we hear them and the latter refers to the physical properties of speech sounds in terms of how we produce them. Even though the study of acoustic phonetics can reveal many interesting properties about speech sounds, we will, in this course, concentrate on articulatory phonetics. For more information about acoustic phonetics, see such introductory texts as William O'Grady, Michael Dobrovolsky and Mark Aronoff's *Contemporary Linguistics* (New York: St. Martin's Press, 1993, 2nd edition).

²Another interesting fact about speech sounds comes from language acquisition. In the very earliest stage of native-language development, known as the *babbling stage*, children produce sounds that may or may not be represented in the language to which they are exposed. What makes the set of sounds that children produce during this stage interesting, though, is that this set appears to fall within the range of human-language speech sounds. That is, what the child produces during speech-like production at this stage apparently includes only sounds that exist in human languages. Chapter Five provides more information on language acquisition.

1.1 Lungs

Starting from the lowest organ, we can identify the lungs, from which an airstream is forced upwards. The central muscles involved in forcing this airstream upwards are the *diaphram*, a sheet of powerful muscles that separate the chest cavity from the abdominal organs, and the *intercostals*, muscles that surround the lungs and are located between the ribs. With these muscles, we are able to force an airstream upwards, out of the lungs.

Note, though, that just forcing air out of the lungs does not necessarily result in speech sounds. After all, if we want to blow out candles on a birthday cake, we have to perform exactly the same function: force air out of the lungs.

1.2 Larynx

One of the more important speech organs, then, is located above the lungs, specifically in the throat region. This organ is known as the *larynx*, more commonly known as the *voice box* or the *Adam's apple*. The larynx is located at a strategic junction in the throat, at exactly that point where the throat divides between the *trachea* (your *wind pipe*), which leads to the lungs, and the *esophagus* (the tube that carries food or drink to the stomach). Covering the larynx is a flap that acts to close off the trachea when we swallow, thereby preventing us from 'swallowing' food into our lungs. This flap is known as the *epiglottis*.

When the epiglottis is folded back out of the way, though, we observe the 'real' speech-making parts of the larynx. These are two sheets of tissue called *vocal folds* that stretch in a V-shaped fashion from the front to the back of the larynx. When these folds are close together and air from the lungs is forced past them, they vibrate in a way not entirely unlike the way that a comb harp vibrates. This vibration produces an audible noise, a speech sound.

More particularly, though, we are able to produce several different speech sounds with the vocal folds in the larynx. Two of these speech sounds will be of particular importance: The *voiced* sounds and the *voiceless* sounds. Voiced sounds are those that are produced when the vocal folds are close together and vibrating (i.e., producing audible noise); voiceless sounds are those produced when the vocal folds are far apart and not producing the vibrating noise.

Voiced and voiceless sounds

Because we will be referring to the difference between voiced and voiceless sounds many times in the coming weeks, it is important that you understand this difference. To get this difference, I want you to produce two sounds. First, produce the 'SSSSSS' sound that you'd make when you imitate the sound of air leaking out of a soon-to-be flat tire. Then produce the 'ZZZZZZZ' sound that you'd make to imitate the sound of a flying bee. Go back and forth between the two for a second or two. Now do something important: Put your fingers on your neck right over your 'voice box' and produce the 'SSSSSS' sound. Then, keeping your fingers over your 'voice box,' produce the 'ZZZZZZZ' sound. You'll feel the difference immediately: When you produce the 'ZZZZZZZ' sound, you'll feel a vibration; when you produce the 'SSSSSSS' sound, the vibration will stop.

What you're feeling when you do the 'SSS' versus 'ZZZ' exercise is the difference between voiceless ('SSS') and voiced ('ZZZ') sounds. The voiceless one is where the vocal folds are apart and not vibrating, and the voiced one is where the vocal folds are close together and making sound through the vibration. You can repeat this exercise with the following sounds:

Voiceless	vs.	Voiced
f-sound		v-sound
*p-sound		*b-sound
*t-sound		*d-sound
*k-sound		*g-sound

The asterisk "*" before some of the sounds above means that they might be a little tricky for you. When you produce the p-sound, for example, produce not something like *pee*, but just the 'p' itself. That'll be the voiceless sound. And when you produce the b-sound in the exercise, produce not something like *bee*, but only the 'b' itself. That'll be the voiced sound.

So far, we've seen (and felt) that we can produce voiced and voiceless sounds with the vocal folds in the larynx. Two other types of sounds are important, too, though we'll be refering to these two somewhat less often than to the usual voiced and voiceless sounds. One of these other two types is known as a *whispered* sound, for the obvious reason that we produce this type of sound when we whisper. This type of sound is produced when the vocal folds are closed off in the front (anterior); only the very back part (anterior) is open. In fact, whispered sounds are voiceless, a fact that can be ascertained by trying, very silently, to whisper the buZZZing-bee sound (which turns out to be impossible). The other type of sound is known as a *murmur*, also sometimes referred to as a *whispery voice*. Murmured sounds are, in fact, voiced sounds, but the vocal folds are somewhat relaxed, allowing them to vibrate rather more loosely and to allow more air to past them.

We've thus seen that we can produce a variety of different sounds by varying the positions of the vocal folds. These varying positions of the vocal folds are known, together, as *glottal states*. The *glottis* itself is the technical name for the open space between the vocal folds.

1.3 Oral Cavity

The cavity oral extends from the lips back to the *pharynx* (the beginning of the throat). It is in this area of that we are able to make large numbers of discriminations in speech sounds. The most important regions are the following:

- 1. Labia (lips)
- 2. Teeth (*dental* sounds)
- 3. Roof of mouth
 - a. *alveolar ridge* (rough bony region just behind the upper teeth)
 - b. *palate* (bony roof of mouth, sometimes called the *hard palate*)
 - c. *velum* (soft roof in the rear, sometimes called the *soft palate*)

4. tongue

- a. *apex* or tip
- b. *blade* (upper tongue just behind the tip)
- c. *body* (rear portion)

1.4 Nasal cavity

The final area of importance is the nasal cavity, with which we produce nasal sounds like the n-sound or the m-sound.

