Swarthmore College

Works

Linguistics Faculty Works

Linguistics

1996

Linguistics: An Introduction

Donna Jo Napoli

Swarthmore College, dnapoli1@swarthmore.edu

Follow this and additional works at: <https://works.swarthmore.edu/fac-linguistics>

Part of the [Linguistics Commons](#)

[Let us know how access to these works benefits you](#)

Recommended Citation

Donna Jo Napoli. (1996). "Linguistics: An Introduction". *Linguistics: An Introduction*.
<https://works.swarthmore.edu/fac-linguistics/58>

This work is brought to you for free by Swarthmore College Libraries' Works. It has been accepted for inclusion in Linguistics Faculty Works by an authorized administrator of Works. For more information, please contact myworks@swarthmore.edu.

1

Phonetics

Chapter Organization—for the Instructor

This chapter covers a lot, probably more than you've ever tried to handle in an introductory course, where phonetics is often the area of linguistics that gets short shrift. But I urge you to try to cover as much of it as you can. A solid knowledge in phonetics is an invaluable foundation for phonology. With it, the current approaches to phonology seem insightful and delightful. Without it, they can seem arbitrary.

We begin with articulatory phonetics, going through the consonants of English, from the front of the mouth back toward the throat, introducing distinctive features as they become necessary for distinguishing among the consonants. We then look at consonants from other languages with the goal of being representative rather than exhaustive. Then we move to an inventory of the vowels of American English, adding in new distinctive features and recognizing the variety of types of syllable nucleus. We augment this with vowels from other languages. Finally we move into acoustic phonetics, covering pitch and intensity. We study formants and learn to read several kinds of information from spectrograms. An appendix displays the International Phonetic Alphabet (IPA) and distinctive features of commonly found consonants.

Okay, let's go

Here's the situation. You live in America and you speak American English (the English spoken by native speakers in the United States of America). You have a pen pal in New Guinea. Your pen pal writes and reads English but has never heard the language spoken. Your pen pal asks you in a letter how to pronounce English. What do you do?

You make a tape recording and mail it to New Guinea. But guess what? Your pen pal has no tape recorder—and no source of electricity anyway.

At this point you have to get clever. You decide to describe the sound of English. You might begin word by word, say with the word *dear* (since that's how you start your monthly letters to this pen pal). If you know your pen pal's language (which is Ku Waru, since your pen pal is a Papua New Guinea Highlander of the Nebilyer Valley in the Western Highlands Province), you might try comparing the sounds of *dear* with the sounds of some word or words in Ku Waru. Nice try, but I won't let you get away with it. I

declare you are a less well-traveled American and, in fact, you had never even heard of Ku Waru before you began writing to your pen pal.

At this point you might try to find out what languages, if any, you both have a speaking knowledge of. Let's say you discover one and it's Italian (an Indo-European language in the Romance family). So now you decide to compare the sounds of the word *dear* (as you say it) to sounds in Italian words. If you are a good speaker of Italian, you have a little problem right off with "d." That's because the Italian "d" of *dirò* 'I will say,' for example, isn't quite the same as the American one. You might not be sure what the differences are, but your ear detects them. If you are a poor speaker of Italian, you might not even know the two "d"s are not the same, and you might go blithely on to the next sound in *dear*.

Let's say somehow or other you get to this next sound.

And now the real trouble starts, for unless your ear is like a stone, you have to notice that the VOWEL sound in the English word lasts longer and changes somehow in quality, while the vowel sound in the Italian word, which starts out the same (or close to the same) as the sound in the English word, maintains a steady quality. What are you going to do? What is the change in vowel QUALITY in the middle of this English word that isn't happening in the Italian word?

Let's say somehow or other you get past this problem and hit the "r."

Yikes. How can both the English "r" of *dear* and the Italian "r" of *dirò* go under a single title or rubric? They don't sound very much alike at all.

If you don't know what the Italian word sounds like, this discussion may have left you cold. But think of Spanish, French, Chinese, Vietnamese, Swahili, Amharic, whatever language you have ever heard even in passing in a shopping mall or in a line at the movies. Whether you ever found out its name or not, you must have heard another language at some point in your life. So you know very well that the sounds of the world's languages are not all the same. That's the point. That's exactly the point.

So what do you do?

Articulatory Phonetics

You talk about what we all have in common: physiology. And that's precisely what linguists do. They describe language sounds based on how the body produces them. This is called ARTICULATORY PHONETICS. Let's see how it works.

To begin, list twelve MONOSYLLABIC words of English (that is, words consisting of precisely one syllable). Any twelve. Here's a typical list:

car	dog	house	law	my	knee
rip	I	walk	few	ma	sleep

Look at them. Look at all those letters. Say each word. Is there any word there that is written with a vowel that you can't hear? Sure: *house*—just look at that final "e." But look a little closer. See the MEDIAL vowel combinations (in the middle of the word) in *sleep* and *house*? Are there really two equally prominent vowels in the middle of each of these words? We will find out that the answer is no.

Is there any word there that is written with a CONSONANT that you can't hear? Sure: *knee* (the "k") and, for most of us, *walk* (the "l"). Again, though, a closer look should

make you stop at *law*. Say it. Compare it to *few*. Unless you are from one of a few areas in the American South, you probably don't pronounce a final "w" in *law*.

So ORTHOGRAPHY (spelling) can include symbols (here, Roman alphabet letters) that don't correspond to any sound in the word.

Now look again at these words. Do you hear any sounds (whether consonant or vowel sounds) that aren't represented by a written letter? Sure: in *my*. Where's the letter corresponding to the vowel sound(s) we hear? Perhaps you had a high school English teacher who told you "y" could be a sort of honorary vowel, so you feel sure that "y" does, indeed, represent the vowel sound(s) that follows "m" in *my*. You may be confident that the letter "y" in *my* does not represent the same sound as the letter "y" in the word *you* at the start of this sentence. One is the honorary vowel (so to speak) and the other is a consonant. In fact, you're right. And you can demonstrate this difference yourself.

To see this we need to consider one more word: *high*. Say *my* and *high*. They rhyme, right? So they end with the same sound. Now put two fingers in your mouth sideways, so that one pushes up and the other presses down. Try to say *high*. It doesn't sound too bad. Keep your fingers there and try to say *you*. It sounds awful. Do it again. The final sound of *high* can be said with your two fingers in your mouth like that, but the initial sound of *you* can't be. That means that these sounds are not identical. So the final sound of *my* (which is the same as the final sound in *high*) and the initial sound of *you* are not identical. In other words, "y" does not always represent the same sound, and your high school English teacher was not a fool.

Before going on with our discussion of the sounds in *my*, stop for a moment and explain why, in the experiment we just did, we used *high* and *you* instead of contrasting *my* and *you* directly. What irrelevant interference does *my* present? Sure, you got it. The "m" can't be said with the fingers in the mouth like that, because to say "m" we have to close our lips. So we had to switch to a new word that didn't have sounds like "m."

Returning to *my*, though, we find that even if we allow "y" to represent something vowel-like, that still can't be the whole story for most speakers of English. Unless you speak a southern variety of English, you have three sounds in the word *my*. To see this, compare *my* to *ma*. Both words start with the same consonant and then go into the same vowel sound. But the sounds of *my* keep going after the sounds of *ma* have stopped. There are three sounds in *my*, but only two sounds in *ma*. Can you hear that? Say them quickly, then slowly, then quickly. Do you see? When you say both words quickly, it can be very difficult to hear that *my* involves two vowel sounds. But the slow pronunciation should help you. When you say *ma* very very slowly, you simply hold onto a single vowel sound for a long time. But when you say *my* very very slowly, you can hear the change from one vowel sound to another and you can feel the tongue change position from low in the mouth to high. If you are still having trouble, look into a mirror as you pronounce each word slowly. You will see the jaw stay quite open at the end of *ma*; but the jaw will move from very open to more nearly closed at the end of *my*. That's because *ma* consists of a consonant and one vowel, but *my* consists of a consonant and two vowels. So no matter how clever we might be about categorizing "y" (as sometimes vowel-like and other times a consonant), we still have to face the fact that *my* has only two letters, but three sounds. Somehow a sound is not represented by a letter (or else a single letter is representing more than one sound).

If you are a southerner, you may feel left out in the cold right now. You pronounce *my* with only two sounds: a consonant and a vowel. The remarks below, however, should

make the point for your speech as well (the point being that the orthography of a word can have fewer letters than the sounds of that word), so please bear with me.

What's going on with *few*, another word on our list of monosyllabic words of English? Compare it to the child's exclamative *phoo*. The initial sound in both is [f] (so orthography fails us once more) and they share a vowel. But *few* is longer than *phoo*—it has more sounds. Listen to what comes between the initial consonant and the vowel that is common to both words. *Few* has another sound after the initial consonant that is missing from *phoo*. But *few* has no letter between the “f” and the written letter “e” to represent that sound.

So the orthography of a word can fail to represent sounds that actually occur in a word.

And you already noticed that the [f] sound can be represented by the letter “f” or the letters “ph” (in our discussion of *few* vs. *phoo*).

And now, to complete the picture, think about the initial sounds of *cat* and *celery*. No identity there. So a given letter can represent more than one sound.

We have seen four shortcomings of orthography in English. First, letters can appear that correspond to no sound. Second, sounds can occur that correspond to no letter. Third, different letters can correspond to a single sound. Fourth, a single letter can correspond to different sounds. In other words, there is no beautifully clean correspondence between sounds and letters in English.

This is no accident. It has to do with the history of our language and the fact that our spelling reflects pronunciations used when the spelling conventions were adopted. Language changes over time but orthography is fixed at a certain time and reflects only the pronunciation of a certain group of people at that certain time. (You'll get a chance to discuss some issues related to this fact in Problem Set 1.1.)

What matters to us now is that our Roman alphabet will not allow us to have a one-to-one correspondence between the sounds of English and written symbols. For this reason, we are not going to use that alphabet, but, instead, a new system called the INTERNATIONAL PHONETIC ALPHABET (or IPA), in which there is a one-to-one correspondence between sounds and symbols. When we write words in the IPA, we say we are TRANSCRIBING those words into IPA. Fortunately for us, many of the symbols in the IPA are based on the Roman alphabet, so we will have many fewer new symbols to learn than a linguistics student who uses a different alphabet or a nonalphabetic writing system. In the appendix at the end of this chapter, the IPA is explained. The meaning of the DIACRITICS (that is, the marks added to these symbols to indicate various modifications in sounds) introduced in this chapter are also given. Feel free to consult the appendix whenever you need to.

From here on out, every time we talk about a sound, we will use the IPA symbol for that sound. Furthermore, when we speak of consonants and vowels, we will be speaking of sounds, not of letters.

Minimal Pairs and Consonants

Okay, let's return to our list of twelve monosyllabic words. Let's try to come up with all the consonants of English. We can start on this task by considering the word *my* again. I claim that *my* is a separate word of English from *by*. You agree with me, right? They mean two different things—that's what makes them different words. However, they also sound different—and the difference is minimal: They contrast on only one sound. (Be careful here: Sometimes two different words can sound the same, as in *red* and *read*. These are

HOMONYMS. What we are concerned with in the text is words that differ in both their meaning and their sound.) For this reason pairs like *my* and *by* are called **MINIMAL PAIRS**. **The fact that *my* and *by* are different words means that the difference between them—which amounts to the difference between [m] and [b]—is DISTINCTIVE in English:** this difference matters to us in identifying and distinguishing between words. **Single sounds that are distinctively different from other sounds in whatever language we are considering are called PHONEMES.**

What other consonant phonemes does English make use of? Take the word *my* and form as many minimal pairs with it as you can. You should come up with several. Write them all down. Then organize them according to the physiology used to speak them. I'll help you. Right now please put your hand over the paragraph immediately below, so that you can't see it. Now, answer: what parts of the body do you use in making sounds that are parts of words? List them. Once you've done that, read the next paragraph.

You should have come up with at least these: lips, teeth, tongue, roof of mouth.

Don't limit yourself just to considering the mouth, however. What else? Your nose, definitely. To see this, say the words:

cat can

Now pinch your nose shut and say them again. *Cat* sounds the same; but *can* doesn't. Focus on the vowel and what follows it. Say these two words again. With your nose pinched, *cat* has three sounds made in the **ORAL CAVITY** (the mouth), but *can* is different. With your nose pinched, the initial consonant of *can* sounds the same, but the vowel and the final consonant sound different. That's because while you are still making the vowel, you prepare to make the [n] by opening up the **NASAL CAVITY** (in a way you will learn later—you do this whether you have pinched your nose shut or not). The vowel is therefore **NASALIZED**. Feel how that nasalized vowel resonates inside your nasal cavity? But what makes this nasalized vowel different from the vowel if your nose was not pinched shut is that the air that passes into the nasal cavity cannot escape through the nose. The final consonant is also different for the same reason—the air in the nasal cavity cannot pass out through the nose, as it normally does in making [n]. Now take the little hum that we can make to signal a positive response, as in “mmmhmmm.” Say it. Now pinch your nose shut and try to say it. See the difference? So the nasal cavity is involved in making [m], as well.

What else is involved in making language sounds?

Warning: By this time you have learned that when I ask a question, I quickly answer it. That can lead you to simply stop bothering to try answering for yourself. Please don't get lazy. This book will be much more valuable to you if you read along with one hand ready to cover the material that follows a question mark. Okay? Be a sport—try it my way.

I'll repeat the question: What else is involved in making language sounds? Your lungs, of course. And the part of your throat that leads into your windpipe. In fact, different languages can use sounds made by activity in different areas of the throat all the way down to and including your voice box (the **LARYNX**, the front protruding part of which you might call the Adam's apple). The air passages above your larynx are known as the **VOCAL TRACT**.

Other factors are involved in making and distinguishing language sounds. But for now I want you to concentrate just on the oral cavity. Starting at the very front of the mouth and moving inward toward the throat, organize the minimal pairs you have come up with

according to whether they do something significant at some given point, which we will call the **PLACE OR POINT OF ARTICULATION**. You should come up with a list that includes:

(1)	m	my	p	pie	b	by
(2)			f	fie	v	vie
(3)			θ	thigh	ð	Thy
(4)	n	nigh	t	tie	d	die
			s	sigh		
					l	lie
(5)			ʃ	shy		
(6)					ɹ	rye
(7)					ɡ	guy
					w	why
(8)			h	high		

The letters and symbols are from the IPA. The words following them begin with the sounds these symbols stand for. You should have all these consonant phonemes that I've listed here (unless you find *nigh* so totally archaic that you skipped it), in the same numbered lists I've put them in, although you may well have ordered them differently within those lists. However, you might have been unsure whether to put [w] in the list in 1, since it uses the lips, or in the list in 7 (where I put it), since it uses the back of the tongue. In fact, [w] has two places of articulation, a **PRIMARY** one inside the mouth (your tongue bunches and the thick back of it rises toward the top of the back of the mouth—can you feel it?) and a **SECONDARY** one at the lips. We'll return to the idea of secondary articulations later.

Some of you might have put words like *cry*, *fly*, and *sly* on your lists. Actually, while spelling is often not terribly helpful in English in figuring out how many sounds we have in a word, in examples like these the spelling can be instructive, for each of these words begins with a consonant **CLUSTER** (a series of more than one consonant). Still, you were being sensitive to the fact that there is something unusual about the second consonant in each of these words, unusual with respect to place of articulation. When we discuss "r"-like and "l"-like sounds later, you can return to these three words and then you can explain for yourself why you might not have recognized at first that they begin with two consonants.

Each of the sounds in 1–8 is made at the same place of articulation as the other sounds in that same example. We say that each example in 1–8 contains **HOMORGANIC** sounds (sounds made at the same place of articulation). So [m], [p], and [b], for example, are homorganic.

We now have seventeen consonant phonemes of English. But there are more. To show there are more, we'll have to change the standard word that we've been using for forming our minimal pairs. Let's take the word *moo* and try forming minimal pairs with it by varying the initial consonant. Write them down. Let's repeat the lists above and fill in the new consonants where they fit (and for those of you who were unhappy with *nigh*, you can now use *new* to fill in the consonant [n]):

(1)	m	my	p	pie	b	by
(2)			f	fie	v	vie
(3)			θ	thigh	ð	thy

(4)	n	nigh/new	t	tie	d	die
			s	sigh	z	zoo
					l	lie
(5)			ʃ	shy		
			tʃ	chew	dʒ	Jew
(6)					r	rye
					j	you
(7)			k	coo	g	guy
					w	why
(8)			h	high		

We have added in five more consonants: [z], [tʃ], [dʒ], [j], and [k]. And this is as far as we can go with minimal pairs that contrast by the initial consonant. (You might notice that two of these symbols, [tʃ] and [dʒ], are complex in the sense that they look like a combination of two symbols. In fact, the sounds these symbols signify are, arguably, equally complex. We will go into this when we talk about affricates.)

However, there are still two more consonant phonemes of English. One of them occurs only in a syllable-final position in words in English, which is why we couldn't find it using minimal pairs that contrasted on their first sound. So now let's form minimal pairs by varying the final consonant rather than the initial one. Take the word *sip*. Vary the final consonant and you'll come up with many words that are only minimally different from it, including:

sib (short for *sibling*), sin, sit, Sid (a proper name),
sis, sill, sing, sick, cig (short for *cigarette*)

The new consonant is the final sound in *sing*, and it's made in the same place [k] and [g] are made. Its IPA symbol is [ŋ], which is called "eng" or "engma," so enter it in the list in 7 like so:

(7)	ŋ	sing	k	coo	g	guy
					w	why

The last consonant phoneme for us to find occurs word-initially only in words that are borrowed from French (another Romance language like Italian, spoken in France) and maintain a strong French flavor. You know it as the first sound in the name *Zsa Zsa Gabor* (at least as Americans say it) and in the word *genre*. However, it can occur medially or in word-final position in many words that we do not today associate with French, such as *vision* and *garage*. Minimal pairs for this sound are hard to find, but consider the medial sounds in:

confusion Confucian
allusion Aleutian

This sound is written as [ʒ], and you can now fill it into the list in 5:

(5)			ʃ	shy	ʒ	genre
			tʃ	chew	dʒ	Jew

Our final list of the consonants of English contains these twenty-four:

(1)	m	my	p	pie	b	by
(2)			f	fie	v	vie
(3)			θ	thigh	ð	thy
(4)	n	nigh/new	t	tie	d	die
			s	sigh	z	zoo
					l	lie
(5)			ʃ	shy	ʒ	genre
			tʃ	chew	dʒ	Jew
(6)					ɹ	rye
					j	you
(7)	ŋ	sing	k	coo	g	guy
					w	why
(8)			h	high		

When we refer later to the examples in 1–8, please use this final list.

Places of Articulation and Articulators

Each of the lists in 1–8 groups homorganic sounds. Move around the various parts of your mouth. What can move? Your lips, your tongue, your bottom jaw. But your top jaw stays fixed. For this reason, we talk about an **ARTICULATOR**, which moves, and a point or place of articulation which that articulator moves toward. The articulators are along the bottom of the mouth. The places of articulation are along the roof of the mouth.

What are the various places of articulation on the roof of the mouth? You have a top lip and top teeth, of course. Now stick your tongue behind your teeth and run the **TIP** of it along the roof of your mouth as far back as you can go. The bony ridge behind your upper teeth is called the **ALVEOLAR RIDGE**. The hard part of the roof of the mouth that rises up behind the ridge is called your **HARD PALATE**. The tip of your tongue can easily reach to the point where the palate starts to get more fleshy, although it can't travel very far onto that fleshiness. That part is called the **SOFT PALATE** or the **VELUM**. And when your tongue is curled back like this, the tongue is said to be **RETROFLEXED**.

Go to a mirror now and look in your mouth. You can see a little fleshy part hanging down back there. That's called the **UVULA**, and sounds made at the uvula or directly behind it are called **UVULAR** sounds. The area inside the throat deeper than the uvular area is called the **PHARYNX** and sounds made there are called **PHARYNGEAL**. The area below the pharynx is called the larynx, and it produces **LARYNGEAL** sounds.

Now consider the parts of the bottom of your mouth. Again you have a lip and teeth. But you also have a tongue, which can be used in many different ways. We will talk about the tip of the tongue, the **BLADE** (the wider part behind the tip), the **BACK** or **DORSUM**, and the **ROOT**. If you close your mouth and rest your tongue flat inside it, the tip will be right behind the bottom teeth, the blade will line up with the hard palate, and the back will line up with the velum.

In Figure 1.1 you can see all these areas labeled.

The drawing shown in Figure 1.1 is called a **MIDSAGITTAL SECTION**, because it shows what we would see if we could take a picture of the inside of the middle of the head.

Figure 1.1. The numbers in this figure correspond to the following points of articulation and articulators: (1) lips; (2) teeth; (3) alveolar ridge; (4) hard palate; (5) velum; (6) uvula; (7) pharyngeal area; (8) epiglottis; (9) laryngeal area; (10) larynx with vocal cords; (11) tip of tongue; (12) blade of tongue; (13) back of tongue; and (14) root of tongue.

Okay, now look at the sounds listed in 1, which I have augmented with [w]:

[m]	[p]	[b]
		[w]

What is their place of articulation? (top lip) What is their articulator? (bottom lip) These are called the **BILABIAL SOUNDS**. We say that any sound which makes use of either or both of the lips has the **FEATURE** of [+LABIAL].

As our discussion progresses, we will introduce more and more features, which we will use as classification devices for sounds. **All of the features we talk about will be BINARY** (so [+labial] contrasts with [−labial], for example, and every relevant sound has the value of + or − for each feature). And a given sound will have several features, which we group together into what we call a **FEATURE BUNDLE**. The usefulness of this notion will become apparent later.

All four of the sounds discussed above, [m], [p], [b], and [w], are made with the same articulator and same place of articulation. So what makes them different sounds? What other features are we using to distinguish between them?

We have already noted that [w] is doubly articulated. (It has two places of articulation and two articulators.) So we can hold off on further discussion of [w] for the moment.

But what distinguishes [m], [p], and [b]?

We noted above that in making [m] the air flows through the nasal cavity. In fact, for [p] and [b] the whole velum is raised so that it touches the back wall of the vocal tract and it closes off the flow of air into the nasal cavity, forcing all of it to escape through the oral cavity. (This might surprise you. You might have thought the velum was stationary. It's not.) But for [m] the velum is lowered, and both the oral and nasal cavities fill with air from the lungs. This contrast is shown in Figure 1.2*a* (with [m]) and 1.2*b* (with [b]). Compare the position of the velum in Figure 1.2*a* and *b*. In Figure 1.2*b* the velum is raised in preparation for making the oral sound [b].

We call sounds with a lowered velum **NASALS** (which is the name of the feature that distinguishes sounds made with airflow through the nasal cavity). So we can distinguish [m] from [p] and [b] by the fact that [m] is nasal but the other two are oral. We say that [m] has the feature of [+NASAL], while [p] and [b] have the feature of [-nasal].

Now we need to distinguish [p] from [b]. Take a piece of paper and hold it up flat in front of your mouth less than an inch from your lips. Say *pad*. Now say *bad*. The paper should get pushed away from your mouth more at the start of *pad* than at the start of *bad*. Now put your hand, palm facing you, right in front of your mouth. Repeat the experiment. You can feel a puff of air at the beginning of *pad* that you don't feel at the beginning of *bad*. (Some speakers can feel a puff of air at the beginning of both, but it is stronger and more explosive at the beginning of *pad* than at the beginning of *bad*.) That strong puff or explosion is called **ASPIRATION**. We say that the [p] of *pad* is aspirated, but the [b] of *bad* is not. The feature we use to indicate aspiration is [+SPREAD GLOTTIS], because the spreading of the **VOCAL CORDS** at the release of a complete closure causes the aspiration. (The vocal cords, the space between them, and the area around them together are called the **GLOTTIS**.) We write an aspirated [p] in this way: [p^h]. In fact, in my speech every occurrence of [p] above should have been written as [p^h].

Phonemes and Allophones

We now have a nice difference between those two sounds. Unfortunately, if you are an Italian speaker or a French speaker or a German speaker or a Japanese speaker or a speaker of many other languages and you try that same experiment (with the paper and then with the hand), you will not be able to reveal a difference between [p] and [b]. That's because these languages don't aspirate either sound.

And, in fact, there are times when the speaker of English doesn't aspirate this sound. Repeat our experiments, saying the minimal pair *lap* and *lab*. The paper doesn't move for either sound. You don't feel a puff on your palm for either sound. Neither [p] nor [b] is aspirated in these words. That's not to say that you can't aspirate the [p]. If you want to, you can add aspiration, perhaps to give an emphatic focus. But normally you don't. And this contrasts with the situation in *pad* vs. *bad*, where the aspiration is obligatory on the first consonant in the first word.

Make a list of several minimal pairs that contrast [p] and [b] in initial position and then several that contrast [p] and [b] in final position. What do you notice? Now don't do the experiment—just consider the minimal pairs *pill* and *Bill* and *tap* and *tab*. In which word do you expect [p^h]? You should expect it in *pill*.

Now make a list of several **POLYSYLLABIC** words that have [p] and [b] somewhere other than in word-initial or word-final position—words such as *repeat*, *imbecile*, *compound*,

Figure 1.2. (a) The lips meet. The velum is lowered. The vocal cords vibrate (as indicated by wavy lines in the larynx). (b) The lips meet. The velum is raised (so air does not pass through the nasal cavity). The vocal cords vibrate.

rebuke, etc. Is [b] ever aspirated? (no) Is [p] in these words? Okay, so in general, is it word-initial position or syllable-initial position that we find [p^h] in? (syllable-initial)

Now consider another question: When [p] occurs in syllable-initial position, is it always aspirated? To answer this one, consider words such as:

potato	punctual	potentate	preponderance
coupon	rampage	capacious	apiary
campus	tapping	reparations	copulate

You should feel quite sure that in word-initial position we find [p^h]. And I hope you feel rather sure that in the second row of words we also find [p^h]. But in the third row you probably do not aspirate the [p]. Can you hear that? In all of these words, the [p] or [p^h] is syllable-initial (you will learn about how to syllabify words later and in the section “Syllable Types” in Chapter 2), yet not all of them require aspiration. We find that when [p] is initial to (that is, begins) a truly unstressed syllable, the [p] is typically not aspirated.

In other words, it is predictable that in the initial position of a stressed syllable we will find [p^h] and not [p] in English (although we’ll find [p] in Italian). And it is predictable that [b] will never be aspirated in English. For this reason we say that the feature of [+spread glottis] in some instances of [p] is **not distinctive in English—it is REDUNDANT. It is a difference that occurs automatically in a certain PHONOLOGICAL ENVIRONMENT OR CONTEXT OR CONDITION.** (Here initial position of a syllable is the phonological environment for aspiration.) We say, therefore, that [p] and its aspirated counterpart [p^h] are not separate phonemes, but, instead, ALLOPHONES of a single phoneme. That is, our basic phoneme is /p/, which we write between slanted lines (a point I’ll return to). But when /p/ finds itself in the initial position of a stressed syllable, it gets aspirated and it comes out as [p^h]. (A symbol between brackets represents the sound as we say it.)

The simplest way to recognize whether two sounds are separate phonemes or allophones of a single phoneme is to look for minimal pairs. If you are lucky enough to find a minimal pair, the two sounds in question are separate phonemes. If you don’t find a minimal pair, you need to check further, since not every sound will participate in a minimal pair with every other sound. (You already knew that. Remember how we started forming minimal pairs using *my* for a standard, but then we had to switch to using *moo* because not all of our English consonants participated in a minimal pair with *my*?)

What you’ll be looking for is predictability. If you can see that one sound occurs only in a given environment or in a restricted set of environments, and another very similar sound occurs in all other environments, then you are probably dealing with allophones, where the sound that occurs in “all other environments” is the BASIC or UNDERLYING form, and the sound that occurs only in a given environment or in a restricted set of environments is the DERIVED form. When we write our sounds using the IPA, we make a distinction between what we know must be the underlying form and what we actually say. What we know must be the underlying form we write between slanted lines, and we call it our PHONEMIC OR UNDERLYING REPRESENTATION (UR). What we actually say we write between square brackets, and we call it our PHONETIC REPRESENTATION (PR).

The importance of the distinction between UR and PR will become clear when we turn

to Chapter 2 (and see particularly the section called “Why Have UR and PR?”). But even now it’s worth it for you to be aware of this distinction.

To see how we use this terminology, let’s contrast *sip* with *pass*. The last sound of *sip* can be pronounced as [p] or [p^h] (or, actually, a couple of other pronunciations, as well—but let’s consider only these two here). If we say it as [p], then it has an identical UR and PR: /p/ and [p]. The first sound of *pass*, on the other hand, must have a different UR and PR: /p/ and [p^h].

An important fact about phonemes and allophones is that two sounds which may be allophones of a given phoneme in one language might well be separate phonemes in another language. For example, there are languages in which the feature of [+spread glottis] on any consonant is distinctive, and thus /p/ and /p^h/ are separate phonemes. We will be talking more and more about allophones in this chapter and throughout Chapter 2, so don’t worry if you’re not yet comfortable with this idea. You will be.

Voicing

Let us return now to our discussion of the distinctive difference between /p/ and /b/. We have learned that we cannot point to aspiration as that difference. There must be another difference that is a distinctive difference, for even though [p] is not aspirated in syllable-final position in English, we still can distinguish it from [b] in that position. And even though Italians and Japanese do not aspirate [p] in any position, they still can distinguish it from [b]. So there has to be some other difference between the two—a difference that is not predictable from context.

Let’s start again with *tap* and *tab*. Put your hand on your throat so your palm covers your larynx. Say the words. Can you feel more vibrations in your hand at the end of *tab*? You should be able to. Try putting both hands over your ears, forming a suction with your palms. Now say them. See how *tab* seems to reverberate more inside your head?

Actually, because the [p] and the [b] follow vowels in these examples, it may be hard for you to feel the difference with either of those tests. So if you are having difficulty, you could try getting ready to say *pan*, but don’t actually say it. Just close your lips and do whatever it is you do before you say the initial consonant. Now do the same for *ban*. With each, put your hand over your larynx, and then put your palms over your ears. Do you feel the difference? If you are still having difficulty, just believe me for now that the [b] reverberates more in your head than the [p] does, and when we get to the discussion of [f] vs. [v], you can try these two tests again, simply holding the [f] sound for several seconds and then holding the [v] sound for several seconds, first with your hand over your larynx and then with your palms over your ears. I guarantee you will feel the difference then.

The difference here is in your vocal cords. They are vibrating quickly with [b], but they are not vibrating or only barely vibrating with [p]. This distinction is called VOICING. We say that [p] (and [p^h], as well) is VOICELESS and [b] is VOICED; [p] has the feature [–voiced] and [b] has the feature [+voiced]. Try whispering the words *pad* and *bad*. Can you tell the difference? Aspiration allows us to distinguish between the two words. But now whisper the words *tap* and *tab* with no aspiration on the final consonants. When we whisper, the vocal cords don’t vibrate, so the voicing distinction in the final consonants is lost. Can you hear how the difference between word-final [p] and [b] is removed with whispering?

Many people can still distinguish between *tap* and *tab* when whispered, but that is

because they are paying attention to another difference between the words. Say *tap* and *tab* again, this time without whispering. Pay close attention to the vowel sound. Can you see a difference? The vowel in *tab* is slightly longer than the vowel in *tap*. It appears to be a language universal that vowels have greater DURATION or QUANTITY before voiced consonants than before voiceless consonants. When people whisper, they maintain a lengthened vowel before a consonant that would be voiced if they weren't whispering. It is the vowel LENGTH difference that allows you to distinguish *tap* from *tab* when they are whispered, since the rest of these words are identical when whispered.

For English speakers the distinction between voiced and voiceless consonants is relatively light. For speakers of Italian it is a strong distinction. That's one reason why when an Italian says something like, "Bello" ('pretty'), people often think the remark is exclamatory. But even said as a cool judgment, the Italian word *bello* simply starts with a fully voiced [b]—in contrast to the lightly voiced [b] of English.

The whole idea of having a feature to a certain degree (such as light voicing vs. full voicing) is disturbing. Most discussions of phonetics treat the features as being binary in value—a sound either has a feature or it doesn't. And in the rest of this book we will follow that convention. However, if you ever do your own research in phonetics and phonology, you might well have to pay attention to varying degrees of a feature.

We have now seen three of the major ways sounds can be distinguished. One is by their articulator and place of articulation. (So far we've discussed only bilabial sounds, with a passing remark on affricates.) Another is by whether the air flow goes into the nasal cavity (that is, whether or not the sound has the feature [+nasal]). And a third is by whether the vocal cords are vibrating (that is, whether or not the sound has the feature [+voiced]). So we have the following feature bundles for these three consonants:

- [p]: [+labial, -nasal, -voiced]
- [b]: [+labial, -nasal, +voiced]
- [m]: [+labial, +nasal, +voiced]

(These feature bundles are incomplete. As we learn more features, you can fill in their value—plus or minus—for these three consonants.)

Actually, we have seen a fourth way that sounds can be distinguished—by the feature of [±spread glottis]. In English this feature is not distinctive (as we argued above when discussing /p/), but in other languages it often is.

All right, we're ready to move on to additional features. Look back at our examples in 1–8. Is there any other group of phonemes which, like /m/, /p/, and /b/, have the same place of articulation and same articulator, but differ by nasality and/or voicing? Sure. Look at 7 again:

- [ŋ] [k] [g]
- [w]

This situation is entirely parallel to the bilabials. Here we have a nasal, a voiceless oral sound, a voiced oral sound, and the doubly articulated [w]. These sounds are called VELAR. The back of the tongue is the articulator, and, in fact, the only part of the tongue that we ever move toward the velum is the back of the tongue. Which of the velar sounds is the nasal?

((ŋ)) Which of the other two is [−voiced]? ([k]) Is either of them [+spread glottis]? ([k] is when in syllable-initial position—so in that position we write it as [k^h].)

If we want to distinguish the sounds in 1 (the bilabials) from the sounds in 7 (the velars) in terms of features, we can say that the former are [+labial], while the latter are [−labial]. This doesn't work beautifully, however: the problem is [w], for [w] is a LABIALIZED velar, so it is [+labial]. Therefore we need to add another feature, and we have already discussed it: the back of the tongue pulls toward the velum in making all the sounds in 7, so we say that the velars are [+BACK]. Thus the velars in the top row of 7 are distinguished from the bilabials by at least two features: they are [−labial] and they are [+back].

Before going on, let's linger a moment over [w]. Give a feature bundle for [w] with respect to the features we know so far:

[w]: [+labial, −nasal, +voiced, +back]

Is [w] ever aspirated? No. So it is also [−spread glottis]—but in English we would not list this feature, since it is not distinctive. As you continue meeting new features in this chapter, you should return to the consonants we have already discussed and enrich their feature bundles accordingly.

Okay, let's go on. When I asked you to find a group of sounds that seemed to contrast internally in the same way the sounds in 1 did, you might well have also wanted to list the examples in 4:

[n]	[t]	[d]
	[s]	[z]
		[l]

The problem here is that we have more than three sounds made with the same articulator and at the same place of articulation. The articulator is the tip of the tongue (or, for many speakers of English, the blade of the tongue). The place is, for English, back toward the bony alveolar ridge, and for the Romance languages and many of the languages of Europe, immediately behind and touching the teeth. These sounds are often called the DENTAL sounds, even though in English they are made on or just in front of the alveolar ridge. (Sounds made on the alveolar ridge are called ALVEOLAR sounds.) As with the velar sounds, the parts of the tongue that can be involved are automatically limited since the tip and blade of the tongue are the only parts we can bring to the area behind the teeth or to the alveolar ridge. (If you have been playing around with these sounds, you may realize that the remarks in this paragraph do not pertain as much to [l] as they do to the other five sounds. Please have patience. We will get to a discussion of [l].)

For each of the sounds in 4 now, give a feature bundle with respect to the features we have discussed above—that is, give a list with the values for the features labial, nasal, spread glottis, voiced, back. All of them are [−labial, −back]. Which is [+nasal]? ([n]) Which are [−voiced]? ([t] and [s]) Which is [+spread glottis] in syllable-initial position? ([t]—which is then written as [t^h]) But that is as far as we can go. We have no way yet to distinguish between the voiceless sounds [t] (in positions in which it is [−spread glottis]) and [s], and between the voiced sounds [d], [z], and [l], yet we know our language cares

about the distinctions between these sounds. So the dental sounds in 4 force us to find more features that distinguish between sounds.