2. IPA and American English speech sounds

After inventorying the crucial speech organs, one would think that one could proceed directly to the sounds themselves. Alas, we need to introduce one last twist, namely, some means of representing sounds on paper (or blackboards). The system we will use to represent human-language speech sounds is known as the *International Phonetic Alphabet*, which is a system of symbols such that each symbol uniquely represents one sound, that is, that there is a one-to-one relationship between individual sounds and individual symbols. Some of the IPA symbols may be familiar from American English; others will be completely new and will have to be memorized.

Don't be dismayed!

Some students are dismayed when they learn that they must learn a new alphabet. It is important to realize, then, why it is that such an alphabet is necessary. The basic problem is that the English (French, German, etc.) alphabetic system is not unique enough to be able to represent sounds in a one-to-one way. For example, the English alphabetic symbol 'a' may be pronounced in different ways in different words (compare, for example, the pronunciation of *a* in *tap* and *tape*). In fact, what we find in English is (1) that the same sound may be represented by several different English alphabetic symbols; (2) that the same symbol may represent several different sounds; and (3) that some sounds are represented by combinations of symbols.

We are now in a position to begin inventorying the sounds of American English and describing their physical structure in terms of the vocal tract. In fact, we divide sounds into the traditional categories *consonants* and *vowels*. However, there is a group of sounds that fall somewhere between the consonants and vowels. These are knows as the semi-vowels, which we will include among the consonants.

2.1 Consonants

To describe the consonants, we divide the sounds into several categories and then, for the

sounds in each category, employ locational and voicedness information to subdivide even further. The major categories for the consonants, known as the *manners of articulation*, are the *stops*, the *fricatives*, the *affricates*, the *liquids*, the *nasals*, and the *glides* (semi-vowels).

2.1.1 Manner of articulation: Stop

Stops are formed when the air stream is completely closed off by, for example, closure of the lips. When we produce a p-sound, for example, we note that the lips are completely closed, albeit for only a very short time. In other words, the air steam is completely interrupted when we produce a p-sound -- or any other stop consonant.

Among the stop consonants, we can differentiate between those that are voiced and those that are voiceless. The voiced stops include [b], [d], and [g] while the voiceless stops include [p], [t], and [k].

Note in passing that we have put the symbols in square brackets, like this: [p]. The brackets mean that we have done phonetic representations with the IPA symbols. The significance of this type of representation will become more clear as we proceed; for now, just remember to use the square brackets yourself when you represent sounds.

Not all of the stops are articulated in the same place in the oral cavity, however. Compare [t] and [d]. They differ on voicing, as we already noted above, but for both sounds, the tongue is in the same place vis-a-vis the alveolar ridge. Now compare [t] and [k]. Both are voiceless, as we pointed out above. However, [t] is articulated at the alveolar ridge while [k] is articulated at the velum (soft palate).

We thus observe that consonants can differ on their *glottal state* (voiced vs. voiceless) and that they can differ on their *place of articulation* (e.g., alveolar vs. velar). We can put this information all together as in the chart below. (The notation "NA" under the IPA Symbol column means that there are no such sounds in American English.)

Manner of articulation: Stop		
Place of articulation	Glottal state	IPA symbol
bilabial	voiceless	[p]
	voiced	[b]
labiodental	voiceless	NA
	voiced	NA
interdental	voiceless	NA
	voiced	NA
alveolar	voiceless	[t]
	voiced	[d]
palatal	voiceless	NA
	voiced	NA
velar	voiceless	[k]
	voiced	[g]
glottal	voiceless	[?]
	voiced	NA

Again, most of the IPA symbols are familiar from American English--although the IPA will *always* be pronounced in exactly the same way, no matter where they occur. In addition, we also see a symbol that is not familiar from American English, namely, the symbol [?]. This symbol stands for the *glottal stop*, which is produced as a stop immediately at the beginning of *Oh-oh!* (e.g., when something is spilled).

Before going on, note that we can refer to particular sounds with a three-part descriptive statement. For example, we can refer to [p] as a voiceless bilabial stop.

2.1.2 Manner of articulation: Fricative

The fricative sounds have a hissing noise associated with them, as, for example, we hear when we produce the s-sound. As with the stops above, we divide the fricatives according to the state of the glottis and the place of articulation.

Manner of articulation: Fricative		
Place of articulation	Glottal state	IPA symbol
bilabial	voiceless	NA
	voiced	NA
labiodental	voiceless	[f]
	voiced	[V]
interdental	voiceless	[θ]
	voiced	[ð]
alveolar	voiceless	[s]
	voiced	[z]
palatal	voiceless	[š] / [∫]
	voiced	[ž] / [ʒ]
velar	voiceless	NA
	voiced	NA
glottal	voiceless	[h]
	voiced	NA

Again, we observe some symbols that may not be familiar. To get a feel for the sounds that these symbols represent, consider the following examples:

Voiceless interdental fricative	[θ]	thing
Voiced interdental fricative	[ð]	that
Voiceless palatal fricative	[š]	ship
Voiced palatal fricative	[ž]	azure

Above you note that we've listed two different symbols for the palatal fricatives. Older varieties of the IPA use the symbols [\check{s}] and [\check{z}] while more recent varieties use the symbols [\check{s}] and [\check{z}]. Because Cowan & Rakušan (*Source Book for Linguistics*) use the older symbols, we will assume them here. If you have learned the more recent IPA symbols in some other courses, you are also free to use them here as well.

2.1.3 Manner of articulation: Affricate

The affricate sounds are often informally described as a stop+fricative combination, as in the initial sound in the word *chair*. More specifically, these sounds are characterized by the slowed release of their (initial) closure. American English has only two affricates, as shown below.

Manner of articulation: Affricate		
Place of articulation Glottal state IPA symbol		
bilabial	voiceless	NA
	voiced	NA
labiodental	voiceless	NA
	voiced	NA
interdental	voiceless	NA
	voiced	NA
alveolar	voiceless	NA
	voiced	NA
palatal	voiceless	[č]/[t∫]
	voiced	[j]/[dʒ]
velar	voiceless	NA
	voiced	NA
glottal	voiceless	NA
	voiced	NA

To memorize what these two symbols represent, consider the following examples:

Voiceless palatal affricate	[č]	charm
Voiced palatal fricative	[j]	jar

Note, once again, that we have provided two different symbols for these sounds; the symbols [č] and [ž] are the somewhat more traditional ones, as used in Cowan & Rakušan.