Manner

Let's start by examining [t] and [s]. Compare *toe* to *sew*. Certainly we have the aspirated [t^h] here, while [s] is not aspirated. But there's more to it than that—and you know it, since *mat* and *mass* are distinctly different, yet both end in a consonant that is [–spread glottis]. Try to hold onto the initial consonant of *toe* and then of *sew*. That is, try to increase its duration or length. Can you? For *toe*, lengthening the [t^h] means having a longer silence before the sound is RELEASED (that is, before the closure formed by the tongue against the alveolar ridge is opened). For *sew*, lengthening the [s] means hissing for a longer period. The air flow is completely stopped in the oral cavity in making both [t] and [t^h], but the air flow is continuous in the oral cavity in making [s]. For this reason, we say that the MANNER of air flow differs between these sounds. We call /t/ (in both its allophones [t] and [t^h]) a STOP (or PLOSIVE) and we say that it has the feature [–CONTINUANT], while [s] has the feature [+continuant].

Let me warn you about terminology here. We say that sounds are [–continuant] if the air flow in the oral cavity is completely cut off at some point in making the sound. Thus even though the air may flow continuously through the nose on some sounds (as it does with [n]), a nasal sound will be [–continuant] if the air flow in the oral cavity stops. Not all linguists call nasals stops, however. So the terms *stop* and *noncontinuant* are not interchangeable. I will often refer to oral stops to distinguish them from nasals.

Look again at our bilabial sounds in 1 for a moment now:

[m] [p] [b]

Are they [+continuant] or [–continuant]? All three are [–continuant] because the air flow in the oral cavity is completely stopped. (Did this throw you off? Remember: while the air flow is never blocked in the nasal cavity for [m], the important question for the determination of this feature is what happens in the oral cavity.) What about our velar sounds in 7?

[ŋ] [k] [g]
[w]

Again, all these are [–continuant], except [w]. Since the lip closure is not complete for [w], the air flow in the oral cavity is never totally stopped.

Which of the sounds in 4 are [–continuant]?

[n] [t] [d]
[s] [z]
[l]

Yes, the top row. That's why I organized these sounds into rows.

Make a full list of the stops of English now and of the nasals:

stops: p, b; t, d; k, g

nasals: m, n, ŋ

Look again at 4, repeated above. Both the second and third row of 4 contain sounds that are all [+continuant]. We have already noticed that we can distinguish between [s] and [z] by the feature of voicing. But we are lacking a way to distinguish [z] from [l].

What differs between these two sounds? (If you do not make the [l] of *lie* with the tip or blade of the tongue touching the alveolar ridge, again, I ask for your patience. Many people do. So we are now on a hunt for a difference between [z] and [l] for those speakers. Bear with me and we will get to a more detailed discussion of [l], I promise.) We are going to get subjective here as we find another difference in manner of air flow, so loosen your shoulder muscles and let it happen. Hum a tune you like with the sound [z]. Now hum that same tune with the sound [l]. Which one sounds more melodic? Which one, in contrast, sounds more buzzy? Why is [z] buzzy? Make a [z] and hold it. What's happening to the air flow to make it so turbulent? It's being forced against the inside of the teeth and the oral cavity is being reduced because of the shape and placement of the tongue. Can you feel that? In contrast the air flow in making [l] is smooth and unobstructed. Yet the tip of the tongue is touching the roof of the mouth as you say the initial consonant of *lie*, just as it is in saying the dental stops. How is that air flow escaping? The air passes freely around the sides of the tongue. See?

We call [+continuant] sounds which involve air turbulence **FRICATIVES** or **SPIRANTS**. Look across your twenty-four consonants of English:

[m]	[p]	[b]
	[f]	[v]
	[θ]	[ð]
[n]	[t]	[d]
	[s]	[z]
		[l]
	[ʃ]	[ʒ]
	[tʃ]	[dʒ]
		[ɹ]
		[j]
[ŋ]	[k]	[g]
		[w]
	[h]	

Which of them are fricatives? You should be asking which of them you can hold (that is, which are [+continuant]) and have a hiss or buzz to them. Why are some hissy while others are buzzy? That difference is voicing. List now the fricatives in pairs that differ only by voicing. You should come up with four pairs:

fricative pairs: f, v; θ, ð; s, z; ʃ, ʒ

(If you listed [tʃ] and [dʒ], you were correctly noticing their hissy or buzzy quality. But neither of these sounds begins as a [+continuant]. We'll discuss them in detail later.) The last two pairs of fricatives in the list here are called the **SIBILANTS**. The sibilants are our most turbulent sounds. Compare the pair of [f] and [v] to [θ] and [ð]. Can you hear a

difference in turbulence between these pairs? You may well not be able to. I can't. Yet rules that affect sounds (PHONOLOGICAL RULES, which you will learn about in Chapter 2) often group [f] and [v] with the sibilants, treating them as though they are highly turbulent or STRIDENT. We make note of the difference here with the feature of STRIDENCY. We say that [f] and [v] and all the sibilants are [+STRIDENT], while the fricatives [θ] and [ð] are [-strident].

There is one more fricative of English that does not come in a pair: [h]. If you came up with it yourself, great. If you didn't, that's not surprising—it is a difficult sound to describe. I'm going to hold off on discussing it until after we've looked at the other consonants of English.

Laterals

Go back to [l] now. Earlier I treated [l] in the word *lie* as though it had the same place of articulation as [t], [d], [n], [s], and [z] in the appropriately minimally contrasting words. But for many speakers, this is not true. In saying the word *lie* there are many speakers whose tongue does not make contact with the top of the mouth. And for all speakers, we could easily find words containing /l/ that didn't involve this place of articulation. Describing the sounds associated with /l/ is more complicated than describing any other phoneme we have yet faced, because /l/ has allophones that can be remarkably different from one another. Say the word *little*. There are two instances of the phoneme /l/ in this word, each having a different PR (phonetic representation)—that is, each being a different allophone. The initial instance is like the initial consonant sound of *lie*, which is the allophone we have already considered. But the second instance is new. We call the first LIGHT OR CLEAR OR HIGH L (because the tip of the tongue is high), written as [l^h], and the second DARK OR VELARIZED OR LOW L (because the back of the tongue is bunched a little toward the velum and the tip of the tongue is low), written as [l^l]. What is the shape of your tongue as you make the two sounds? Certainly the blade of the tongue is fairly wide and flattened out in both, though not much else of the shape may be similar. But the crucial common feature is that the air flows freely around the sides of the tongue—this is the defining feature of “l”-like sounds, and we call it LATERAL. We say that both allophones are [+lateral]. Thus you should not feel terribly bad about admitting them as allophones of the single phoneme /l/ in English. Go through the list of English consonants. Are any of the others [+lateral]? No.

Dark and light *l* are not the only allophones of /l/ in English. Say the word *million*. Now try to isolate the INTERVOCALIC sound (the sound between the vowel sounds). Say it over and over. What is its place of articulation? Yes, the palate. Consider the shape of the tongue as you make this sound. It isn't terribly similar to the shape of the tongue as you made either of the other allophones of /l/, because now the blade of the tongue is high. But still the blade is wide and flattened. And can you tell that the air flow is still lateral? This palatal allophone of /l/ is written as [λ]. While we will be sure in our transcriptions of English to distinguish between [l] and [λ], we will lump light and dark /l/ together and transcribe them simply as [l]. There are two reasons why we will not be conscientious about distinguishing between light and dark allophones of /l/. First, it will ease transcription. But the second reason is the more important one: no language I know of makes a phonemic distinction between light and dark laterals, although many make a phonemic distinction between other laterals, such as the dental and the palatal. So it is important that

you note the distinction between dental and palatal laterals, but less important for you to note the distinction between light and dark laterals.

All of our various “l”-like sounds are lateral; that’s what unifies them. Nevertheless, [l] cannot be said to have a distinctive place of articulation, as we have just seen. We will say that [l] is **UNDERSPECIFIED** for place of articulation. Instead, its place of articulation varies according to the sounds we find preceding or following it. In fact, the phoneme we call “l” is likely to have allophones in any language that has this sound. This is going to become an important point when we talk about underspecification in the sections “Dissimilation” and “Underspecification Theory in Lexical Phonology” in Chapter 2.

When we first met the idea of allophones above, I told you that sounds which were allophones in one language can be separate phonemes in another. Let me demonstrate that point here. Italian has three lateral consonant sounds. One is a dental; one is a long dental (a dental of double duration, indicated with a colon after the consonant symbol); and one is a palatal. Now consider the following words of Italian, where I have written before each one the IPA symbol for the intervocalic lateral sound and I have written after each translation the IPA transcription for the whole word (with a vowel you will meet later):

[l]	pala	‘shovel’	[pala]
[l:]	palla	‘ball’	[pal:a]
[ʎ]	paglia	‘straw’	[paʎ:a]

These words are a **MINIMAL TRIPLET** (like a minimal pair), differing only in the lateral consonant. What does this say about these three lateral sounds in Italian? Are they separate phonemes or are they allophones of a single phoneme? If you don’t know the answer right off, reread the earlier paragraphs concerning allophones. Do you know now? The fact that the words above differ in sound by only the one **SEGMENT** and also differ in meaning (that is, they form a minimal triplet) is the key. I expect you to say that these three lateral sounds are three separate phonemes in Italian: /l/, /l:/, and /ʎ/. Actually, that isn’t exactly how linguists say it. Notice that the quality of /l/ and /l:/ is identical—they have all and only the same features. What differs is duration. So we say that Italian, unlike English, has distinctive length on its consonants (although not on its vowels).

The point of the above discussion of Italian is that the palatal lateral [ʎ] is an allophone of /l/ in English, but it is a separate phoneme from /l/ in Italian.

But another point has fallen out of that discussion, a fact about language that we could not have learned simply from looking at English alone: duration can make a sound distinctively different from another sound. We saw above that Italian has distinctive length on consonants. A colon after an IPA symbol means a lengthened segment of sound. Lengthened segments are called **GEMINATES**. Thus the consonant [l] differs distinctively from the lengthened or geminate consonant [l:] in Italian. Many languages, in fact, use duration to distinguish between sounds, particularly between vowel sounds. The Slavic language Czech has distinctive length on vowels, for example, and the Uralic-Yukaghir language Finnish of Finland uses duration distinctively for both consonants and vowels.

Inventory of Features Thus Far

At this point we can benefit from making an inventory of the new terminology we have been developing. First, we talked about articulators and places of articulation. We’ve

looked at three places of articulation thus far and discussed the sounds found there: bilabials, dentals or (in English) alveolars, and velars. In this discussion two features were relevant to the place of articulation and articulators:

labial, back

Second, we talked about whether the vocal cords were vibrating and whether they were spread to make aspiration. The relevant features were:

voiced, spread glottis

Third, we talked about the manner of the air flow, noting that we have oral stops, nasals, fricatives, and laterals. There were four relevant features here:

continuant, nasal, strident, lateral

Are you comfortable with that? If not, I suspect your worry is about calling nasality a manner of air flow. The other three features all deal with how the air moves in the oral cavity, but nasality deals only with the fact that air moves through the nasal cavity, not with how it moves in that cavity.

But think about that fact a moment. Certainly there is nothing analogous to the tongue inside the nasal cavity. And the tongue is one thing that allows us to stop the air flow to create sounds that are [–continuant] and to direct the air flow around it to create sounds that are [+lateral]. There is also nothing analogous to the teeth or lips. So we cannot direct the air flow against an obstruction to create sounds that are [+strident]. Therefore, saying a sound is [+nasal] says all that is relevant about its manner of air flow in the nasal cavity.

However, saying a sound is [+nasal] may not be all that is necessary to describe the manner of the air flow in its entirety. If the air flow were continuous in both the oral and nasal cavities at once for a given sound, we would need to say more about the manner of air flow for that sound. Does any such sound actually occur? In other words, do we ever get a nasal that does not correspond to an oral stop? Certainly not among the nasals we have discussed thus far. ([m] corresponds to the oral stop [b]; [n] corresponds to the oral stop [d]; and [ŋ] corresponds to the oral stop [g].) And in most languages of the world that have nasals, most of those nasals in most of their uses are [–continuant]. This is not, however, a necessity. We have already noted that vowels can be nasalized. (Look back at our discussion of *cat* vs. *can*.) And you will learn in Problem Set 1.3 that at least one of our English nasals has an allophone that does not correspond to an oral stop. Nevertheless it is an unusual or MARKED occurrence for a nasal consonant to be [+continuant]; the UNMARKED nasal consonants are all [–continuant]. In other words, generally, nasal consonants correspond to oral stops. And you may come across the term “nasal stops” in your readings.

Also, as you will see in Problem Set 1.3, a sound can be both [+nasal] and [+strident].

What about the features of nasal and lateral? Are they mutually incompatible? If we were to have a sound that was [+nasal, +lateral], that would mean that the velum was lowered and that the air that passed through the oral cavity went around the sides of the tongue. If laterals are [+continuant], then such a sound would also be [+continuant], so it would be marked (given that nasals are typically [–continuant]). But the question I want you to consider is whether such a consonant is possible at all. Try to nasalize the initial

sound of *lie*. See? You can do it. But it's certainly odd. I don't know if any languages of the world make use of nasalized laterals.

Since it came up, I will indulge in a brief aside here. In the previous paragraph, there is a sentence that begins, "If laterals are [+continuant] . . ." According to our definition of continuancy, laterals are continuant because the airflow is continuous around the sides of the tongue. However, in some languages laterals pattern with stops on phonological rules. That is, laterals and stops will behave the same way, when other sounds will behave differently. For this reason, some linguists have defined the notion of continuancy in such a way as to give laterals the value of [–continuant]. We have not done that, and we will not, in fact, look at any languages in which laterals pattern with stops. But you should be aware that this controversy exists, simply because it can help you to understand a very basic point: **while our physiological apparatus is the basis for our description of the sounds of language, the choice of which differences between sounds we will elevate to the status of distinctive features is based not solely on physiology, but on phonological behavior, too.** Without a doubt, as linguists study the sound systems of more and more languages, the features that we choose to list as distinctive features will change somewhat to accommodate the new findings and to accommodate the resulting new ways in which we look at already-familiar data.

Now go back to the set of 1–8. Again beginning at the front of the mouth, let's consider sounds whose articulator and place of articulation we haven't yet discussed. We come to 2:

[f] [v]

What is the place of articulation for these sounds? (top teeth) The articulator? (bottom lip). These are called the LABIODENTALS. Give a feature bundle for these fricatives. You should come up with [+labial, –back, –nasal, +continuant, +strident, –lateral, –spread glottis]. As for voicing, [f] is [–voiced] and [v] is [+voiced]. Since no other sounds we have discussed thus far have the same feature bundles, there is no need for us to distinguish them further at this point.

Now we arrive at the sounds in 3:

[θ] [ð]

Describe the place of articulation and the articulator. Not surprisingly, these are called the INTERDENTALS. Give a feature bundle for these fricatives. You should come up with: [–labial, –back, –nasal, +continuant, –strident, –lateral, –spread glottis]. Which of them is voiced? ([ð]). Again, these feature bundles are different from those for every other sound we've described thus far, so we can pass on without further details here.

We have already covered the sounds in 4 ([t], [d], [n], [s], [z], and [l]), so let us move on to 5. Consider the sounds in the first row of 5:

[ʃ] [ʒ]

([ʃ] is sometimes written as [š]) and [ʒ] is sometimes written as [ž].) Which of them is voiced? ([ʒ]) What is the articulator for these sounds? (the tip or blade of the tongue) What is the place of articulation? (Right behind the alveolar ridge.) These fricatives are [–labial, –back, –nasal, +continuant, +strident, –lateral, –spread glottis]. But this feature bundle is the same as that of another pair we have already discussed. Which other pair? That's

right—the other sibilants, [s] and [z]. So we need a feature to distinguish between the two pairs. And we already know that they differ by place of articulation. So we now introduce the new feature of [+ANTERIOR]: those sounds that are made with the primary place of articulation being at or in front of the alveolar ridge are [+anterior]; those behind are [−anterior].

Now consider the sounds in the next row of 5:

[tʃ] [dʒ]

([tʃ] is sometimes written as [t͡ʃ] and [dʒ] is sometimes written as [d͡ʒ].) Which of them is voiced? ([dʒ]) Are these really made at the same place of articulation as the postalveolar fricatives [ʃ] and [ʒ] in the first row of 5? Many linguistics books say these sounds are made on the hard palate—that is, further back than the postalveolar fricatives. But that isn't the full story. These sounds are complex in nature. Say them as slowly as you can. Can you see? They start out with a complete closure and then slowly release. The complete closure is on the alveolar ridge and along the sides of the palate. The DELAYED RELEASE (a term that we will now include in our features and abbreviate as DEL REL) means that the air is forced through a gradually enlarging passageway, so these sounds are definitely [+strident]. **Complex sounds that begin as a stop and end as a fricative are called AFFRICATES.** In English each of these two affricates behaves as though it is a single consonant phoneme with respect to phonological rules, and their complex nature is reflected in their IPA symbols ([tʃ] and [dʒ]). In other languages, however, they may better be analyzed as a sequence of two consonants, a stop and a fricative. We would then transcribe them as [t] followed by [ʃ] and [d] followed by [ʒ]. Their feature bundle as single consonants, not considering voicing, is: [−labial, −back, −nasal, −continuant, +strident, −lateral, −spread glottis, +del rel, −anterior]. Notice that even without introducing the feature of delayed release we could have distinguished these sounds by their feature bundles, since the fact that they are [−continuant] makes them different from the postalveolar fricatives. It's worth it to learn the feature of delayed release, however, since this one feature can pick out all affricates as a class.

“r”-sounds

We now come to the sounds in 6:

[ɹ] [j]

Let us first consider [ɹ]. No book on phonetics that I know of groups [ɹ] with the palatal sounds. Instead, they typically list [ɹ] with the alveolar sounds. Is the [ɹ] of *rye* an alveolar sound the way you say it? Not if you use the actual point at which the tongue touches the top of the mouth for your criterion. The sides of the tongue are raised against the inside of the top teeth on the sides of the hard palate. Perhaps the tip of your tongue is pointing upward, and if it is, it is pointing toward the alveolar ridge (and, hence, the classification as alveolar in many texts). Probably your lips are protruding. We know that /ɹ/ is a phoneme of English. Is every PR of /ɹ/ identical? Of course not. (You may have already realized that the PR of any phoneme can vary somewhat from its UR.) Compare the initial sound of *rye* to the final sound of *car*. Notice that your lips are not protruding on *car*. Yet I

know of no symbol to distinguish between these two allophones of /ɹ/. What that suggests is that linguists have not found it necessary to have an additional symbol here. Why? Probably because there is no language in which the two sounds differ distinctively. Can you understand the logic here? If not, reread this paragraph.

My description of the tongue, teeth, and lips when making [ɹ] may not have matched the description you would have given. Please do not get overly concerned about the differences. This sound, like the laterals, and much more so than the other sounds we have yet discussed, varies in its place of articulation based on the sounds around it, so much so that it can even be said not to really have its own distinct place of articulation. So again we are dealing with a phoneme whose place of articulation is underspecified. But one thing we should be able to agree upon is that the blade of the tongue is raised when we make [ɹ]. We can now meet another feature, CORONAL. We say a sound is [+coronal] if we move the tongue blade toward the teeth, alveolar ridge, or hard palate to make it.

List our [+coronal] consonants thus far. This feature is needed because certain phonological rules make use of it. When this feature was first proposed, the interdentalals were considered [−coronal]. Today, however, many linguists call the interdentalals [+coronal]. By this point you are familiar with this sort of controversy, and you can guess that linguists found the interdentalals behaving like coronal sounds in the phonological rules of some languages and they therefore changed their judgment of the value of this feature for these sounds.

But there is another point I'd like to make here: **Not all linguistics texts and articles use the features we discuss in this chapter to mean precisely the same thing.** So be sure to try to figure out exactly how an author is using the term as you follow that author's discussion. Let us return now to our consideration of "r"-like sounds.

Do you know any other language that has an "r"-sound that sounds like English [ɹ]? Be careful here. Don't be led astray by spelling. Many sounds that we identify as "r"-like occur in other languages, but they may well not be identical to [ɹ]. In Spanish, for example, we get a repeated TAP, called a TRILL, on the dental area. The symbol for this sound is [r], and it is truly a dental sound. Say the English word *pudding*. The intervocalic sound here is a single tap or FLAP, very similar to the repeated tap of the trill, although it can be made on the alveolar area; its symbol is [ɾ]. Can you tell that it is a little farther forward than both [d] and [t] in English? In French, on the other hand, the "r"-like sound we get is a voiced uvular fricative. The symbol for this sound is [ʁ].

Many different types of sounds go under the rubric of *r*, then, including trills, taps, a fricative, and the English [ɹ], which is often called an APPROXIMANT. (The term "approximant" covers continuant sounds that produce no audible friction. Many works in linguistics mark the laterals as approximants.) And these so-called "r"-sounds do have some things in common (although the French [ʁ] is an exception). They are [+voiced, +continuant, −strident]. They are also highly melodic. In this way they are like the nasals and the laterals, and unlike the oral stops, affricates, and fricatives.

We can see a major division among consonants here: The first set is the SONORANTS (nasals, laterals, approximants). The second set is the OBSTRUENTS (stops, affricates, fricatives). These features—sonorant and obstruent—are two sides of the same coin. The sonorants are our most vowel-like consonants and we will discuss them further after we have discussed the vowels. For now, just add sonorant to your list of distinctive features. (We could as easily have chosen obstruent for the feature. We chose sonorant because sonority will become an important concept in our discussions.)

The oral sonorant consonants, that is, the "r"-like sounds and the lateral sounds, are

called the LIQUIDS, although the term *liquid* is not typically used as a feature name. The liquids form a natural class in that they participate in phonological rules in a similar or identical way in many languages. (Remember that we'll meet phonological rules in Chapter 2.) In English you can feel relatively secure in grouping them together even from a purely phonetic perspective because they do not have a distinctive place of articulation, as already mentioned. Instead, [ɹ] and [l] vary in place of articulation quite a lot—much more than the stops, for example—according to context.

The second sound in 6 is [j]. Compare [ɹ] and [j]. Go back and forth between them. Their major difference is the shape of the tongue. For [j] the middle of the tongue is high, coming very very close to closure with the palate, yet not closing fully. We say that sounds that raise the blade or back of the tongue high up toward the roof of the mouth are [+HIGH] (which is the name of another feature).

Which other sounds that we've discussed thus far are [+high]? Actually, we are once more in a controversial area. Surely the affricates [tʃ] and [dʒ], and the postalveolar fricatives [ʃ] and [ʒ] are candidates worth considering. But the alveolar fricatives [s] and [z] are not, since even if you say them with the blade of the tongue, the tongue body is distinctly lower than when you say the postalveolar fricatives. You will find some linguists marking these affricates and the postalveolar fricatives as [+high], although others will mark them as [−high]. But everyone marks the alveolar fricatives as [−high].

The controversy over the highness of these affricates and postalveolar fricatives has two parts. First, it is clear that these sounds are made with the tongue blade (so they are [+coronal]), but does the tongue body rise in the process? And if it does, is this raising phonologically important? That is, do languages make use of this information when putting sounds together into larger units? For these affricates and postalveolar fricatives some languages seem to treat them as [+high], whereas other languages don't seem to offer such evidence. In Chapter 2 you will see many instances in which sounds change their features when preceded or followed (or both) by certain other sounds. When you get there, you can ask yourself whether or not the changes you consider offer evidence that a particular affricate or fricative is [+high]. (See in particular Problem Set 2.4.) In languages that lack relevant evidence about highness for these sounds, it is possible that such sounds are underspecified for this feature.

What about the velars with respect to highness? They are clearly [+high].

To return to [j], is [j] sonorant or obstruent? (sonorant) It is neither nasal nor liquid, however. Instead, [j] is an approximant (like the English [ɹ]). It is also often called a GLIDE. (This term is usually applied only to [j], [w], and [h], and we will make only minimal use of it.)

We are now at 7, the velars:

[ŋ]	[k]	[g]
		[w]

You should have no trouble, using all the features we have discussed thus far, giving a feature bundle for these sounds.

So we turn to [h]. Give a feature bundle for it. Is this sound back? No, because the back of the tongue is not bunched up toward the velum. If this sound has to fit into one of our groups of oral stops, nasals, fricatives, affricates, liquids, or approximants, which one does it seem most like? Certainly it is a continuant and oral, so the only real choices should be fricative, liquid, or approximant. But notice that the other liquids and approximants which we have discussed are voiced, while [h] is as close to silence as any of the sounds we've

discussed thus far gets. In fact, [h] is produced merely by spreading the glottis and letting the air flow out from the lungs (so it is [+spread glottis]), without any point of articulation or articulator in the oral cavity. So [h] is a fricative (though it's [-strident]).

Many books list [h] as [+sonorant]. Given how we've developed this chapter thus far, you probably think of fricatives as being [-sonorant]. There are at least two reasons for considering [h] a sonorant. First, acoustically, [h] appears to be equivalent to the voiceless counterpart of the vowel [ə]—a vowel we will discuss later. (If you have a phonetics laboratory available to you, you can demonstrate that for yourself.) Second, sonority is fundamentally a feature of sounds that have freedom of airflow in the vocal tract. Now the fricative sound of [h] is made in the larynx. If the larynx is an articulator of the vocal tract, then [h] should be considered [-sonorant]. But if the larynx is not an articulator of the vocal tract, then [h] should be considered [+sonorant]. It appears that languages vary as to whether their phonological rules treat the larynx as an articulator of the vocal tract or not. So [h] sometimes participates in phonological rules as though it is a sonorant consonant, and we therefore often find [h] listed as a glide. The most sensible guide then is how [h] behaves in the language that you are analyzing. You may mark it as [+sonorant] in one language and [-sonorant] in another.

In our discussion we have come up with **thirteen distinctive features for English: voiced, spread glottis, labial, back, nasal, anterior, continuant, strident, lateral, delayed release, sonorant (which is the opposite of obstruent), coronal, and high.** Of these, eleven were needed in order to distinguish one sound from another. But two of them, sonorant and coronal, were introduced for other reasons. There are no two sounds of English that are identical except for their value on the feature of sonority or for their value on the feature of coronality. There are other languages (as I will mention later for sonority), however, in which these features may be distinctive. And we will see that these two features turn out to be very useful when we write phonological rules.

These are all the consonant features we are going to discuss. Certainly there are more ways that we could classify consonant sounds if we wanted to. But the question for us is not how detailed we can get in describing sounds, but what the minimal amount of detail is that we need in order to handle whatever linguistic phenomena we want to look at.

Think about that point a moment. If we discussed, say, two dozen features of sounds, when only perhaps three were relevant to a given phenomenon, the twenty-one superfluous features would be in the way. They might tend to mask the real nature of the phenomenon. For that reason, linguists opt for the minimal set of features (or any other mechanism) necessary to handle a problem. The features we have discussed above turn out to be adequate for our purposes in discussing phonological rules, so we will not introduce new ones. But if you approached other sorts of problems in the study of language sounds, you might well want to consider additional features.

Beyond English

At this point we have discussed all the consonants of English. (See the Appendix at the end of this chapter, which lists all these consonants with their place and manner as well as with their distinctive feature values.) But we haven't discussed the full range of consonants found in the languages of the world. Among the things we've left out are many fricatives and some oral stops and nasals, particularly in the rear parts of the vocal tract. Fricatives can be made at every possible place of articulation, so let's start at the front of the mouth and fill in those that are missing.

Put your lips together lightly. Now force air through them. If you let a pocket of air form behind the top lip, you will wind up making a trill. This bilabial trill, [ʙ], is a sound you might make if you were fed up or disgusted with something. But it is also a consonant in a very few languages, such as Kele (also spelled “Gele”) and Titan, Indo-Oceanic languages spoken in the Admiralty Islands of Papua New Guinea.

If, instead of forming an air pocket behind the top lip, you keep your lips taut as you force the air between the lips, you will make a bilabial fricative, and you can choose to voice it or not. [ɸ] is voiceless; [β] is voiced. Bilabial fricatives are commonly found: Japanese (a language isolate in the Altaic phylum, spoken in Japan) makes use of the voiceless one; Spanish (another Romance language, spoken in Spain), of the voiced one.

Now make an alveolar [l]. Then try forcing the air very hard around that lateral. You’ll come up with a lateral fricative, which you can voice ([ɬ]) or not. (To indicate lack of voicing on a normally voiced sound, you can put a diacritic—a small circle—under the sound.) Welsh (a Celtic language of the Indo-European family spoken in Wales) and Zulu (a Bantu language spoken in South Africa) both use lateral fricatives, as do many languages. The approximant alveolar lateral is [+sonorant], while the fricative lateral is [−sonorant]. This is the first instance we’ve seen in which the feature of sonority can be the distinguishing feature in a minimal pair.

Now put the tip of your tongue at the front of your hard palate (but well behind the alveolar ridge). You will have to curl it backward to do this. Make an oral stop. If it’s voiceless, it will sound like a strangely muffled [t]. (In fact, this sound is sometimes considered a speech impediment for speakers of English.) Recall that sounds made with the tongue curled back like this are called retroflexed. ([ɮ] is the voiceless retroflexed stop; [ɮ̌] is the voiced one). Make a retroflexed nasal now. ([ɮ̃]) Retroflexed stops and nasals occur in many of the major languages of India, including Hindi-Urdu (an Indo-Iranian language spoken also in Pakistan) and Malayalam (a Dravidian language). You can also make retroflexed fricatives by lowering the tongue so the closure is incomplete. Do it. The symbols for them are [ʂ] and [ʐ]—and Vietnamese (an Austro-Asiatic language spoken in Vietnam) and Malayalam make use of them, as do many other languages. You can make both retroflexed liquids there. Try it. (The “r”-like one, [ɻ], will be a tap or flap, not a trill, right?) The symbol for the lateral is [ɭ].

Now put the blade (not tip) of your tongue to your palate. Make an oral stop. You can voice it ([ɟ]), or not ([tʃ]). This is, of course, where the palatal lateral ([ɭ]), and the approximant ([j]) we have already discussed, are made. A palatal nasal ([ɲ]) is also made here, and you will learn about it in Problem Set 1.3. Italian makes use of the palatal nasal and liquid; Hungarian (a Uralic-Yukaghir language spoken in Hungary) makes use of both palatal stops. Now make a fricative, simply by lowering the tongue the slightest bit and forcing the air through the reduced passageway. If you’re having trouble, try saying the word *huge*. Most speakers start this word with a palatal voiceless fricative. Practice voicing ([j]) and devoicing ([ç]) the fricative. German (an Indo-European language of the Germanic family) uses the voiceless palatal fricative. Greek (an Indo-Hittite language spoken in Greece) uses both voiced and voiceless palatal fricatives.

Now make a [k]. Then don’t complete the closure and you’ll be making a voiceless velar fricative, like at the end of the word *Bach* under the German pronunciation. The symbol for this fricative is [x]. Say it and then voice it ([ɣ]). Vietnamese and Greek use both velar fricatives. You can also make a lateral there ([ɭ]). Try it. Some languages spoken in Papua New Guinea use velar laterals, but they are rare. And you can make an approximant similar to the lateral and similar to [j], written with the symbol [ɰ].

You can make fricatives behind the velum in the uvular area, but in order to do it, it

helps to know how to make non-nasal stops back there. Make a [k] between [ɑ] sounds. ([ɑ] represents the vowels in *mamma*.) So say [ɑkɑ]. Now try to swallow it a bit, so the closure is in the uvular area. That's the voiceless uvular stop [q] (which the western Semitic language Arabic makes use of). (If you have trouble making this, then wait until I've discussed ejectives and come back to it then.) Now make the voiced uvular stop [g]. Once you can make those stops, make the uvular nasal [ŋ]. All three of these consonants occur in Inuit (an Eskimo-Aleut language spoken in Alaska and Canada), Quechua (an Amerind language spoken in the Andes), and some other languages native to the Americas, which are in the Amerind family. Okay, now raise the velum (so it's no longer a nasal sound) and try easing up on the closure so that you get the voiced ([ɣ]) and voiceless ([χ]) uvular fricatives. (Recall that the voiced one is the French "r"-sound.) And now sort of gargle a bit while you voice and you will be making the uvular trill [ʀ]. Amazing, huh?

Now go even deeper in your throat. You won't be able to make stops in the pharyngeal area, but you can make fricatives there. Make a sound like you're clearing your throat. Now voice it. You've just made the pharyngeal fricatives [ħ] ([−voiced]) and [ʕ]. Fricatives are the only types of sounds that are made in the pharynx, and both Classical and modern varieties of Arabic use these fricatives, as do Hebrew (a Semitic language spoken in Israel) and some of the languages of the Caucasus Mountains.

You already know how to make a glottal fricative—that's what [h] is. Now make an [h], but add voicing, to get [ɦ]. You should sound like a dying person or a monster. You can also make a stop here. Make an extended [ɑ] and stop the air in the middle of it at the glottis. The glottal stop [ʔ] is always voiceless and you can hear it in the medial position of American English and British English (English spoken by native speakers in Great Britain) pronunciations of words like *button* and the exclamation *uh-oh*. You can also hear it when teachers read words off a list in a spelling test; they will often put a glottal stop before a word that begins with a vowel, such as *apple*, so that the word seems to get a jumpstart. The feature [+CONSTRUCTED GLOTTIS] is used for any sound that involves complete closure of the vocal cords, such as [ʔ] (and we will meet more such sounds later).

Some languages use other sounds made deep in the vocal tract, called EPIGLOTTAL sounds. Without a tape recording and/or a teacher to help you, I doubt you can make these sounds on your own. The symbol for the voiceless fricative is [ħ]; for the voiced fricative is [ʕ]; and for the voiced stop is [ʔ]. The Semitic languages (a family in the Afro-Asiatic phylum) are the main users of these sounds.

Furthermore, some languages make a double articulation of [ʃ] and [x] written as [ʃx]. This one I think you can produce on your own; try it. You are already familiar with double articulations from having looked at affricates and at the labialized velar [w] (which, by the way, is another approximant, like [ɹ], [j], and [ɰ]). You will meet more doubly articulated sounds later in this chapter.

Where I have failed to give an example of a language that uses a particular sound, this gap is due to my lack of knowledge. But the IPA introduces all these symbols, so I assume that there must be languages that make use of all these sounds.

Ejectives, Implosives, and Clicks

Languages also have other stop consonants which are made in the same areas as the ones we've discussed, but for which other aspects of the air flow vary. One such type of sound is called an EJECTIVE. The stops we've looked at thus far all depend on PULMONIC air

pressure—that is, air pressure from the lungs. To make an ejective, the glottis closes so that the air pressure builds up behind. Then you make a momentary closure somewhere else in the vocal tract and release it along with the glottis, thus getting GLOTTALIC pressure. The sound sort of pops from your mouth. All ejective sounds are, by definition, [+constricted glottis].

Let's try to make some ejectives. Perhaps the easiest one to start with is an ejective velar, written as [k']. (An ejective air pressure is indicated by an apostrophe after the regular IPA symbol.) Put it between [ɑ]s again and go back and forth between the pulmonic and the glottalic stop, like so: [ɑkɑ] [ɑk'ɑ]. Now try it for the palatal, retroflexed, alveolar, and bilabial stops. All of them are voiceless. You can also make a uvular ejective. It sounds very much like the "glug" sound from an emptying bottle. (And if you had trouble making uvular stops earlier, now make this ejective and then try opening the glottis and doing a regular pulmonic uvular stop.) Ejectives are very common. They are found in many Amerind languages (such as Lakhota) and Athapascan languages (such as Navajo), in several African languages (such as Hausa, a Chadic language of Nigeria, and Swahili, a central Bantu language), and in some languages of the Caucasus region (such as Georgian).

You can also close the glottis, and when you release it, suck air into the lungs instead of letting it go out the vocal tract. That is, the AIRSTREAM is INGRESSIVE rather than EGRESSIVE. Sounds made in this way are inverse ejectives and are typically called IMPLOSIVES. Like the ejectives, they are typically [+constricted glottis] (although they do not have to be, as mentioned in the next paragraph). Unlike the ejectives, both voiced and voiceless implosives are found in language, although voiceless implosives are rare. Again, it may take some practice to learn how to make these sounds, and, strangely enough, you might have an easier time doing it if you start in the uvular area. Make a uvular ejective (the bottle "glug"). Then make a uvular implosive and go back and forth between them as quickly as you can, ejecting and sucking, in alternation. Once you get the hang of it, you can make implosives at all the other stop points, although retroflexed implosives do not have symbols in the IPA, so I assume they do not occur in language. An implosive air pressure is indicated by a rightward curved hook on the upper left or right portion of the regular IPA symbol, so the voiced bilabial implosive, for example, has the symbol [ɓ] and the voiced uvular implosive has the symbol [ɠ]. Zulu makes use of both ejective and implosive bilabials, as do many other languages of Africa and some languages of the Americas and of India, such as Sindhi (an Indo-Iranian language).