2.1.4 Manner of articulation: Nasal

The nasal sounds are produced with the oral cavity closed off, the airstream passing through the nasal cavity instead. American English has three such sounds:

Manner of articulation: Nasal			
Place of articulation	Glottal state	IPA symbol	
bilabial	voiceless	NA	
	voiced	[m]	
labiodental	voiceless	NA	
	voiced	NA	
interdental	voiceless	NA	
	voiced	NA	
alveolar	voiceless	NA	
	voiced	[n]	
palatal	voiceless	NA	
	voiced	NA	
velar	voiceless	NA	
	voiced	[ŋ]	
glottal	voiceless	NA	
-	voiced	NA	

The sounds associated with [n] and [m] are those that we usually associate with the *n* and the *m* in American English. The third sound, represented by [ŋ], is what we make in the *-ing* sound as in *blasting*.

2.1.5 Manner of articulation: Liquid

The liquids are produced when the airstream must pass either between or around the tongue. There are two liquids in American English, the [1] and the [r]. Note, though, that while both sounds are voiced, and both are alveolar, the tongue isn't in the same position when we pronounce them. We call these two tongue positions *lateral* and *retroflex*, as shown below:

Manner of articulation: Liquid			
Place of articulation	Glottal state	IPA symbol	
bilabial	voiceless	NA	
	voiced	NA	
labiodental	voiceless	NA	
	voiced	NA	
interdental	voiceless	NA	
	voiced	NA	
alveolar	voiceless	NA	
lateral	voiced	[1]	
retroflex	voiced	[r]	
palatal	voiceless	NA	
	voiced	NA	
velar	voiceless	NA	
	voiced	NA	
glottal	voiceless	NA	
	voiced	NA	

2.1.6 Manner of articulation: Glide

The glides are the semi-vowels in that they have both vowel-like and consonant-like qualities. The American English glides are shown below:

Manner of articulation: Glide	e	
Place of articulation	Glottal state	IPA symbol
bilabial	voiceless	[M]
	voiced	[w]
interdental	voiceless	NA
	voiced	NA
alveolar	voiceless	NA
	voiced	NA
palatal	voiceless	NA
	voiced	[y] / [j]
velar	voiceless	([M])
	voiced	([w])
glottal	voiceless	NA
	voiced	NA

Several things are of interest here. First, the sounds themselves. In fact, the two symbols [w] and [M], may not even be present in all dialects of American English. If you can hear the difference between the initial sounds of *which* and *witch*, then you have this difference in your dialect. Another good example is from the pun on the well-known bumper-sticker: *Visualize World Peace* versus *Visualize Whirled Peas*. Note the pronunications of *world* and *whirled*. If there is a difference, then your dialect of American English has these two sounds.

The second interesting fact about [w] and [M] is that we've listed them in two different places, that is, both as bilabial sounds and as velar sounds.

Finally, note that we've provided two symbols for the palatal glide. Older varieties of the IPA use the symbol [y] (as in Cowan & Rakušan) while more recent incarnations of the IPA use [j].

2.2 Diacritics and narrow transcription

In fact, our inventory of American English consonant sounds is not really complete, because we may observe in English several speech sounds that are not obvious from the preceding presentation. To represent these, we employ a series of special marks added to the IPA symbols. These marks are known as *diacritics*.

2.2.1 Aspiration. Compare the following uses of the sounds [p], [t], and [k]:

Aspirated		Unaspirated
pat	versus	spat
tub	versus	stub
cope	versus	scope

If we compare these uses carefully, we discover that the ones in the Aspirated column have a small puff of air immediately following the sound itself. (A good way to test this is by holding your hand in front of your mouth and then saying these sounds outloud.) We indicate the *aspiration* on these uses of the consonants with a diacritic, in this case, a raised "h":

Aspirated		Unaspirated		
pat	$[p^h]$	versus	spat	[p]
tub	[t ^h]	versus	stub	[t]
cope	$[k^h]$	versus	scope	[k]

2.2.2 Syllabic liquids and nasals. In English, syllables include a vowel as the central element. Sometimes, however, we find that the vowel disappears. This is common in the world's languages;³ in English, we find that it happens only with liquids like [r] and nasals like [n]. We use the *tick mark* to indicate this. Following are some examples of words that have *syllabic consonants*, either liquids or nasals:

³ And when we talk about the world's languages, we're not talking about just a few hundred of them or so. Estimates on the number of different languages range anywhere from 3500 to 5000, some 20 percent of which are spoken by only a few million people in the South Pacific. Alas, it is also the case that languages are currently being lost forever at an ever accelerating rate.

bird	no vowel sound; [r] becomes [r]	
her	no vowel sound; [r] becomes [r]	
bottle	no vowel in the 2nd syllable; [1] becomes [1]	
funnel	no vowel in the 2nd syllable; [1] becomes [1]	
button	no vowel in the 2nd syllable; [n] becomes [n]	
rhythm no vowel in the 2nd syllable; [m] becomes [m]		

2.2.3 Flapping. In American English, we generally pronounce [t] by solidly placing the tongue against the alveolar ridge. However, sometimes the tongue tip touches the alveolar ridge much more briefly. The *flap* sound is transcribed as [D], as in the following contrasts:

bit	not flapped	[t]	
bitter	flapped		[D]
but	not flapped	[t]	
butter	flapped		[D]

2.2.4 Devoicing. Sometimes we make a voiced sound into a voiceless one, and when we do so, we use a diacritic to indicate *devoicing*. Consider the following examples:

lip	voiced lateral	[1]
please	devoiced	[ĵ]
limb	voiced lateral	[1]
climb	devoiced	[1]

2.2.5 Palatalization. *Palatalization* occurs when we produce a sound in a position that is further forward, specifically, when the place of articulation moves forward to the palatal region. Palatalization is inciated by a raised-y diacritic, as in the following example:

cot	unpalatalized	[k]
keep	palatalized	$[k^y]$

Finally, it is worthwhile noting that we have learned techniques that allow for much more exacting representations of speech sounds. When one attempts to represent speech sounds in these more exacting terms, one traditionally speaks of *narrow transcription*.