By the way, it is perfectly possible to have ingressive pulmonic air pressure, but it is rare. The Penutian language Maidu (of northern California) is the only language I have found listed as making use of it.

Finally, there is one more thing we can do with air that doesn't involve the lungs or the glottis. We can form a complete closure somewhere in the oral cavity and then move the articulator to produce a suction. The release of this suction is called a CLICK. Sounds made with this sort of suction are called VELARIC pressure sounds (as opposed to pulmonic or glottalic pressure sounds). You already know how to make a bilabial click: it's identical to the kiss sound and it's written as [ɔ̤]. You are also familiar with the affricate click that touches both the alveolar and palatal areas and sounds like a scolding noise or a noise that means, "What a shame!" It is written as [ʄ]. You can move that click forward to behind the dental area and make [ɭ], also written as [ɮ]. And you are familiar with the alveolar lateral that sounds like encouragement to a horse to speed along, written as [ɬ]. (This is made by opening only one side of the mouth and clicking that same side of the

tongue.) You can also easily make an alveopalatal click that sounds very much like a cork popping from a bottle, [≠]. Make all of them now. You can find clicks in many languages used as interjectional noises. But they are common as phonemes primarily in some languages of southern Africa: Zulu, Xhosa, Nama, Bushman, and Hottentot use a variety of clicks.

The clicks you've just made (with the exception of the bilabial and lateral ones) involve pulling the tongue down and back to rarefy the air behind the closure. But you could also push the tongue upward and compress the imprisoned air to make an inverse click. Try going back and forth with the alveolar affricate click, regular then inverse then regular and so on.

Breathy Voice, Whispery Voice, and Creaky Voice

At this point you have been introduced to all the consonants listed in the IPA. However, there are other things you can do to these consonants. Compare the pair *bloom* and *plume*. Focus on the lateral sound. Say the words quickly, then slowly. What's the difference between the two laterals? Certainly that in *plume* has a more airy quality to it. Can you hear it? Try it again. It is very difficult for an English speaker who is not trained in phonetics to detect this difference. Try getting ready to say *bloom*, but then say it without the [b]. Then do the same for *plume*, without the [p]. If you really have yourself all prepared to make the [b] first, and then all prepared to make the [p] first, you should be able to notice that the lateral starts differently in each. That's because the lateral in *bloom* is voiced, while that in *plume* is voiceless. The IPA marks voiceless sounds that are typically voiced with a small circle under the regular IPA symbol, such as [l̥]. (Another way to mark them is to use the corresponding capital letters, for example: [L].) For those of you who are still having difficulty hearing the difference, don't despair—you're not alone. A more noticeable [l̥] is after an [s] in an unstressed syllable, as in *whistling* [wɪs̺lɪŋ]. Can you hear it now? Once you learn to hear that one, then you should be able to hear the [l̥] of *plume*, too.

Can you guess why this lateral is voiceless in *plume* but voiced in *bloom*? If you don't know how to begin guessing, then say many words with a lateral and put them in two lists according to whether they are voiced or not. You might come up with something like:

Voiced	Voiceless
lose	slow
alter	claim
pail	please
glue	flavor
pearl	schlep (from Yiddish)

(Yiddish is a High Germanic language spoken by Jews of eastern Europe.) Now can you see? The voiced [l] occurs in many phonological environments, but the voiceless one in English occurs only as part of a consonant cluster where the first consonant is voiceless. What does this tell us about the status of the voiceless one? Is it a phoneme of English or is it an allophone of /l/? It's an allophone (and now we have met yet one more allophone of /l/). But in another language it might be a separate phoneme.

Now that you know how to make and look for voiceless laterals, find instances in English of voiceless “r”-like sounds and voiceless nasals.

The difference between voiceless and voiced sounds is a difference in PHONATION. For voiceless sounds our vocal cords are generally separated and the air passes between them without causing vibration. For voiced sounds our vocal cords are generally close together and the airflow causes them to vibrate.

We can also separate the glottis and use a substantial enough airflow to make the vocal cords vibrate a bit, despite their separation. Say [h] again. Now keep saying it, but increase the airflow as much as you can. You have now made a BREATHY or MURMURED [h]. And you can add this breathy quality on top of your ordinary speech—that is, to all consonants and vowels. Some linguists call this Marilyn Monroe talk, and it is indicated by two horizontally arranged dots under the regular IPA symbol. Breathiness on a vowel is in a sense analogous to aspiration on a voiceless oral stop. The language Kuy, a minority language of the Mon-Khmer family spoken in eastern Thailand and some areas of Cambodia and southern Laos, uses breathy vowels, which contrast distinctively with nonbreathy vowels.

We can also add a WHISPERY quality by removing all voicing. This is something you are probably quite used to doing. Many languages have voiceless vowels (such as French, Japanese, and Comanche, an Uto-Aztecan language in the Amerind phylum of the western United States), but typically they are allophones of voiced vowels, rather than phonemes.

And we can speak in a CREAKY voice, as if we were making the low sound of a rusty hinge on a door in telling a horror story. (This is indicated by the diacritic tilde, ~, under the regular IPA symbol.) Here all sounds involve vibration of the glottis, but only one end of the vocal cords can vibrate. When we discuss formants, we will talk more about exactly what’s happening in the larynx as we make a creaky voice. For now, it is enough that you can make a creaky voice and detect one when you hear it.

Try taking a word with all voiced sounds in it and imposing breath on it. Now try giving a whispery quality to the word. Be careful not to totally whisper, because then all the sounds will be voiceless. There’s a difference. And now say it in a creaky voice.

Another common way consonants in various languages can be different from those we’ve already discussed is in the addition of a secondary articulation. You are already familiar with the idea of secondary articulation from our discussion of [w]. We called the labialization here secondary. Take [k] and [g] and give them a secondary labialization. You come out with the initial sound of words like *queen* and *guana*. Labialization is indicated with a raised w: [k^w], [g^w]. Try labializing other consonants. Do you find any that seem to occur in English? Certainly labialized alveolar stops occur in English. Find examples. (the initial consonants in *twin*, *Dwight*, and the like) Typically it is stops that are labialized, but other sounds may be, as well.

Consonants can also have a secondary articulation in the palatal area (Russian makes use of palatalized stops), in the velar area, and in the pharyngeal area. Again, this typically happens with oral stops. And the release of such a stop can be accompanied by a quick lowering of the velum (to give a nasal release) or by a widening and flattening of the tongue (to give a lateral release). (Here you may need an instructor’s demonstration to help you make the sounds.)

This is not the whole story. Other fine distinctions can be made. The point is not for you to memorize what goes on in language, but for you to be sensitized to the sorts of things that can happen, so that you can listen for them.

Minimal Pairs and Vowels

Just as orthographic consonants in English do not have a one-to-one relationship with consonant sounds, the same is true for vowels. So please fight off tendencies to revert to any short list of vowels your grade school teacher may have given you.

With that caveat in mind, let's begin by trying to understand what a vowel is and how a vowel differs from a consonant. We have already introduced the vowel [a] (as in *mamma*). Say this [a] and compare it to the consonant sounds. What do you notice about it? Which consonant sounds does it seem to have the most in common with? Certainly [a] is a continuant and it's voiced—and we will find that all our vowels are (by definition) continuants, and voiced in the unmarked situation. Yet the vowels do differ from the consonants. With respect to articulatory phonetics, the vowels are made with no obstruction in the vocal tract whereas the consonants have varying degrees of obstruction, from very little (with the approximants and liquids) to complete (with the stops). Still, you may have trouble swallowing the claim that the first vowel in the word *ear*, for example, involves any greater or smaller obstruction than the obstruction made by the liquid at the end of that same word.

Fortunately, we don't have to rely solely on articulatory phonetics to see the differences between vowels and consonants. Vowels differ quite remarkably from consonants in ways we can see both from acoustic phonetics (which we will talk about later) and from phonological rules. That is, the vowels play different sorts of roles from the consonants when we put sounds together to make words. That's what we'll focus on first.

To see this, compare [a] to [z], for example, which is also [+continuant, +voiced]. (Notice that in this discussion the sounds are consistently between square brackets. That is because all we care about is how they sound as we say them, not what they might have been in underlying representation.) Say monosyllabic words that contain [a] and monosyllabic words that contain [z]. What do you notice about them? Say a monosyllabic word that contains both [a] and [z], such as *Oz* [ɔz] (as in the wizard's land) or *shah's* [ʃaz] (as in having to do with the shah of Iran). Maybe you're starting to get an inkling of the difference. Try comparing [a] to another voiced continuant, such as [v]. Look at the word *von* [vɔn] as in the last name *von Beethoven*. Now make up monosyllabic nonsense words that contain three sound segments, choosing only from the list of [z], [v], and [a]. You have undoubtedly made up [zaz], [zav], [vaz], [vav], and [avz]. That is, you can have two consonants, but only one [a]. What happens if you put two instances of [a] into your nonsense word? English does not form minimal pairs in which one word has a short vowel (a single vowel) and the other has a long vowel (a sequence of two identical vowels.) (But some languages, do, as we'll discuss later.) So if we're going to make a nonsense English word with two instances of [a], those instances will be in separate syllables. The result in English is a polysyllabic word such as [aza] or [ava].

Syllables

Give me any two-syllable word of English. You might have given *father*, or thousands of others. Give me any three-syllable word of English. You might have given *unlikely*, or thousands of others. Give me any four-syllable word of English. You might have given *American*, or thousands of others. How are you choosing which words have three syllables

and which have four? What are you counting? In musical terms, you might say you are counting beats. Each syllable consists of a single beat.

And now you have found one of the major differences between consonants and vowels: in the unmarked case vowels are crucial to the beat of a syllable. We say they form the NUCLEUS of the syllable and that they are [+SYLLABIC]—they carry the beat. And in the unmarked case consonants are [−syllabic] and they appear either preceding the nucleus of the syllable, in the ONSET, or following the nucleus, in the CODA. Thus in a syllable like [van], the onset consists of [v], the nucleus consists of [a], and the coda consists of [n]. Is it possible for a syllable to have no coda in English? Certainly. Think of the word *pa* [pʰa]. Is it possible for a syllable to have no onset in English? Again, of course. Think of the word *Oz* [ʌz]. But every syllable must have a nucleus, by definition, since every syllable has a beat.

Syllabic Consonants

Now while consonants typically are not SYLLABIC (that is, they typically don't carry the beat of a syllable) and vowels typically are syllabic, we find that there are some syllabic uses of certain consonants in English as well as some nonsyllabic uses of certain vowels. The vowel [a] is always syllabic, so we'll hold off on a discussion of nonsyllabic vowels for a moment. But we can now discuss syllabic uses of consonants. Do you remember which consonants we claimed above were the most vowel-like? The sonorant consonants. Consider first the liquids. Can you think of a word that contains a syllable that has no vowel, but instead has a syllabic liquid? Let me give you one for [l]: *bottle* [ˈbɑ.rəl]. (The diacritic of a small vertical line under a consonant shows that this consonant is [+syllabic]. So [l] is nonsyllabic and [ɫ] is syllabic. A period indicates a syllable boundary. And an acute accent, ´, over a vowel, or, in some books, a raised ´ before a syllable, indicates that that vowel (or syllable) receives PRIMARY WORD STRESS. By convention when transcribing monosyllabic words, we leave out the stress mark, even though the word may be fully stressed.) Now you can easily find others. What about a syllabic use of [ɹ]? Consider the word *work* [wɜɹk]. Again, you can now easily find others.

Turn now to the nasals. Can you find a syllabic use of any of them? Examples with the bilabial and alveolar nasals are easy to come by (such as the minimal pair *prism* [pʰɪ.ɹ̥z̩m], *prison* [pʰɪ.ɹ̥z̩n]), both of which use a vowel in the first syllable that we haven't yet discussed). But for some speakers the velar nasal can also be syllabic, as in the first syllable of the word *ingot* [ɪŋ.gət] (which uses a vowel in its second syllable that we haven't yet discussed).

The only other sonorant consonants are the glides, [h], [w], and [j], the last two of which are often called the SEMIVOWELS. No glides can be syllabic in English, however, a point we return to later.

The liquids and nasals—that is, the sonorants minus the glides—are the only consonants that can be syllabic in English. However, other languages may make use of no syllabic consonants or of a range of other syllabic consonants. The Imdlawn Tashlbiyt dialect of Berber, an Afro-Asiatic language family spoken in various parts of northern Africa west of Egypt, particularly in Morocco, demonstrates one extreme case, for in this language any consonant (as well as any vowel) can be syllabic, whether it is sonorant or obstruent. This is a highly marked language for syllabicity, of course.

Sonority

Let us take a moment now, before we go on with our discussion of the vowels, to look a little more closely at the feature of sonority and its role in the shape of the syllable. Consider monosyllabic words in English that contain clusters of two consonants in either onset or coda, or both. List them. Start with the bilabial consonants and ask what clusters they can initiate. Then go to the labiodentals, and so on, into the mouth. Your list should begin with such examples as:

onset clusters

[p ^h l], [p ^h ɹ]	please, proud
[bl], [bɹ]	bloom, bread
[fl], [fɹ]	fly, fry
[θɹ]	thread

Considering only the contrast of obstruent vs. sonorant, what can you say about the ordering of these two types of sounds within an onset cluster? Obstruents precede sonorants. Now consider coda clusters. Make another list. It should contain examples like:

coda clusters

[lp], [ɹp]	help, burp
[lb], [ɹb]	bulb, suburb
[lf], [ɹf]	elf, smurf
[lθ], [ɹθ]	health, earth

Once more you can see a contrast between obstruents and sonorants, but this time the sonorant comes first. Can you put these two statements of the distribution of sonorants vs. obstruents together into one generalization? Sonorants occur closer to the nucleus of the syllable; obstruents occur at the periphery of the syllable.

In fact, there is much more that we could say. Sometimes languages allow clusters containing more than one sonorant and/or more than one obstruent, or a sonorant plus an obstruent. English does this. And there are rules regulating the order of different consonants within a cluster. It has been proposed that there is a SONORITY HIERARCHY among sound segments:

vowels, glides, liquids, nasals, fricatives, affricates, and oral stops

In fact, within the fricatives and affricate/stop groups, the voiced segments are more sonorous than the voiceless segments. When languages form syllables, they usually arrange sounds within a syllable so that those closer to the nucleus are more sonorant. English presents exceptions to this generalization that you will explore a bit in Problem Set 1.4. English is not alone in being exceptional. Thus this generalization should be considered a kind of metric for evaluating the naturalness of a syllable, rather than as a principle for creating syllables. In fact, we will see in the section "Syllable Types" in Chapter 2 that the sonority hierarchy can be helpful in understanding how languages syllabify words.

Okay, at this point I hope you feel comfortable with our treating the vowels as different sorts of animals from consonants. And if you don't, then perhaps after we have discussed acoustics a bit later, you will. But for now let's turn to an articulatory phonetic description of the vowels of English.

Vowels and Diphthongs

Think about the features we've learned for the consonants. Are any of them relevant to [a]? Certainly. [a] is voiced; it doesn't use the lips in any way; the tongue is not bunched in the back of the mouth; it is not nasal; it is continuant; it is not strident; it is not lateral; it has no aspiration so it is not spread glottis; it can't have a delayed release because it has no closure to start with; the tongue body is not high in the mouth; it is a sonorant sound; and the tip or blade of the tongue is not raised in any way so it's not coronal. In other words, all of the features we have talked about for consonants could easily be used in describing vowels. Nevertheless, only some of the features used to talk about consonants are used to talk about vowels, and other features particular to vowels will need to be introduced. That's because there is little point in saying a vowel is [+continuant] or [+sonorant], for example, if all vowels have those features. (But if we had cause to gather together all the vowels and all the continuant consonants with a single feature, for example, we could use [+continuant].) Furthermore, what distinguishes vowels from one another more than anything else is the shape and size of the RESONATING CHAMBER—that is, the oral cavity. Vowels are resonant sounds made with much less closure in the oral cavity than most consonants, and this is what we will focus on now.

Say the baby-talk words *meeme*, *mama*, and *moomoo* (as in a child calling attention to self, calling to the mother, and mimicking the noise of a cow). Try to make the vowel sound in each syllable as short and clipped as you can. (In musical terms, give it a staccato quality.) With a short vowel in each one, we transcribe these three words into IPA like so:

[mí.mi] [má.ma] [mú.mu]

The three vowels we are hearing are [i], [a], and [u].

Probably you did not make a short vowel for two of these three words until I asked you to. Go back to your original pronunciation of *meeme* and *moomoo*. Say them several times. Then take off the initial nasal consonant and say the vowels in isolation. Contrast these longer vowels to the short staccato vowels I asked you to produce. What are the differences? One difference, of course, is duration. The short vowel takes up one TIMING SLOT (or unit of time), while the longer one takes up two timing slots. But there is another difference. Can you hear that the quality of the longer vowels changes? I asked you to pronounce them with staccato style so that you could isolate a single vowel sound here, without the vowel change at the end.

Say the words *beat* and *bit*. Now say them with staccato vowels. You have now met a fourth vowel, the vowel in *bit*: [ɪ] (often written as a script "i," and called *iota*). We transcribe *bit* as [bɪt]. When you say *beat* staccato style, with only one timing slot for the vowel, we have [bɪt]. When many of you say *beat* in the normal style, with two timing slots for the vowels, we have [bi:t]. (If you are having trouble saying and hearing [bɪt] in

contrast to [bɪɪt] and my directions to use staccato style don't help, try starting the word and holding the initial vowel without ending the word. Can you hear and produce the difference now?)

For many speakers of American English the normal pronunciation of *beat* involves a DIPHTHONG, a combination of a syllabic vowel and a nonsyllabic vowel. Which part of the diphthong in [bɪɪt] is the [+syllabic] part? The first vowel. We will indicate that the second vowel is nonsyllabic with the diacritic _̚ below the vowel, like so: [bɪɪ̚t]. That is, just as syllabic consonants are marked with a special diacritic, so nonsyllabic vowels are marked with a different special diacritic—because both situations are the less-typical case. The [–syllabic] part of a diphthong is often called a glide. In fact, you will find linguistics articles that transcribe the second part of this diphthong with the glide [j]. However, you know that this part of a diphthong is different from [j]. (If you don't remember, go back to our discussion early in this chapter of the final sound in *my* in contrast to the initial sound in *yes*.) The final sound in *my* is a [–syllabic] instance of the vowel [ɪ]; but the initial sound in *yes* is the consonant [j]. Transcribe *my*. You can do it. ([maɪ̚]) (We will return to thorny issues of glides and diphthongs in Chapter 2, particularly in Problem Set 2.2. So if you are full of questions now, please hold onto them.)

A similar situation holds for the vowels in *moo moo*. Say the words *cooed* (as in what a pigeon does) and *could*. If you say them normally, the first one has two timing slots for the vowel sounds, but the second one has only one timing slot. Say the first one staccato. You come up with [kʰʊd]. This contrasts with the second one: [kʰʊd]. Now say *cooed* in your normal style (not staccato). Can you hear the diphthong? Many American English speakers have a diphthong here, which would give us the transcription [kʰʊɪ̚d].

Not all American English speakers, however, have a diphthong in *beat* or *cooed*, and even for those who do, the diphthongization is slight and hard to hear, because the two vowels in each of these diphthongs are so similar to each other. ([i] is only slightly higher than [ɪ], and [u] is only slightly higher than [ʊ]—which is also often written as closed omega [ɔ̞].) Some speakers instead have a long vowel in each of these words. (And there are some words that perhaps all speakers have only a long vowel in, such as *cease* [si:s].) If you are among the speakers who say [bi:t] rather than [bɪɪ̚t] and [kʰu:d] rather than [kʰʊɪ̚d], then go to a movie or turn on the radio and listen closely. You will almost assuredly be able to hear the diphthongs in words like this within a few moments if you are looking for them, because this diphthongization, while delicate, is prevalent in many varieties of English.

One way to think of diphthongs is to consider them vowels in motion. You start out with one vowel and move into another vowel. In the two examples with diphthongs that we have discussed thus far, the nonsyllabic part is second. This is called an OFF diphthong or glide (or, depending on your terminology, a FALLING diphthong or glide). You can also have the nonsyllabic part come first in an ON diphthong (or a RISING diphthong), as in the word *piano* ([pʰɪæno]). And you can have TRIPHTHONGS in which nonsyllabic vowels surround the syllabic vowel. English has some triphthongs, as in *cute*. Can you hear the rising and falling glide on either side of the [u]? Transcribe *cute*. ([kʰɪuɪ̚t])

English is a difficult language to get started talking about vowels in because English makes such frequent use of diphthongs. But now that you know what a diphthong is, you can be sensitive to them and you can revert to staccato speech to avoid them. So we're ready to face vowels head on now.

Front, Back, High, Low

Let's contrast [i] to [u] to [ɑ].

For [i], describe the shape and position of the tongue. Is it high in the mouth (that is, close to the roof) or not? Yes, it's as high as vowels get in that part of the mouth. What do I mean by "that part of the mouth"? The palate. So [i] is often called a palatal vowel. Now consider the bunched up part of the tongue (the part that comes up toward the palate) and contrast that position to the position of the tongue in [u] and [ɑ]. Would you say the bunching for [i] is in the front of the mouth or the back in contrast to [u]? Certainly the bunching is not very far forward at all, but [i] is definitely in the front when compared with [u], right? [i] is called our **FRONT**, high vowel. And we have now met two features that are relevant to vowels: front and high. Front is not to be confused with anterior for consonants. Anterior is a feature of sounds made at or in front of the alveolar ridge. But the vowels need a resonating chamber—so they can't be made too far forward in the mouth. The hard palate is as far forward as the vowels get. Likewise, high for vowels is not as simple as the feature high for consonants. With consonants, we mark them as + or – for highness and that is the only feature we have for tongue height. That's because consonants call for some sort of partial or total obstruction of the airflow somewhere in the mouth, so if the tongue is crucially involved in making a given sound, it won't be terribly low in the mouth. But for vowels, the important part is the shape and size of the resonating chamber. So a very low tongue position will be relevant to the quality of the sound. Therefore, with vowels we will speak not only of highness, but of lowness. [i] is a high vowel. Is any vowel we've discussed yet low? Certainly: [ɑ]. These two vowels are at the extreme ends of the scale from high to low. But many other vowels come in along that scale. And some of these vowels are nearer to the middle of the scale than to either end. They are called the **MID** vowels and we say they are [–high, –low].

What about [u] with respect to the features of high and front? [u] also is a [+high] vowel. But now the tongue is bunched up in the velar area. Is [u] as high as [i]? Look back at our midsagittal section in Figure 1.1. The roof of the mouth slopes downward from the hard palate to the velum. So, because of our physiology, [u] is lower than [i]. But the important point is that [u] is as high as it can be for the velar area, so it is [+high]. Now about frontness, [u] is [–front]. But just as we made a three-way distinction with consonants between [+anterior], [+back], and [–anterior, –back] sounds, we will have reason to do the same for vowels (when we meet the central vowels). So [u] is a [+back] vowel.

Go back and forth between [i] and [u] several times. They contrast on frontness and they are similar on height. But that isn't the full story. Can you notice anything else different between them? Listen to them. What about **PITCH**? If you can't hear a difference, whisper the two sounds. Can you hear it now? [i] has the higher pitch of the two. Oh, certainly you can change the pitch on a vowel. Sing [i] from a low to high pitch. Do the same for [u]. But if we are just talking normally and not making any attempt to control our pitch, [i] has a naturally higher pitch than [u]. Why is that? Do an experiment now. Take two bottles of the same shape and size. Fill one with water half way. Fill the other with water one-fourth of the way. Now blow across the tops of the bottles so the air resonates in the space in the bottles not filled with water. Which has a higher pitch? The one that has more water in it. Why? Smaller resonating chambers give higher pitch. We will discuss why this is so when we turn to acoustics. But for now, draw your own version of Figure 1.1 and draw in the tongue as it would be for making [i] (bunched up at the palatal area). Now draw in the tongue as it would be for [u] (bunched up at the velar area). Where is the

resonating chamber for each sound? It is the space in front of the bunched up portion of the tongue. So now you can see why [i] is naturally higher in pitch: the resonating chamber in front of it is smaller than the resonating chamber in front of [u].

Vowels definitely differ on their natural pitch. We will not, however, discuss this matter further here, because in this section we are talking about what we can notice in producing sounds (articulatory phonetics), not what we can notice in receiving sounds (acoustic phonetics). So please just lay this sort of difference in vowels aside and consider only the features of front, back, high, and low.

Tense and Lax

Put your tongue in the position to say [i]. Surround that sound with [b] and [t]—say [bit]. (We're back to staccato vowels here to avoid diphthongs.) Now say *bait*. Does it have a diphthong or not? Probably you say it with a falling diphthong. (And this is true even if you say *beat* with a long vowel rather than a diphthong.) If so, say it staccato. You can transcribe the staccato pronunciation with the new vowel [e]: [bet]. (Some books call this CLOSED "e".) How would you transcribe the normal pronunciation? ([beɪt]). Is the bunched-up part of the tongue high in the mouth or not? It's not very high, but it's also not flat like with [a]. So [e] is our first front mid vowel; its features are [+front, -high, -low]. Now say [i] and [e] in succession. Then try to slide between the two. Stop midway on that slide. Surround the vowel here with [b] and [t]. Do you recognize the word you're producing? It's *bit* again: [bit]. We say that this front vowel is also [+high]. So how can we distinguish [i] from [ɪ] with a feature bundle?

Unfortunately, we cannot see the articulatory difference here, but perhaps you can feel it. Say [i] and [ɪ] staccato style. Does one seem to involve more tension in the tongue? In fact, to make [i], inside the pharynx we project the tongue root forward, creating a larger pharyngeal opening. We say the [i] has the feature of [+ADVANCED TONGUE ROOT], or just [+ATR] for short. We also say that [i] is a TENSE vowel, whereas [ɪ] is LAX.

Okay, now pronounce in staccato style the words we've talked about so far: *beat*, *bit*, *bait*—[bit], [bit], [bet]. Say just the vowel sounds in succession: [i], [ɪ], [e]. Now let the tongue lower just a little. What word do you get? *Bet*—[bet]. We call [e] EPSILON or OPEN "e." Here we have another [-high, -low] vowel that is [+front], just like [e]. Which of the two is [+ATR]? ([e])

If you find yourself unable to feel the tension difference between the [+ATR] vowels ([i] and [e]) and the [-ATR] ones ([ɪ] and [ɛ]), you're not alone. There is another way you can help yourself remember them. Say all the words we've been working on to find these vowels. What can you say about vowels with respect to diphthongs? When we are dealing with monosyllabic words, the tensed vowels typically occur only in diphthongs with a lax vowel, but the lax ones can occur in a syllable that does not have a diphthong. When we have a diphthong in English, which vowel is the syllabic one: the tense or lax vowel? The tense one.

Okay, lower the tongue even more. Put the new vowel that you are making between [b] and [t]. What word do you have? *Bar*, written as [bæt]. This vowel is [+low, +front]. Can you tell if it is tense or lax? If you can't feel whether the tongue root is advanced, listen for a diphthong. If it has a diphthong, it must be tense. If it doesn't have a diphthong, it could be tense or lax. It has no diphthong, right? And, in fact, this vowel is [-ATR]. We name [æ] ASH, or DIGRAPH.

So we have discussed five front vowels. We have also discussed the high back vowels [u] and [ʊ], as in the words *cooed* and *could*. (The first has a diphthong.) Which of these vowels is [+ATR]? ([ʊ]) Now lower the tongue to midway and keep the same consonants surrounding the vowel. What word do you get? *Code*. When you say this word in ordinary style, does it have a diphthong or not? Yes—[kʰoʊd]. (And, again, even if you say *cooed* with a long vowel rather than a diphthong, you probably have a diphthong in your pronunciation of *code*.) Get rid of the glide and just say the vowel in isolation: [o], sometimes called “closed o.” Now lower the tongue a bit and what new word do you get with this new vowel? *Cawed* (as in what a crow did): [kʰɔd], where [ɔ] is called OPEN “o” OR BACKWARDS “c.”

At this point we come to a major split in American English speech. While NATIVE SPEAKERS along the east coast, including the South, as well as speakers in Los Angeles and San Francisco, pronounce the word *cawed* with [ɔ], most of the rest of the speakers in America pronounce this word with [ɑ]. (A native speaker of a particular language is one who learned the language before the onset of puberty. A grammar of a language is always based on the linguistic behavior of native speakers only, since language learned after the onset of puberty is rarely of the same competence as native competence.) In fact, the vowel [ɔ] does not occur in the speech of these American English speakers at all. So if you are a speaker who does not have [ɔ] in your variety of English, you will have to treat this vowel like you would any other sound from a foreign language and simply learn how to produce it and how to recognize it. If you let your tongue slide between [ɑ] and [o], you should be able to pick out a middle point that is close to [ɔ]. Do that now until you feel comfortable producing this vowel.

Now compare [o] with [ɔ]. Both are [+back, -high, -low]. So which one is [+ATR]? ([o]) (Here your diphthong test should have been a big help, right?)

Okay. We now have four back vowels. At this point we have one more front vowel than back vowel. Many books in linguistics list [ɑ] as the missing low back vowel. Other books call it a CENTRAL vowel. We will call it a back vowel because in many languages phonological rules group [ɑ] with the back vowels. However, there is one feature that the other back vowels have that [ɑ] does not have. To see this, hold a mirror in front of you and say the front vowels, and then the back vowels, leaving out [ɑ]. What do you notice about the lips? The back vowels in English are [+ROUND]—they call for rounding of the lips. But [ɑ] is [-round].

There are, however, more vowels than just these front and back ones. Say the word *but*. This vowel is [ʌ], called CARET or UPSIDE-DOWN “v.” Can you feel where the tongue slightly bunches? Is it front, central, or back? I hope that you didn’t say it was front. Many books classify it as central, others as back. We’ll return to this later.

Now say the word *roses*. In the second syllable of this word we meet yet another vowel, called SCHWA [ə]. Is this vowel higher or lower than [ʌ]? It’s a bit higher, but it may be hard for you to convince yourself of that because we have no minimal pairs for these two vowels. In fact, for most speakers of English, [ə] occurs only in unstressed syllables and [ʌ] occurs only in stressed syllables. You might want to adopt that convention for yourself as you transcribe English words. We say that [ʌ] is a low vowel (though not as low as [ɑ]), while [ə] is a mid vowel. Is [ə] front, central, or back? This is one vowel that linguists agree should be classified as central. Some people pronounce the second syllable of *roses* with a higher central vowel, written as [ɪ]. One way to see if you do is to say “Rosa’s roses” (with an American English pronunciation of *Rosa*: [ɹoʊzə]). If the words differ, you probably use [ə] in the final syllable of *Rosa’s* and [ɪ] in the final syllable of *roses*.

Beyond (Standard) English

These vowels pretty much cover the vowels we find in what you might call standard American English. However, there are other vowels that occur in natural language, some of which occur in various regional varieties of American English. You probably are very aware of that fact already if you have been contrasting my pronunciation of the words I've discussed thus far with your own—for the quality of vowels is the most noticeable difference between varieties of English. For example, in Boston some people might pronounce the word *Harvard* without either "r"-sound and with both vowels pulled back and slightly rounded, written as [ɒ]: [hɑ.vɒd]. Contrast that to the more central vowel of a midwesterner saying [hɑ.ɪ.vɪd], or another type of Bostonian using a front vowel, as in [hæ.vɪd], or an even lower front to central vowel, written with [a], as in [hɑ.vɪd]. And notice that here the Bostonian is using the same vowel that southerners use in words such as *my* [ma].

Another way we can find more vowels is to play with the feature of roundness. In American English, front and central vowels are [—round]. But they don't have to be that way. So make [i]. Now round your lips and say the new front round vowel: [y] (which is often written as [ü]). Do you know a language that makes use of the high front round vowel? French does (as in the word *tu* 'you' [ty]); so does German (as in the word *Tür* 'door' [tyr]). German also makes use of the lax high front round vowel [ʏ] (which you make by rounding [ɪ]). Now take your mid vowel [e] and round it. You get the new vowel [ø] (often written as [ö], and found in the French word *peu* 'a little' and the German word *Söhne* 'sons'). Now take your lax mid front vowel [ɛ] and round it. You get the new vowel [œ] (often written as [ö], and found in the French word *peur* 'fear').

You can also take your back vowels and unround them. Try it for the tense vowels first. You get the high vowel [u] and the mid vowel [ʏ] (which is also written as [ë]), both found in Korean. You can also unround your lax mid back vowel. What you get may well be identical to [ʌ], and some books list this vowel not as a central vowel but as a back unrounded vowel, as I said above.

As with consonants, this isn't the whole story, but it's enough to get you sensitized to the sorts of distinctions you should listen for. The appendix at the end of this chapter gives details of these vowels and their features.

Acoustic Phonetics

Distinctions in the physical properties of the sounds that your ear receives are our next focus.

There are at least three different audible characteristics of language sounds: pitch, loudness (or INTENSITY), and quality. So far when discussing how we make sounds, we have concentrated on the quality of those sounds. Let us turn now to the other two characteristics, both of which are called SUPRASEGMENTALS. In a language like English, only quality is essential to recognition of the phoneme: we can hold the quality of a sound steady and change its pitch and intensity without changing its identity as a phoneme. Nevertheless, the pitch and intensity of a sound are important, and when we put sounds together to make words, we pay particular attention to these suprasegmentals, as you will learn in our discussion of Metrical Phonology in Chapter 2. For now let's just try to understand what these two suprasegmentals are.

Consider a pendulum. If you set it moving, it will swing upward to one side, then back toward its resting point, pass through that point, and continue on in an upward arc to the other side. At a certain point it will reverse direction again and go back to the resting point. At this point we say the pendulum has completed one **CYCLE** of movement. Now it passes through the resting point and heads upward on the first side again, beginning the second cycle of movement. Repetitive movement like this is called **PERIODIC**.

Now pluck a guitar string. Your pluck has set the string in motion, and the motion is periodic. Do you see that? Give me other examples of periodic movement.

Not all motion is repetitive or periodic. Swing your arm once through the air. That's an **APERIODIC** movement. Drop a book on the floor. That's an aperiodic movement.

Okay, now pluck the guitar string again (or any other string that is tightly attached at both ends). You can see its periodicity. But what you don't see is that your pluck has also set the air in motion. That string moved to one side, displacing the air on one side of it—thus compressing it. Then it moved to the second side, rarefying the air on the first side and compressing it on the second side. (Of course, much more air is compressed if the guitar string causes a larger surface, called a soundboard, to move, also. But let's just assume our guitar string isn't even attached to a guitar, but only to a peg at each end so that all we have to discuss is the movement of the string and its affect on the air.) Each time the guitar string makes a cycle of movement, the air likewise makes a cycle of movement. But the molecules in the air touch other molecules, which are likewise compressed and rarefied in a pattern. We say the guitar string has set up a periodic **SOUND WAVE**, which is a regularly increasing and decreasing sequence of air pressure that is transmitted out in every direction from the string.

Let's take a moment here to consider that fact. All sound must travel through a **MEDIUM**. The air is our typical medium for speech sounds. Can you guess what happens if we hang a bell inside a closed jar and then suck out all the air in the jar and ring the bell? There is no sound. That's because there is no medium to carry the sound wave. What do you think happens if we hit a spoon on a metal table? Try it. Hit the bowl of a spoon very hard on the edge of a metal table, then hit it very hard on the edge of a wooden table. What do you hear besides the smack of the spoon? The metal table probably set up a buzzing sound. The wooden table probably did not. That's because metal is a more efficient medium for carrying sound waves than wood is; metal absorbs less of the sound internally than wood does. When you hit the spoon on these two tables, you set up other vibrations, called sympathetic vibrations, that travel through the metal or wood and are prolonged by the metal or wood.

Pitch

Okay, now let's return to our discussion of the vibrations of the guitar string. We say that the amount of time it takes for a vibration to complete a cycle of compression and rarefaction is its period. (And for the sake of a simpler exposition, we will hereafter use the term "cycle" to mean one compression and rarefaction.) The number of periods that a sound wave has in a given amount of time is its **FREQUENCY**. For speech sounds we use one second as our measure of time. So the frequency of a sound would be how many cycles a sound wave executes in one second (or how many periods it has in one second). The unit of measure here is the **HERTZ**, abbreviated as Hz. For example, if a cycle of a sound wave takes 0.01 second to complete, that means that in a full second it will repeat 100 times. We say that sound has a frequency of 100 Hertz. In case you haven't used much

mathematics in your life recently, let's go over this. You can take what we just did above and restate it as: One second divided by 0.01 second = 100 Hz. (An [l], for example, is a sound that has a frequency in the range of about 100 to 125 Hz.) If a cycle of a sound wave repeats every 0.1 second, what would its frequency be? (10 Hz. One second divided by 0.1 second = 10 Hz.)