2.3 Some Consonants from Other Languages

In the various charts above, we used the notation NA to indicate that no relevant sound exists in American English. It should come as no surprise that many of these NA-blanks are, in fact, filled in other languages. Some of these sounds are indicated by new symbols. To provide but one example here, consider German, which has a voiceless palatal fricative (transcribed as [c]and seen in words like *mich*, [mɪç], English "me") and a voiceless velar fricative (transcribed as [x] and seen in words like *Fach*, [fax], English "field of study"). Other sounds are indicated not by new symbols like [c], but by the addition of diacritics. Some of these diacritics were introduced above; others will be introduced as the discussion proceeds or in exercises. A glance at the consonant chart on page XXV of Cowan & Rakušan will show some of the sounds found in other languages.

2.4 Consonant chart

We are now in a position to assemble all of this information together into a single, traditional consonant chart for American English.

Place of Articulation	Glottal State	Bilabial	Labiodental	Interdental	Alveolar	Palatal	Velar
Manner of Articulation							
Stop	voiced	b			d		g
	voiceless	p			t		k
Fricative	voiced		v	ð	z	ž/3	
	voiceless		f	θ	S	š/∫	
Affricate	voiced					j/dʒ	
	voiceless					č/t∫	
Nasal	voiced	m			n		ŋ
	voiceless						
Liquid Lateral	voiced				1		
Retroflex	voiced				r		
Glide	voiced	w				y/j	(w)
	voiceless	м					(M)

2.5 Vowels

The *vowels* are those syllabic speech sounds that have a more resonant quality and are produced with less obstruction than that seen with the consonants or semi-vowels. We divide vowels into two general types, the *simple vowels* and the *diphthongs*. Simple vowels do not show a change in their quality as they are pronounced in a single syllable; the vowels in words like *cat* or *pit* seem to retain their same quality all the way through. By contrast, diphthongs have an audible change in quality, as in the vowels in words like *cow* or *boy*. In this course, we will represent the American English diphthongs as combinations of glides and vowels; hence, the vowel in *cow* will be [aw], the vowel in *buy* will be [ay] (or [aj]), and so forth.

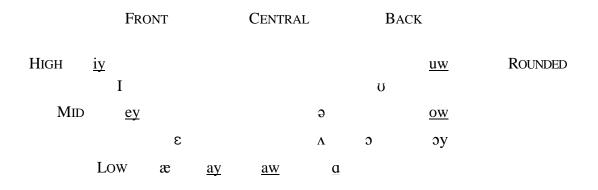
The basic parameters for describing (all) vowels (not just simple vowels or diphthongs) involve (i) the general position within the mouth where the vowel is articulated, (ii) the roundedness of the lips, and (iii) the general constrictedness of the musculature.

As for the general position within the mouth, one can discriminate vowels that are articulated in the front of the mouth from vowels that are articulated in the back. Hence, the vowels in words like *beat*, *bit* and *bat* are front vowels while the vowels in words like *boat*, *boot*, and *father* are back vowels. Looking at the vertical dimension, one can also discriminate vowels that are articulated high from vowels that are articulated low. *Beat*, for example, contains a high vowel while *bat* contains a low vowel; likewise, the vowel in *boot* is high while the vowel in *father* is low.

As noted above, lip rounding is one of the other basic parameters that is employed to describe the articulation of vowels. In American English, for example, the vowels in *boot* and *boat* involve rounded lips. Other vowels in American English are unrounded.

Finally, one can distinguish between vowels whose articulation includes a greater degree of constriction in the oral cavity and those that include less constriction. Compare, for example, the words *beat* and *bit* by saying them out loud. Note with *beat* that the oral cavity is somewhat more constricted than it is with *bit*. We call vowels that involve more constriction *tense*; those with less constriction are called *lax*.

Putting these dimensions together into a chart, we can observe the positions of the various vowels of American English:



American English Vowels Tense vowels are <u>underlined</u>; Rounded vowels are enclosed to the right.

To get a feel for the sounds that the symbols in the chart above represent, consider the following examples:

<u>Symbol</u>	<u>Examp</u>	ole	More examples
[iy]	heat	[hiyt]	b <u>ea</u> t, cr <u>ea</u> m, am <u>oe</u> ba
[I]	hit	[hIt]	b <u>i</u> t, b <u>ee</u> n, admon <u>i</u> tion
[ey]	bait	[beyt]	f <u>a</u> te, cl <u>ay</u> , gr <u>ai</u> n
[٤]	bet	[bɛt]	l <u>e</u> t, s <u>ay</u> s, s <u>e</u> ver, <u>gue</u> st
[æ]	bat	[bæt]	h <u>a</u> t, r <u>a</u> cket, l <u>au</u> gh
[ay]	tide	[tayd]	l <u>ie</u> s, m <u>y</u> , th <u>i</u> gh
[aw]	crowd	[krawd]	b <u>ow</u> s, pl <u>ow</u> , (to) h <u>ou</u> se
[ə] "schwa"	sofa	[sowfə]	Can <u>a</u> d <u>a</u> , ros <u>e</u> s, <u>a</u> fford, tel <u>e</u> graph
[\Lambda]	shut	[šʌt]	b <u>u</u> tt, <u>u</u> dder, t <u>ou</u> gh, l <u>u</u> cky
[a]	father	[faðr]	p <u>o</u> t, r <u>o</u> b, c <u>o</u> t
[ɔ]	ought	[ɔt]	b <u>ou</u> ght, <u>aw</u> ful, <u>o</u> ral
[U]	book	[buk]	sh <u>ou</u> ld, p <u>u</u> t, <u>goo</u> d
[y]	boy	[by]	l <u>oy</u> al, c <u>oi</u> n
[uw]	boot	[buwt]	t <u>o</u> , t <u>wo</u> , br <u>ew</u> , thr <u>ough</u>
[ow]	vote	[vowt]	n <u>o</u> te, th <u>ough</u> , <u>oa</u> f, <u>O</u> 'Neil

The two vowels $[\mathfrak{d}]$ and $[\Lambda]$ may appear to have very similar sounds. Note, though, that $[\mathfrak{d}]$ only appears in word positions that are unstressed whereas $[\Lambda]$ appears in stressed positions. We use the term *schwa* to refer to the symbol $[\mathfrak{d}]$; it is a so-called *reduced vowel*.