Let's put this into a formula. Let 1 stand for one second. Let p stand for period (the amount of time it takes for a single cycle). Let f stand for frequency (Hz). So we have just realized that in general:

$$\frac{1}{p} = f$$

Now let me ask you a new question that draws on the same information: If a sound had a frequency of 1,000 Hz, how long would it take a cycle of that sound wave to repeat? In other words, what would the p of that sound be? (0.001 second) You can answer this by plugging in values to the formula above. And you can manipulate the formula so that you isolate p on one side:

$$\frac{1}{f} = p$$

Let's talk a bit more about the experiment you did above when we were discussing the relative pitches of [i] and [u]. You filled one bottle with water halfway and another with water one-quarter of the way. Now think about what you did to the air inside the bottle when you blew across the top. You set the air in motion. When the moving air molecules hit an obstruction, they bounced off and traveled in another direction, until they hit another obstruction and then bounced off again. If you keep blowing, the air motion inside will keep repeating. If you have two cavities of different sizes (like two bottles filled to varying degrees with water, or two bottles of different sizes), the frequency of the air motion will be faster in the smaller one. That's due to a fact of physics that I feel ill-equipped to explain to you. Essentially, smaller bottles make higher frequencies because the stiffness of the air inside the bottle is inversely proportional to the volume.

Consider the specific case of our bottles that are partially filled with water. The space in the bottle that is not occupied by water is the resonance cavity for the air. Now do you remember which cavity produced a higher pitch? (the smaller one) Okay, now can you state the relationship between frequency and pitch? The higher the frequency (that is, the shorter the periods of a sound are), the higher the pitch. Thus Hertz can be thought of as a measure of pitch.

However, the relationship between pitch and Hertz isn't really perfect. That's because even sounds for which Hertz cannot really be measured—that is, even sounds that have no frequency—have pitch. We will return to this point. But first, let's talk about these other types of sounds.

Intensity

Not all sounds have a frequency. What sorts of sounds would set up a sound wave that was not repeating? Aperiodic sounds, of course. A slap of your hand on a table, for example,

sets a sound wave in motion, but it is not a repeating sound wave, because the slap is nothing like a vibrating string. In Figure 1.3 we can see the sound wave set up by a tuning fork (1.3A) and the sound wave set up by a clap of hands (1.3B). The boxed area in 1.3A and 1.3B is enlarged at the bottom of each figure. The wave for the tuning fork is periodic; that for the hands clapping isn't.

The horizontal axis measures time, starting from zero, which marks the moment when the sound begins. The vertical axis measures air pressure, where zero is the pressure before the sound begins. The movement of this wave above the horizontal axis going through the center of each graph shows compression of air. The movement of this wave below the horizontal axis shows rarefaction of air. We call the height of the wave its **AMPLITUDE**. Tell me—if you want to make a louder sound with your guitar string, what can you do? You pluck it harder, right? That means you displace the string more. Therefore, the air around the string is displaced more. Therefore, the compression and the rarefaction are both greater. What does that mean about the amplitude of the sound wave? How will it be affected? It will increase. Do you see that? So the vertical axis on Figure 1.3 is measuring the intensity of the sound. We measure intensity in **DECIBELS**, abbreviated as dB. We say a given sound is so-many decibels. If we increase the intensity of a sound by 5 dB, we are doubling the intensity of the sound. Typically increases of less than 1 dB are not noticed by our ears.

Say these continuous sounds for five seconds each: [m], [z], [e], [l]. Can you hear a difference in intensity among them? As with pitch, we can easily vary the intensity of a sound: say [z] softly and then say [z] loudly. But without any instructions about varying intensity, you might have noticed that these sounds have different natural intensities, just like the vowels and consonants have different natural pitches. We have an **INTENSITY CLINE** that goes from most intense to least intense:

vowels, glides, liquids, nasals, voiced fricatives,
voiced oral stops, voiceless vowels, voiceless
fricatives, and voiceless oral stops

If you couldn't hear the intensity difference among the four sounds I asked you to say, that's okay. Many people have trouble hearing it. (I'm in that group.) What can you notice about the intensity cline? Does it remind you of anything else we've discussed? Certainly. **It starts out very much like the sonority hierarchy, although it differs once we get past the middle of the cline. This tells us that intensity is a correlate or factor of sonority in many instances.**

Back to Pitch

Let us return now to our discussion of pitch. Look back at Figure 1.3. You can count cycles by counting the highest peaks. What is the frequency of the sound wave of this tuning fork? To figure this out, take a close look at the enlargement of the boxed part of 1.3A. Start at any fixed point and estimate how many seconds it takes for a cycle to repeat. For example, if we start at time 0.033, we find a completed cycle by 0.0353. So it takes approximately 0.0023 second for a repetition. Actually, if your eye is good enough, you can see that it takes closer to 0.00228 second for a cycle. So how many cycles would we have in one full second? To find this, you divide one second by 0.00228. The answer is 439. That's pretty close; this tuning fork gives the note A, and its frequency is 440 Hz.

Figure 1.3A. The sound wave of a tuning fork. (The note is A.) The lower figure is an enlargement of the boxed in area of the upper figure. The horizontal axis measures time from the moment the sound begins in seconds. The vertical axis measures air pressure, where 0 is the pressure before the sound begins.

Figure 1.3B. The sound wave of clapping hands. The lower figure is an enlargement of the boxed in area of the upper figure. The horizontal axis measures time in seconds. The vertical axis measures air pressure.

We say the tuning fork has a pure sound (at least when the sound travels through air, not through other media) because what we are hearing is its **FUNDAMENTAL FREQUENCY**, uncluttered by any sympathetic resonances.

But even aperiodic sounds can be said to have pitch, although they have no defined frequency. You know this; your ear tells you this. Say [s] and then say [a]. [s] is a voiceless sound, so the vocal cords aren't vibrating. (You'll learn the relevance of this fact immediately below.) Furthermore, the cavity in the mouth is highly occluded, so we have an irregular, random pattern of noise. But you can hear a pitch difference between [s] and [a] all the same. Which is higher? [s], very clearly, no? So [s] has a pitch, even though no **COMPONENT** of its sound wave is periodic. And that's why Hertz do not really equal pitch.

So how do we measure the pitch of an aperiodic sound? We do it by using a mathematical formula. We take a fixed interval of time and we look at the sound wave during that interval. In effect, we declare that that interval is the period of the sound. We choose how long an interval of time to look at based on certain properties of the particular sound we are analyzing (and I will not go into those properties here). We do this not because the sound actually has a period, but because in this way we can then speak of the Hertz of the sound. (Tell me, what would be the Hertz of a sound that had a period of 0.0005 second? 2,000 Hz, right? What would be the Hertz of a sound that had a period of 0.00025 second? 4,000 Hz. Many consonant sounds of English fall in the 2,000- to 4,000-Hz range.

Now, let's go back to periodic sounds. Your vocal cords are a lot like a guitar string. And your vocal tract contains the body of air through which the sound wave set up by your vocal cords passes. When you make a [+voiced] sound, the vocal cords vibrate in a periodic way. The sound wave moves through the oral and, sometimes, the nasal cavity. But inside those cavities, the air bungs against the various obstacles and sets up other vibrations with their own frequencies—**SYMPATHETIC RESONANCES**. These additional sound waves form a complex sound wave with the one set up by the vocal cords. And what that wave looks like depends directly on the shape of the vocal tract, which you control by movements of your tongue and teeth and velum and lips. In other words, the sympathetic resonances are major contributors to what makes voiced sounds distinct from one another. Consider a sound wave for [l], given in Figure 1.4.

Figure 1.4. Sound wave of [l].

Compare this sound wave to those for the tuning fork in Figure 1.3A and the hands clapping in Figure 1.3B. You can certainly see that the sound in Figure 1.4 is periodic. But you can also see that the sound wave for [l] is complex: The minor peaks between these highest peaks show the sympathetic resonances due to the shape of the vocal tract.

Think now about the shape of the vocal tract. For which kinds of sounds do we have the least occlusion and hence the greatest opportunity for sympathetic resonances? The vowels, of course. But also the nasals, the liquids, and the approximants.

Formants

Consider first the vowels. Acoustically they are complex because they have a fundamental frequency (like the tuning fork has), plus they have sympathetic vibrations at whole-number multiples of their fundamental frequency. For example, if the fundamental frequency is 100 Hz, there will be sympathetic vibrations, called HARMONICS, at 200, 300, 400, etc., Hz. The vocal cords produce the fundamental frequency, and this frequency determines the perceived pitch of the sound. The vocal cords also produce all the harmonics. Most of the harmonics are not significant; however, all vowels have at least four harmonics that are more prominent, called OVERTONES or FORMANTS. The formants for a particular vowel correspond to the resonances of the vocal-tract shape as you say that vowel. We can say that the harmonics are shaped by the resonances of the vocal cavities. The formant with the lowest frequency is called the FIRST FORMANT.

When we ended our discussion of consonants, we talked briefly about different kinds of phonation. One was called the creaky voice. Make a noise like the creaking of a door in a horror story. Now say this whole sentence with that creaky voice. You cannot see what you are doing, but can you guess where the significant changes are taking place? In the larynx, and that's why speaking in a creaky voice is also called LARYNGEALIZATION. Inside your voice box at one end you have the arytenoid cartilages. To make a creaky voice, you pull the arytenoid cartilages tightly together so that the vocal cords at that end of the voice box are also pulled together so tightly that they cannot vibrate. But the vocal cords do vibrate on the opposite end. The result is a very low pitched sound. Sing [a] from as high up as you can go to as far down as you can go comfortably. Now go even lower. Probably you are now laryngealizing the [a].

Creaky voice is of interest to us now because it can help us hear the overtone known as the first formant. Say the back vowels with a creaky voice, sliding from [a] to [u]. Can you hear a pitch change? It is gradual and somewhat delicate, but if you do this a few times, maybe you can hear it. **We are finding a falling pitch as the back vowels rise.** Now say the front vowels with a creaky voice, sliding from [æ] to [i]. Can you hear the pitch change? Again, **the pitch falls as the vowels rise.** In other words, **the first formant has an inverse relationship to vowel height. This is very nice, for it tells us that our vowel height distinction, which we made on articulatory grounds, has an acoustic correlate.** (Indeed, fine distinctions in vowel height are a bit suspicious on purely articulatory grounds, so this finding is welcome.)

The SECOND FORMANT of vowels can also be heard by using a special phonation, that of whispering, whereby the vocal cords don't vibrate (so there is no fundamental frequency). Start at the high front vowel [i], go down the front of the mouth gradually to [æ], then move on to [a] and go up the back of the mouth to [u], whispering the whole way. What do you notice about the pitch? It falls continuously.

I know of no way to hear the **THIRD OR FOURTH FORMANTS** of the vowels, but this is not a terrible problem since these two formants are quite high and they vary little from vowel to vowel (although you will discover something about them if you do Problem Set 1.7). Furthermore, the fourth formant can vary considerably according to the quality of individual voices. (In fact, the fourth formant is useful for speech recognition and identification systems.) So the important formants for recognizing the quality of vowels are the first and second (although if you do Problem Set 1.7, you will discover that the third formant can tell us about the feature of rounding).

Spectrograms

Speech scientists have found a way to give us a visual representation of the component frequencies of sounds: the spectrogram. In Figure 1.5 you see a spectrogram for the vowels [i], [a], and [u], as Bill Reynolds (an American phonologist) says them.

The horizontal axis measures time, just as the horizontal axis for our sound waves in Figures 1.3 and 1.4 did. The vertical axis, however, measures hertz, and we have markers every thousand hertz. Intensity is now shown by increased darkness. The resonant frequencies of a sound are the loudest parts, so they show up as the dark horizontal bars. Now you can easily see the four formants we've just been talking about; I have labeled them for each vowel. You can see that the first formant of [a] is higher than the first formant of both [i] and [u] (as you discovered by saying the vowels in a creaky voice), and the second formant of [a] is midway between the second formant of [i] (which is higher than it) and the second formant of [u] (which is lower than it, as you discovered above by whispering the vowels).

There is another important piece of information we can glean about these sounds from this figure. Look at the vertical **STRIATIONS** (the vertical marks). These show momentary increases in acoustic energy due to a single cycle of movement. If the cycle is very short, the striations will be closer together than if the cycle is long. So higher-pitched sounds (for which the cycles are short) will have the striations more densely packed together than lower-pitched sounds. That means that the formants on the front vowels will typically be

Figure 1.5. Spectrograms of the vowels [i], [a], and [u], with the four most prominent formants labeled.

Figure 1.6. Spectrogram for the sounds [l̥ ɛ l̥].

denser (and thus appear darker) than the formants on the back vowels. This contrast between [i] and [u] is particularly striking in Figure 1.5.

Now consider the spectrogram in Figure 1.6. Here the speaker is going from a light lateral to a dark lateral with the open mid front vowel between—that is, [l̥ ɛ l̥]. The vowel in the middle is easy to see. Contrast the laterals on either side of it to the vowel in the middle. How many formants do you see for these laterals? (Please ignore the striations at the very bottom for now. We will account for them when we discuss Figure 1.7.) Only three of significant intensity, right? As compared to the vowels, which formant seems to be missing? The third—can you see that? For most speakers in most contexts the third formant is of such weak intensity that it does not appear clearly on a spectrogram. What about the comparative intensity of the liquids' formants as compared to the vowels' formants—do you see a difference? The intensity is less on the liquids.

You are going to explore formants for nasals in Problem Set 1.6, so we won't consider them here. But what do you guess happens on a spectrogram when we nasalize a vowel? Both vowels and nasals have their own formants. When we nasalize a vowel, the vowel's formants have the nasal's formants superimposed. That means nasalized vowels will show up as very dark wherever the two sets of formants overlap. It also means that you may see more than four formants for the vowel, if one of the nasal's formants does not overlap with any of the vowel's formants. We call the extra formants **PSEUDOFORMANTS**. While English does not nasalize vowels in a distinctive way (in contrast to French, for example), certain regional varieties of American English are characterized by partially nasalized vowels, particularly before a nasal consonant, and many speakers of English may partially nasalize their vowels just as an individual (rather than regional) characteristic.

What do you think the spectrograms of [s] and [z] will look like? Will the striations on the spectrogram be relatively high or relatively low? Think back to the pitch of [s] as compared to [a]. So you expect striations at the upper end. All fricatives, in fact, have quite high pitch. Say [s] and compare it to [f]. Which is higher? [s] is. So even though all fricatives will have striations relatively high on the spectrogram, you will be able to see a difference between [s] and [f] in that [s] will be higher and its striations will be denser (hence looking darker). In fact, [s] and [z] have quite dark striations on spectrograms. Do you expect to see formants for [s] and [z]? No, these sounds don't resonate in the oral

Figure 1.7. Spectrograms for [s] and [z].

cavity. They are hissy and buzzy, not melodic. We can see their spectrograms in Figure 1.7.

Can you guess what the dark stuff at the very bottom of the spectrogram for [z] represents? It runs along the time bar and extends up to around 250 Hz. Notice that this is a major difference between [s] and [z]—so that fact should help you to guess what this dark mark is. This is called the **VOICE BAR** and it shows the frequency of the vocal cords. All voiced sounds will have a voice bar on their spectrogram, while voiceless ones will not. (I didn't point out the voice bar in Figs. 1.5 and 1.6 both because it is difficult to see on those figures and because all of our sounds in those figures were voiced, so no contrast could be made.)

What do you think the spectrograms of [k] and [g] will look like? Try to make a [k] in isolation of any other sound. If you're having trouble, say the word *tack* and then say it again, without pronouncing anything but the final consonant. There's complete silence, and then a voiceless release. How about with [g]? Say the word *tag*, and then say it again, pronouncing only the final consonant. You are voicing those vocal cords before you release. So there will be nothing on the spectrogram for [k], followed by relatively high striations. But there will be a voice bar for [g], followed by similar high striations. What if you pronounce *cool* and *ghoul*, and isolate the initial consonants? The aspiration on the initial voiceless consonant will come out as a moment of silence, before the onset of voicing (of the following vowel). In Figure 1.8 you can see the spectrograms of these four words. Just based on how the words begin and end, can you label which spectrogram represents which word? (The words are in the order *tack*, *tag*, *cool*, *ghoul*.)

Stop for a moment and look back at the spectrograms of the vowels in Figure 1.5. Compare the spectrogram of [u] there to the spectrogram of the [u] in the middle of *cool* and *ghoul*. Notice that the formants for this vowel are not identical in the two utterances. When [u] is pronounced in isolation, the formants stay relatively steady (and when I recorded the vowels in Fig. 1.5, I asked the person I was recording—Bill Reynolds—to be careful to avoid diphthongization). But when [u] is pronounced as part of a word, the formants begin in a certain way and end in another certain way, with only their middle part reliably close to the formants we see of the vowel in isolation. Why is that? Certainly, there may be a slight falling diphthong for the words in Figure 1.8 for many speakers, and

Figure 1.8. Spectrograms for four words, with attention to velars.

I believe there is for the speaker I recorded here. But there is more difference than just that: The sounds preceding and following any given sound can affect that sound's beginning and ending, respectively. That is, the initial [t^h] vs. [k^h] and the final choice of [k], [g], or [l] affect the sounds in the middle of the words. **Given that utterances are generally made up of syllables, and that syllables generally have vowels as their nucleus, we can think of consonants as different ways to begin or end vowels.** (You will find this way of looking at things useful as you do Problem Set 1.8.)

Figure 1.9 shows a spectrogram of the sentence *I like Staymen apples*. Can you see the diphthong in the first syllable? Cover the figure a moment and ask yourself what you expect to see happen to the first formant of the vowel [a] (in the word *I*) as it moves into the [-syllabic] vowel [ɪ]. Remember, the first formant varies inversely to height. So you expect the first formant to fall. Now what do you expect the second formant to do as we move from the [+syllabic] to the [-syllabic] vowel? Remember that the second formant falls continuously from the high front vowel down to the bottom of the mouth and then keeps falling as we go up the back of the mouth. So we expect the second formant to rise in this diphthong. Now look at the figure. Do these formants behave as you expected? Aren't you happy?