2.6 Vowels from other languages

When we examine the vowel chart above, it is important to keep in mind that these vowels represent only those in American English. Other languages may well have other, complete different vowels. For instance, the rounded vowels [uw], [ow], [o], [o], and [oy] represent American English; by contrast, the vowel in the German word *hübsch* (English "pretty") is a high front rounded vowel. Another example is from the language Mokilese (an Astronesian language spoken in the South Pacific), which has voiceless vowels. We represent these vowels with the diacritic discussed earlier for devoicing. Hence, the Mokilese word for English "firewood" is [supwo], where the first vowel has no voicing. See the chart of vowels on page XXVI of Cowan & Rakušan for a listing of some of these vowels.

You should at this point begin studying the phonetic illustrations provided in **Exercise 1** through **Exercise 74** in Cowan & Rakušan. *In addition, note that we will shortly be having two quizzes. The first quiz will involve reverse transcription (as in Exercise 10 at the end of this chapter), and the second will involve transcription (as in Exercise 11 in this chapter).*

3. Natural classes (first time around)

It is possible to speak of natural groupings of speech sounds by means of articulatory descriptions. One can, for example, meaningfully discuss the sounds [p], [b], [t], [d], [k], [g] and [?] by referring to the class of all stops. Likewise, it is possible to deal with the group of sounds comprising [b], [d], and [g] by reference to the voiced stops. Such groupings of speech sounds are called *natural classes*.

Grouping sounds into natural classes can be advantageous in the understanding of phonetic processes. Consider, for instance, our earlier discussion of aspiration in English. The examples we gave earlier for aspiration are repeated below for convenience.

pat[p^h]versus spat[p]tub[t^h]versus stub[t]cope[k^h]versus scope[k]

Is there a natural class at work here? It appears that aspiration appears just when we have a voiceless stop (a natural class) in initial position.

Sometimes, however, natural classes that are identified by means of the articulatory descriptors we have been employing do not entirely suffice. For this purpose, a number of other classificational terms are required, among them, oral, strident (or sibilant), sonorant, and obstruent.

3.1 *Oral sounds* are those articulated in the oral cavity (hence excluding, for example, glottal sounds like [?]).

3.2 *Strident sounds* (also known as sibilants) are those fricatives and affricates that are distinctly noisier than others when they are articulated. The strident sounds, for English, are [s], [z], [š], [ž], [č], and [j].

3.3 *Sonorant sounds* are those speech sounds that are produced with a relatively open passage for airflow and thus that have more acoustic energy. These sounds include the vowels, the glides, the liquids and the nasals.

3.4 *Obstruent sounds* are those that are produced with an obstruction of the airflow. These sounds include the stops, the fricatives, and the affricates.

Note that these classificational terms can group sounds together in ways that the simple articulatory descriptors used above cannot. Consider, for example, the stridents: Looking at the consonant chart, one discovers that there is no way to group together the sounds [s], [z], [š], [č], and [j] by using the simple articulatory descriptors that were employed to construct the chart.

4. Coarticulation

So far, we've seen that individual words can be broken up into segments that we represent with IPA symbols. Hence, the word *cat* appears as $[k^h æt]$ when transcribed. However, if we look carefully, we discover that individual speech sounds do not appear in isolation; instead, an individual speech sound may be influenced by the speech sounds that surround it (before, after or both). A good example is, again, the English word *cat*, which we might compare to the English word *scat* (as in *scram*!). Above we noted that *cat* is transcribed as $[k^h æt]$, that is, with aspiration on the voiceless stop consonant. Note, though, that when we transcribe *scat*, we find no aspiration: [skæt]. A possible generalization seems to be that aspiration on the voiceless stop consonants appears in English only when that consonant is at the beginning of the word and before the vowel. More generally, this observation indicates that speech sounds do not appear in isolation, that is, unaffected by what is around them; rather, individual speech sounds must be influenced by what is around them, that is, by their *phonetic environment*.

In fact, the articulatory adjustments that are made during the production of speech, which we call *processes*, can be classified into a number of different types, including especially *assimilation*, *dissimilation*, *deletion*, *epenthesis*, and *metathesis*.

4.1 Assimilation

As one might expect from the general meaning of the term, assimilation refers to the articulatory process in which one speech sound becomes more similar to a neighboring speech sound. Consider, for example, the vowels in the English words *cat* versus *can*. By listening carefully to the way in which people pronounce these two words, we discover that the vowel in *can* has a nasal quality to it; that is, the vowel in *can* has become *nasalized*. We thus transcribe these two words as $[k^h æt]$ and $[k^h æn]$, where the diacritic over the vowel in *can* indicates its nasal quality. *Crucially, note that it is the presence of the nasal consonant* [n] *following the vowel that causes the vowel to become nasal.* In other words, here we have a straightforward case of assimilation, where the vowel [æ] becomes more similar to the consonant [n] through *nasalization*.

One can differentiate between two types of assimilation, depending on the direction of the influence. For *can* above, for example, we found that [n] influences the vowel to its left. In other words, the direction of influence is backwards, from right to left. This type of

assimilation is called *regressive assimilation*. In other cases, however, we find that the direction of influence is not from right to left, but from left to right. Consider, for instance, German *infinitive* verbs (i.e., the form of German verbs that one would look up in a dictionary), which are all written with an *-en* ending in standard German orthography (e.g., *machen* "to make" and *leben* "to live"). When pronounced by native speakers, however, this *-en* ending is pronounced as [n], but not if the preceding consonant is a labial sound, in which case we find an [m] sound instead. Hence, native speakers pronounce *machen* as [maxn] because the consonant [x] preceding the [n] ending is not labial. However, the consonant [b] preceding the [n] ending in *leben* is labial, and this labial feature affects the final [n] sound, causing it also to be labial; as a result of this *progressive assimilation*, *leben* is pronounced by native speakers of German as [lebm].

4.2 Dissimilation

Another common process in the world's languages is dissimilation, which makes sounds more different—in effect, the opposite of assimilation. We can readily observe cases of dissimilation in American English. A good example is the word *sixth*, which is often pronounced by native speakers not as [sɪks0], but as [sɪkst], that is, by replacing the sound [θ] with the sound [t]. Why should they do so? The most obvious reason is, of course, that it's easier to pronounce [sɪkst] than it is to pronounce [sɪks0]. More technically, because [s] and [θ] are so close to one another in terms of their features—both are voiceless fricatives—it becomes difficult to keep them apart when pronounced; as a result, dissimilation kicks in to make [θ] into the [t] sound, which is a good deal less similar to [s].

4.3 Deletion

Another very common process among the world's languages is deletion, which (again, to ease articulation) removes a sound. Once again, English provides ample examples of the process. For instance, when we say, in natural speech, *She gave him a notebook*, we might well say something like [siy geyv III Λ nowtbok] where the glottal fricative [h] in *him* is deleted.

4.4 Epenthesis

The process epenthesis, also very common, occurs when a sound is inserted as an aid to articulation. Good examples are from English. Words like *something* or *prince* are, in fact, somewhat difficult to pronounce in slow speech: try saying [$sAm\theta n$] or [prins]. The transition between the sonorant sound (e.g., [m] in *something*) and the following non-sonorant sound ([θ] in *something*) causes difficulties in pronunciation. To help out, natives generally insert a consonant sound in these cases, resulting in more natural feeling pronunciations: A [p] is inserted into *something* to result in [$sAmp\theta n$], and a [t] is inserted into *prince* to result in [prints]. Both of these examples illustrate epenthesis.

4.5 Metathesis

Finally, we have the process metathesis, a process that changes the ordering of speech sounds, again as an aid in articulation. For example, native speakers of some American English dialects reverse the order of the consonant sounds in the word *ask*. As a result, instead of saying [æsk] (as is common in the other dialects), we find in this dialect the pronunciation [æks].⁴

5. Suprasegmentals (prosodic) features

Up to now, we've observed what the articulatory structures of speech sound are, how these segments can be classified into meaningful groups (natural classes), and how these segments may modify—or be modified by—surrounding segments (co-articulation phenomena like assimilation). Note, however, that nothing in what we have learned can explain, for example, the difference in pronunciation between the word *produce* when it is used as a verb (e.g., *Those countries produce oil*) and the word *produce* when used as a noun (i.e., to mean "fruit and vegetables", as in *The produce in that store is always overpriced*). Differences like these, as well as other differences not even present in American English, are referred to as

⁴ A rather special kind of metathesis appears when the initial sounds of different words are reversed. Usual examples of this special type of metathesis include *leak wink* (for *weak link*). The term *spoonerism* is commonly used to refer to this special type of metathesis. The term is in fond memory of a real-life person, the Reverend William A. Spooner, who was the Dean and Warden of New College, Oxford at the beginning of the 20th century and who, according to the popular mythology, produced a great many of this type of metathesis.

suprasegmental properties, which mean something like "properties above, over or beyond the segmental level" and which apply to sounds no matter what their place or manner of articulation is. Suprasegmental properties (sometimes called *prosodic* properties) are classified into three types: *length*, *pitch* and *stress*.

5.1 Length

Length means, quite simply, that a certain segment is held out longer than some other segment, all other factors held equal. For example, Czech includes the property length in the discrimination of vowels. The Czech word for the English word "loge" (i.e., in a theater), for instance, is identical to the Czech word for "bed" except that the vowel in the word meaning "loge" is long and the vowel in the word meaning "bed" is short: [loiže] versus [lože] (the diacritic ":" after the vowel indicates longer duration). Czech thus distinguishes long vowels from short vowels.⁵ By contrast, Italian distinguishes not only between long and short vowels, but also between long consonants and short consonants: [fano] for English "grove" and [fan:o] for English "they do".

5.2 Pitch

Pitch involves the modulation of both muscular tension across the vocal folds and the airflow that passes through the folds. Human languages involve two kinds of pitch, one that may make a great difference in the meanings of otherwise identical strings of sound (a type not found in English) and the other that does not affect the meanings of individual words. These two types are known as *tone* and *intonation*.

5.2.1 Tone. Tone languages are those in which differences in word meaning are modulated by pitch differences. For example, the sound-string [papa] in Akan (a language in the Niger-Congo family of languages) means "palm-leaf fan" when it is spoken with a low tone on the

⁵ English teachers in American high schools often speak of words like *beat* as having "long vowels" and of words like *bit* as having "short vowels".Note, however, that the difference between vowels like those in *beat* and *bit* actually involve the tense-lax distinction as well as the distinction between monophthongs and diphthongs. In Chapter 3 (Phonology), we will observe that English has long and short vowels, but that the distinction in English involves not pairs like *beat* [biyt] vs. *bit* [btt], but rather pairs like *beat* [biyt] vs. *bead* [biÿd].

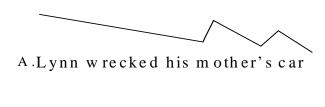
second syllable, but this same string means "father" when the second syllable has a high tone.

The tones seen in Akan are all level tones in that the pitch of the tone does not change across a single syllable. That is, each syllable is all in one tone or all in another tone. Tones of this variety are known as *register tones*. By contrast, in Mandarin Chinese, the sound-string [faŋ], which represents a single syllable, means "house" if it is spoken with a rising tone, and the identical string means "release" if it is spoken with a falling tone. Unlike Akan, Mandarin thus has tones that change in pitch across a single syllable.⁶ Tones of this sort are known as *contour tones*.

5.2.2 Intonation. Differences in pitch that do not modulate the meanings of words are known as intonation. We observe this type of pitch in English. For instance, the word-string *Lynn wrecked his mother's car* can be spoken in a number of ways. Two of these ways are illustrated in the figure below, where the lines above the utterances provide an approximation of relative intonation as the words are spoken. The first example (A.) represents a simple statement while the second (B) represents a rhetorical question (*Lynn wrecked his mother's car?!?*)

- [faŋ] high level tone "square"
- [faŋ] falling tone "release"
- [faŋ] middle falling tone "copy"
- [faŋ] rising tone "house"

⁶ In fact, Mandarin has four different tones, one of which is a register tone and the other three are contour tones. The complete inventory is illustrated below with [faŋ].



B. Lynn wrecked his mother's car

5.2.3 Intonation and tone together. The presentation of tone and intonation above might lead to the incorrect impression that the two do not appear together. A more interesting observation, however, is that while all human languages appear to employ intonation (e.g., English as well as Mandarin and Akan), only some languages employ tone. It is also interesting to note that when the two types of pitch are employed together, it is the *relative* pitch across and within syllables that is important, not *absolute* pitch.

5.3 Stress

Stress phenomena involve the perception that certain segments, generally vowels, receive prominence in their articulation. This perception arises due to the relative *loudness* of these segments. Hence, when native speakers of English articulate the word *stupid*, the vowel in the first syllable is louder than the vowel in the second syllable, and this relative loudness leads to the perception of a stressed first syllable and an unstressed second syllable.

In fact, stress can be somewhat more complex. Consider the word *intelligent*, which has four syllables. When articulated, native speakers place *primary stress* on the second syllable; however, note that there is also a certain level of stress on the last syllable, albeit at a level that is certainly lower than that placed on the second syllable. This is known as *secondary stress*. Linguists indicate primary and secondary stress by means of accent diacritics. Hence, *intelligent* would be transcribed as [Intéləj̈́nt], where the vowel nucleus of the second syllable bears primary stress (i.e., [ɛ́]) and the syllabic consonant making up the nucleus of the final syllable bears secondary stress (i.e., [ǹ]).

6. Phonetics and then ... Phonology

The study of human-language speech sounds and their articulation is known as articulatory phonetics. Note, though, that we have omitted discussion of how these various sounds are perceived by native speakers. For example, we noted earlier that Czech differentiates between long vowels and short vowels, as seen in examples like [lo:že] ("loge" in a theater) versus [lože] ("bed"). This means—as we shall discover in the next chapter—that native speakers of English might perceive [o:] and [o] as different, but essentially irrelevant since it doesn't differentiate between words; for the native speaker of Czech, however, [o] and [o:] are as different as night and day. Without it, they won't know if they had just dreamed about Shakespeare's *Hamlet* while in bed ([lože]), or if the play had managed to put them completely to sleep in the theater ([lo:že]).

Exercises

Exercise 1: Linguistics and Standard Orthography.

It seems uncontroversial to think that two linguists, especially ones who do not speak the same native language, might profit from using a system of sound representation in which *precisely one symbol relates to precisely one sound*. In other words, it would not do if one sound were represented by two or more symbols, or if one symbol represented two or more sounds, or if there are symbols that stand for no sound at all. Would the use of English orthography (the English 'spelling' system) suffice for linguistic work?

A. Consider the symbol comination *ough*. In the word *enough*, for example, *ough* seems to represent sounds like in the word *fluff*. Does the *ough* symbol combination represent any other sounds? Try to locate more:

Example: enough — fluff

1.	2.
3. 5.	 4. 6.
5.	6.

What does this example show?

a. one combination of letters = one sound

- b. one combination of letters = several sounds
- c. several different combinations of letters = one sound
- d. one combination of letters = no sound
- B. The spelling of the English word *occupy* suggests that the letters *cc* represent the 'k-sound'. Do any other single letters or combinations of letters also represent the 'k-sound'? Try to locate any other ways of representing the 'k-sound' for yourself.

Example: occupy (i.e., cc = k-sound)

 1.
 2.

 3.
 4.

 5.
 6.

What does this example show?

- a. one combination of letters = one sound
- b. one combination of letters = several sounds
- c. several different combinations of letters = one sound
- d. one combination of letters = no sound

- C. The English word *flat* suggests that the letter *f* represents the 'f-sound'. Do any other letters or combinations of letters also represent the 'f-sound'? Try for yourself.
 - 1.
 - 2.
 - -. 3.

What does this example show?

a. one combination of letters = one sound

- b. one combination of letters = several sounds
- c. several different combinations of letters = one sound
- d. one combination of letters = no sound

D. The English letters *p* and *b* are 'silent' in words like *clim<u>b</u>*, *dum<u>b</u>er*, and *<u>pneumatic</u>. Are there any other 'silent' letters in English?*

1. 3.	2.
3.	4.
5.	

What does this example show?

a. one combination of letters = one sound

b. one combination of letters = several sounds

c. several different combinations of letters = one sound

d. one combination of letters = no sound

E. What is your final recommendation on the use of the English 'spelling' system for linguists? Be sure to state your reasons for whatever recommendation you make.

Exercise 2: Provide a three-part descriptive statement for each of the following stops:

Example:	[p]	"voiceless labial stop"
	[b]	
	[t]	
	[k]	
	[d]	
	[g]	
	[?]	

Exercise 3: How many consonant sounds are there in each of the following English words?

1. at	2. math
3. psychology	4. cure
5. hopping	6. knowledge
7. awesome	8. elope

Exercise 4: Pronounce the following words and answer the questions below:

1. though	2. thought	3. form	4. view
5. zoom	6. silk	7. pan	8. boat
9. huge	10. choose	11. judge	12. buns

a) Is the first sound voiced or voiceless?

b) Is the last sound voiced or voiceless?

Exercise 5: Compare the following sounds and tell whether they have the same place of articulation:

[s] vs [l]	[f] vs [h]
[k] vs [ŋ]	[w] vs [j]
[p] vs [g]	[b] vs [f]
[1] vs [r]	[č] vs [j]
[m] vs [n]	[s] vs [v]
[j] vs [š]	[θ] vs [t]

Exercise 6: Compare the following sounds and tell whether they have the same manner of articulation:

[s] vs [θ]	[m] vs [ŋ]
[k] vs [g]	[r] vs [w]
[w] vs [j]	[č] vs [j]
[f] vs [š]	[h] vs [?]
[1] vs [t]	[z] vs [ž]
[ð] vs [v]	[č] vs [s]

Exercise 7: Give the IPA symbol for each of the following:

1. voiceless velar stop	2. voiced bilabial stop
3. voiceless glottal stop	4. voiced palatal fricative
5. voiceless palatal fricative	6. voiced interdental fricative
7. voiced palatal affricate	8. voiceless bilabial glide
9. voiced bilabial glide	10. voiced velar nasal

Exercise 8: Give the IPA symbol for each of the following:

1. tense high front unrounded vowel	2. lax high front unrounded vowel
3. tense low central diphthong	4. high back rounded diphthong
5. high front rounded vowel	6. mid front rounded vowel
(see Cowan & Rakušan, p. XXVI)	(see Cowan & Rakušan, p. XXVI)

Exercise 9: For word pairs below, indicate whether the vowel sounds are the same or different.

1. back - sat	2. hide - height
3. bid - key	4. least - heed
5. luck - flick	6. drug - cook
7. ooze - duece	8. sink - fit
9. oak - own	10. pour - port

Exercise 10: Reverse	Transcription.	Give English words	s for IPA re	presentations:

a [riyč]	b [rıč]	c [rĭj]	d [rowt]
e [kruwd]	f [fʌj̆]	g [kayt]	h [ækšņ]
i [piysə]	j [wʌrm]	k [rat]	l [ruwt]
m [krəld]	n [seyl]	o [slowp]	p [neyšņ]
q [krawd]	r [wenzdey]	s [eyžə]	t [klowð]
u [šuwleys]	v [bæŋ]	w [0awznd]	x [kanšəs]

Exercise 11: Transcription. Give the IPA transcription for the following, including aspiration.

1. tog	2. kid	3. attain	4. despise
5. elbow	6. haul	7. juice	8. thimble
9. peel	10. stun	11. Oscar	12. cooler
13. sigh	14. hulk	15. explode	16. tube
17. spell	18. cord	19. accord	20. astound
21. pure	22. wheeze	23. remove	24. clinical

Exercise 12: Provide the sounds (IPA symbols) that make up the following natural classes:

1. voiced labial consonants	2. voiceless oral stops
3. back vowels	4. high front vowels
5. oral sonorants	6. voiced obstruents
7. sonorant consonants	8. nasals

Exercise 13: Provide natural classes for the following groups of sounds:

[p], [t], [k], [?]
 [t], [d], [s], [z], [n], [l], [r]
 [θ], [ð]
 [n], [l], [r]
 [l], [r]
 [m], [n], [ŋ], [l], [r], [w], [𝔅], [𝔅], [𝔅]

Exercise 14: Consider the following questions in terms of natural classes:

1. In English, the plural marker "-s" does not always have the same pronunciation. Consider the words in the following two lists:

<u>A</u>	<u>B</u>
rocks	bees
tops	keys
steps	logs
fits	figs

(a)Transcribe the plural sounds of the words in List A and those in List B (IPA symbols!). (It isn't necessary to transcribe the whole words, just the "s" marker).

(b) What natural classes determine which sound appears?

2. Speakers of German learning English might pronounce the English words in List A below as indicated in List B.

	<u>A</u>	<u>B</u>
	[lowd] "load"	[lowt]
	[hæv] "have"	[hæf]
	[bIg] "big"	[bIk]
	[j̆ʌj̆] "judge"	[ľʌč]
but 🖙	[nuwn] "noon"	[nuwn]

(i) What natural class of sounds precipitates this change in pronunciation?

Exercise 15: Spanish. Consider the speech sounds [b] and [β], which represent the voiced bilabial stop and the voiced bilabial fricative, respectively. Examine the examples below carefully and decide whether the phonetic process that results in the [b] vs. [β] alternation involves assimilation, dissimilation, metathesis, epenthesis, or deletion. (HINT: To solve this problem, look *carefully* at the sounds that surround [b] and [β].)

[ußa]	grape
[futbol]	soccer
[sombra]	shade
[saβino]	cypress
[kaße]	it fits
[bastante]	plenty
[brinkar]	to jump
[sußo]	I climb
[ußo]	there was
[kluβ]	club

Exercise 16: German. Consider the sounds $[\gamma]$ and [g], which represent the voiced velar fricative and the voiced velar stop, respectively. Examine the examples below carefully and decide whether the phonetic process that results in the $[\gamma]$ vs. [g] alternation involves assimilation, dissimilation, metathesis, epenthesis, or deletion. Note that German has *front* rounded vowels: [ö] and [ü]. (HINT: To solve this problem, look *carefully* at the sounds that surround $[\gamma]$ and [g].)

ta:yə	days
na:yən	to nibble
tauyəniçts	a good-for-nothing
fu:yn	to fit together
auyn	eyes
gɛfloːɣn	flown
bo:yən	arch
zoyən	crystalize
ja:yən	to hunt
zi:gən	to conquer, win
bergə	mountains
fö:glajn	bird
mö:gən	to like
re:gən	rain
giŋən	went
ganges	River Ganges
ungarn	Hungary
zigna:l	signal
grok	grog

Exercise 17: English. Examine the transcribed instances of words or phrases in two different speech styles (e.g., slow, careful speech vs. normal speech; adult speech vs. child speech; etc.). Explain, for each, which phonetic process is involved (assimilation, dissimilation, metathesis, epenthesis, or deletion).

Word/Phrase	<u>Style #1</u>	<u>Style #2</u>
1. in his closet	[In hIZ klazət] (slow speech)	[In IZ klazət] (normal speech)
2. spaghetti	[spəgéDiy] (adult speech)	[pəskéDiy] (child speech)
3. police	[pəliys] (most dialects)	[p_iiys] (some dialects)
4. without	$[wi\theta awt]$ (slow speech)	[WIðawt] (normal speech)
5. filthy	$[fil\theta iy]$ (slow speech)	$[filt \theta iy]$ (normal speech)
6. hand-me-down	[hænd miy dawn] (slow speech)	[hæmiydawn] (normal speech)
7. waiter	[weytər] (slow speech)	[weyDr] (normal speech)
8. fifths	$[fif\theta s]$ (slow speech)	[fifts] (normal speech)
9. a drink of water	[A driynk AV water] (slow speec	h) [A wiynk A draDr] (normal speech)
10. eleventh	$\left[\Im l \acute{\epsilon} v \Im n \theta \right]$ (careful speech)	$[\Im l \acute{\epsilon} v n t \theta]$ (normal speech)

Exercise 18: English. For the sentences in *italics* below, draw intonation contours like the examples provided in the earlier discussion of *Lynn wrecked his mother's car*. Compare your rendition of these contours with those drawn by your classmates.

- 1. Frank: I didn't have time to buy any milk today. Harry: What?!? *You didn't go to the store?!?*
- Mary: What do you want to do this evening?
 Georgia: I don't know. *How about if we go to the movies?*
- 3. Tod: Did you see the meteor shower last night? Janet: Yeah. *It was really far out!*
- 4. Tina: I have to go see my advisor today. Tim: *What do you have to see her for?*

5. Customer: *How much is a one-way ticket to Albuquerque?* Salesperson: That's \$245 if you buy before the end of the month.

Exercise 19: Stress. For each of the following words, provide primary and secondary stress markings. Compare your markings with those of your classmates.

- 1. electric
- 2. bombastic
- 3. professor
- 4. jugular
- 5. olfactory
- 6. saccharin
- 7. tangerine
- 8. vindictive
- 9. Manhatten
- 10. coprophagous (If you don't look up the meaning of this one in your dictionary when you get home, you're going to be missing something.)