Figure 1.9. Spectrogram for the sentence *I like Staymen apples*.

Cover the IPA transcriptions below the spectrogram and examine this spectrogram, trying to locate every sound segment in the utterance on the spectrogram. I think you can do it. Let's talk about the obstruents first. Do you see the silence at the [k] of *like*, the [t] of *Staymen*, and the [p^h] of *apples*? Can you see the high-pitched striations that mark the fricative [s] of *Staymen* and the final fricative [s] of *apples*? (In Problem Set 2.9 you will learn that many speakers would have a [z] as the final sound of *apples*, but this speaker had [s].)

Now let's turn to the sonorants. Can you see the same diphthong in *I* and *like*? Do you see the [l] that comes between them? (The [l] should have only three formants of significant intensity. Unfortunately, the second formant shows up more darkly than expected. There's a lesson here: speakers vary so spectrograms vary.) Do you see the diphthong in the first syllable of *Staymen* and the [m] that immediately follows it? Notice that the [m] has formants, like the vowels, but fewer. (You'll deal with this issue in Problem Set 1.6.) Can you see any difference between the nucleus of the first and the second syllables of *apples*? The first one is more intense, right? The same is true for the nuclei of the first and second syllables of *Staymen*. **So we can now see that intensity is a correlate of stress.** How does the spectrogram show you that the [l] in the second syllable of *apples* is the syllabic segment and that there is no vowel here? Can you see that there are only three relatively intense formants, whereas all vowels have four? Compare the [-syllabic] [l] of *like* to the [+syllabic] [l] of *apples*. Do you see any difference there? The syllabic one is more intense. (This is very difficult to see on this particular spectrogram, but if you compare the first formants for the two laterals, you might be able to see the intensity difference.)

That's all we're going to say about acoustics, even though, as you by now know (given the way I tend to end each major discussion in this chapter), there's much more to say. It's time for us to pass on to the job of looking at the principles that govern what happens when we put sounds together into larger units. That's called (PHONOLOGY) and that's the matter of Chapter 2.

Appendix

International Phonetic Alphabet: Consonants

In Appendix Table 1 the consonants discussed in this chapter are shown (including one that comes up in the problem sets). Across the top are labels for place.

Affricates and other doubly articulated sounds are not included since the reader can put together all the relevant information on them by considering their component sounds. The following abbreviations are used: bilab = bilabial; lab-den = labio-dental; inter-den = interdental; alveo = alveolar; post-alveo = postalveolar; retro = retroflex; palat = palatal; uvu = uvular; phar = pharyngeal; glot = glottal; oral = oral stop; tap/f = tap or flap; fric = fricative; lat = lateral; appr = approximant; eiec = ejective; impl = implosive. The symbols I have used for the clicks are not the ones approved by the IPA; however, they are very commonly used, and I prefer them.

Additionally, we have the sounds:

lateral fricative: ɬ

epiglottal fricative: Hʕ

epiglottal voiced stop: ʕ

Appendix Table 1. IPA: Consonants

	Bilab	Lab-den	Inter-den	Dental	Alveo	Post-alveo	Retro	Palat	Velar	Uvu	Phar	Glott
Oral	p b			t	d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
Nasal	m	ɱ			n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			r						ʀ		
Tap/f					ɾ		ɽ					
Fric	ɸ β	f v	θ ð	s z		ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lat					l		ɭ	ʎ	ʟ			
Appr								j	ɥ			
Ejec	pʰ			tʰ			tʰ	cʰ	kʰ	qʰ		
Impl	ɸ ɸ̥			f f̥				ç ç̥	ʁ ʁ̥	ɣ ɣ̥		
Click	ɸ			ǀ	ǃ ǂ			≠				

NOTE: When two symbols occupy the same cell, the first is voiceless and the second is voiced, except in the click row, where all sounds are voiceless.

And, finally, we have these special symbols for some common doubly articulated sounds:

rounded velar approximant: w

alveopalatal affricates: ɽ ʈʂ

double articulation of ʃ and x: ʃʃ

Diacritics

A small circle under an IPA consonant indicates lack of voicing on a normally voiced sound, for example: [t̚].

A tilde, ~, under an IPA symbol indicates creaky voice, for example: [a̰].

A small vertical line under an IPA consonant indicates that this consonant is syllabic, for example: [n̩].

Labialization is indicated with a raised w—for example: [kʷ], [gʷ].

Retroflexing is indicated by a rightward-curving hook at the bottom right of a symbol, for example: [ŋ̠].

Ejective sounds are indicated by an apostrophe following the symbol, for example [pʰ].

Implosive sounds are indicated by a rightward-curving hook at the top right or left of a symbol, for example: [ɖ̥] and [ɓ̥].

Length is indicated by a colon following the symbol, for example, [l:].

Breathiness is indicated by two horizontally arranged dots under the symbol, for example [a̰̰].

A period between symbols indicates a syllable boundary, for example: [la.ka].

Breathy sounds are indicated by two horizontally arranged dots under the symbol.

Appendix Figure 1. Vowel chart giving front (unrounded and rounded), central, and back (unrounded and rounded) vowels.

International Phonetic Alphabet: Vowels

In Appendix Figure 1, where symbols come in pairs, the one on the left is unrounded and the one on the right is rounded.

I have listed [a] as a front vowel, but some texts list it as central. I have listed both [ɑ] and [ʌ] as unrounded back vowels, but some texts put one or both as central or even back vowels.

The vowels found in most varieties of American English are typeset boldface.

Diacritics

The symbol _̚ below a vowel indicates that it is nonsyllabic, for example: [aɪ̚].

A rightward hook under a vowel indicates nasalization, for example: [ǎ̟].

Alternatively, a tilde can indicate nasalization, for example: [ã].

Tone marks above vowels are ˈ for high tone and ˋ for low tone, for example: [à], [á].

The diacritic ˘ over a vowel indicates very short duration, for example: [ă].

A macron (ˉ) over a vowel indicates long duration, for example: [ā].

Alternatively, a colon after a vowel indicates long duration, for example: [a:].

An acute accent (ˊ) over a vowel can indicate primary stress, for example: [à].

Consonants

Appendix Table 2 shows the values for twelve consonant features discussed in Chapter 1. The features of [spread glottis], [constricted glottis], and [ATR] are not included. Consonants heard in most varieties of American English are boldface. (del-rel = delayed release) A list of words illustrating twenty-seven of these consonants with English words is also shown.

English Words that Use These Consonants

[p]: sop	[b]: sob	[m]: psalm
[f]: fat	[v]: vat	
[θ]: thigh	[ð]: thy	
[t]: cat	[d]: cad	[n]: can

Appendix Table 2. Distinctive Features of Most Commonly Found Consonants

	p	b	m	ɸ	β	f	v	θ	ð	t	d	n	s	z	l	r
Anterior	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Back	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Continuant	-	-	-	+	+	+	+	+	+	-	-	-	+	+	+	+
Coronal	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+
Del-rel	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
High	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Labial	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-
Lateral	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-
Nasal	-	-	+	-	-	-	-	-	-	-	-	+	-	-	-	-
Sonorant	-	-	+	-	-	-	-	-	-	-	-	+	-	-	+	+
Strident	-	-	-	-	-	+	+	-	-	-	-	-	+	+	-	-
Voiced	-	+	+	-	+	-	+	-	+	-	+	+	-	+	+	+

	r	ʃ	ʒ	ʧ	ʤ	j	c	ɟ	ç	ɟ	ɲ	ɰ	ɹ	k	g	x
Anterior	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Back	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+
Continuant	-	+	+	-	-	+	-	-	+	+	-	+	+	-	-	+
Coronal	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-
Del-rel	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-
High	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Labial	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Lateral	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-
Nasal	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-
Sonorant	-	-	-	-	-	+	-	-	-	-	+	+	+	-	-	-
Strident	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-
Voiced	+	-	+	-	+	+	-	+	-	+	+	+	+	-	+	-

	ɣ	ŋ	w	ɭ	ɥ	q	ɠ	ɴ	ʀ	x	ɸ	ħ	ʕ	ʔ	h	ɦ
Anterior	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Back	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-
Continuant	+	-	+	+	+	-	-	-	+	+	+	+	+	-	+	+
Coronal	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Del-rel	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
High	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-
Labial	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-
Lateral	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-
Nasal	-	+	-	-	-	-	-	+	-	-	-	-	-	-	-	-
Sonorant	-	+	+	+	+	-	-	+	+	-	-	-	-	-	-	-
Strident	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-
Voiced	+	+	+	+	+	-	+	+	+	-	+	-	+	-	-	+

[s]: sip	[z]: zip	
	[l]: lip	
	[ɹ]: later	
[ʃ]: shade	[ʒ]: illusion	
[tʃ]: chew	[dʒ]: Jew	
	[j]: yellow	
	[ɹ̩]: onion	
	[ʌ]: million	
	[ɹ̩]: rye	
[k]: Rick	[g]: rig	[ŋ]: ring
	[w]: why	
	[h]: high	

Problem Set 1.1: Problems of Orthography and Aspiration in English

1. We noted in the text that there is one consonant phoneme of English that can occur only in syllable-final position ([ŋ]). But there is also one consonant phoneme of English that can occur only in syllable-initial position. Which one?
2. Take the sounds we have written phonetically as:

[p] [tʃ] [ʒ] [f]

and find as many ways as you can to represent these sounds in writing (orthographically) in English. Give a word to exemplify each way. For some of these sounds there may be several ways. For all there are at least a few. Don't go nuts. Just do the best you can.

3. (a) What are the two consonant sounds of English (sounds that you can find displayed in the appendix) that we always spell identically? Make sure that the sounds you choose cannot be written in English orthography with any other letters. (Hint: Be careful to check for sounds that can be written as a single or a double letter.)

(b) How are they spelled?

- (c) Given that these two sounds are always spelled identically, how do we know which sound to make when we read a new word with the letter(s) corresponding to these two sounds? Take a stab at the answer to (c). You won't be graded on this one.

4. Transcribe the intervocalic sound(s) (the sound(s) between the vowels) in *pizza* by using IPA symbols. (Listen carefully as you pronounce the word—you may find some surprises.) Does this sound occur in word-initial position in English? Is it an oral stop, fricative, affricate, liquid, or nasal?

5. We noted in the text that voiceless stops always get aspirated in English in initial position of stressed syllables, but typically not in syllable-final position. What happens in syllable-medial position (that is, when the voiceless stop is neither the first nor the last sound of the syllable)? Make a list of six words you used to answer this question. (You need six. You have three voiceless stops to test, and you need to test each in syllable-onset and syllable-coda clusters. Remember that the onset is the part of the syllable that precedes the nucleus and the coda is the part that follows the nucleus. I'll start you off with both types of examples for /p/: *space*, *apes*.)

6. Find a native speaker of English who pronounces the words *witch* and *which* differently. For most speakers of English, neither of these words involves an *h*, but for some one of them does. Is orthography a help or a hindrance here? Why?

There is something unusual about the [w] in the word *which* involves an [h]. What is it?

7. Say the words *cool* and *key*. There is a difference between the initial consonants. Describe that difference in terms of the place of articulation and the articulator. Is it a redundant difference or a distinctive difference? If you say it's redundant, give the contexts for each of the two sounds.

8. We have learned that there is a discrepancy between orthography and pronunciation in English, and in the text we attributed this discrepancy to the fact that languages change over time and our orthography reflects the pronunciation of an earlier time.

I now have a question that I hope you will discuss in class: Should we adopt new orthography conventions that reflect today's pronunciations? In discussing this, make a list of the advantages and disadvantages of such spelling reform.

Rising illiteracy in the United States is a problem (as of 1995, when I'm writing this). How would learning to read be affected by spelling reform? Think about what would happen when you read books in which characters from Atlanta, Georgia, are speaking, as opposed to what would happen when you read books in which characters from Brooklyn, New York, are speaking. Will your reading experience be enriched or not?

But also be sure to think of what would happen to you if you learned to read in the IPA and then went to the library to read a book published before the spelling reform. And listen to the variation in pronunciations of all the people around you. Whose pronunciation will you pick as the standard to put in your dictionary? What will happen to the new orthography as spoken language continues to change over time?

Problem Set 1.2: Problems of Orthography in Italian and Japanese

1. Look at the following Italian words in their orthography and their IPA transcription.

cera [tʃera]	culla[kul:a]	già[dʒa]
cosa [koza]	amiche [amike]	pace [patʃe]
chiesa [kʲeʒa]	Giorgio [dʒordʒo]	Giulio [dʒulʲo]
ago [ago]	gelo[dʒelo]	chi [ki]
righe [rige]	ghiro [giro]	faccia [fatʃ:a]
amici [amitʃi]	cane [kane]	gufo [gufo]
bacio [batʃo]	ciuccio [tʃuʃ:o]	che [ke]
gatto [gat:o]	Parigi [paridʒi]	laghi [lagi]

(Recall that the colon indicates a geminate. That is, a sound followed by a colon takes up two timing slots. Also, I have used the vowel [a] here and throughout this book. Many Italians, however, use [ɑ] instead.)

Describe the orthographic use of the letter "h" in these examples.

Now consider the orthographic "i" that does not correspond to any vowel in the IPA transcription. Describe the orthographic use of this "i."

2. This question is just for fun. There are no spelling bees in Italy. After about second grade almost no one makes a spelling error. Make a guess why not.

If you know Spanish, do you think there are spelling bees in Spain or Cuba or Argentina?

If you know French, do you think there are spelling bees in France?

3. Now consider the following data on the orthography of Italian vowels (where I have not indicated redundant vowel length):

sella [sél.la]	sera [sé.ra]	slegatura [zle.ga.tú.ra]
dente [dén.te]	bene [bé.ne]	ragazze [ra.gát.se]
esame [e.zá.me]	ieri [íé.ri]	gheppio [gép.pjɔ]
umore [u.mó.re]	che [ké]	Ercole [ér.kole]

(These transcriptions are from Zingarelli 1970. They represent a particular variety of Italian—presumably that spoken by Nicola Zingarelli. Italian, like English, shows variation in vowel quality as you move from one variety of Italian to another. If you speak Italian and your speech is not represented by these transcriptions, please treat the data above as though they were from a language unknown to you and use these transcriptions just for the sake of doing this problem set.)

In this list geminate consonants are repeated so that you can see syllable boundaries, indicated by periods. In *ragazze* the written double “z” corresponds to a long voiceless affricate, which I have indicated with a [t] closing one syllable and an [s] opening the next syllable.

Can you predict when a written “e” will correspond to an [e] and when it will correspond to an [ɛ]? Please assume that these data are representative. Ask yourself whether the sounds around the “e” allow you to predict its pronunciation. Ask yourself whether the shape of the syllable allows you to predict its pronunciation. Ask yourself whether the assignment of stress allows you to predict its pronunciation.

Based on your answer, do you believe that [e] and [ɛ] are allophones of a single sound or that they are different phonemes?

Many speakers of Italian claim that they have no minimal pairs in their speech that contrast these two vowels. Does this fact surprise you? If so, why? If not, why not? Many other speakers claim they do have at least one minimal pair. A common example given is:

venti ‘twenty’ [vén.ti] venti ‘winds’ [vén.ti]

(The facts presented here may make you unhappy, because they don’t go as neatly with the discussion in the text as one would hope. Language is like that. The distribution of [e] and [ɛ] is affected by historical facts which the ordinary modern speaker may no longer have access to. Thus a distribution which may be predictable at one point in time can become unpredictable at a later period of time. This does not mean Italian is capricious. It means languages change.)

4. Consider the following words written in symbols from the Japanese Katakana. In this writing system, each symbol stands for an entire syllable. Under each symbol I have written a Romanization of that syllable that is very close to IPA. (The differences between this Romanization and IPA are irrelevant to this problem.)

ト ラ イ	ト [~] ラ イ	ロ - ス	ロ - ス [~]
to ra i	do ra i	ro o su	ro o zu
コ - ス ト	ゴ [~] - ス ト		
ko o su to	go o su to		

(Actually, the words *torai* and *dorai* consist of two syllables, but three moras. You will learn what a mora is in the section “Syllable Types” in Chapter 2. Here each symbol is showing us a mora, rather than a syllable and this writing system, which has traditionally been called a SYLLABARY, is actually a “MORARY”—a term I put in quotes because I just made it up. That detail should not interfere with your ability to answer the questions.)

- There is a diacritic that indicates voicing. What is it?
- One voiced consonant here does not have that diacritic. Suggest a reason why.
- Do you think voicing is a distinctive feature in Japanese? Why? (I have not given you the meaning of any of these words, so your answer has to be based on what you think a reasonable relationship would be between a writing system and the phonemes of that language.)

Now consider the following Japanese words. I have given you the word in katakana, a Romanization, and its translation into English.

ミスリー	テリー	ホール	スкуль
misutaa	terii	hooru	sukuuru
‘Mr.’	Terry	Hall	‘school’
テリー	リー	スкуль	‘Terry Lee School’
terii	rii	sukuuru	
メリー	ホール	スкуль	‘Mary Hall School’
merii	hooru	sukuuru	

These examples are all of English words or names that have been borrowed into Japanese.

- Japanese conflates two phonemes of English into one. Which two in English get conflated into which one in Japanese? How is this a natural conflation?

(e) What does the symbol - indicate in Japanese katakana?

- Do you think vowel length is distinctive in Japanese? Why or why not?

*Problem Set 1.3: Articulatory Phonetics

- Give the (a) place of articulation

(b) articulator

and (c) manner of articulation

for the sounds corresponding to the italics in these words:

fat *long* *garage* *tunnel* *thanks*

- For the sounds of the italicized parts of the words in 1 give the voiceless or voiced counterpart if such a counterpart exists. Say whether it is voiced or voiceless.

3. In order for a person to be able to carry a tune in a single, particular sound, what features must that sound have?
4. Give the (a) place of articulation
(b) articulator
(c) manner of articulation
and (d) voicing

of the sounds corresponding to the italics in the words that follow.

window young house

Now give other words in which at least one of (a)–(d) is different for each of these three sounds. That is, you are trying to find allophones which vary from the three you have described by place of articulation, or by articulator, or by manner, or by voicing. (Notice that here you will be using orthography as your guide. That is, you will consider other words with the written letters “w,” “y,” and “h.” At this point you probably don’t trust orthography very much, but while orthography for English has an extremely poor relationship to phonetic transcription, or PR, it certainly has a much better relationship to the underlying phonemic representation—that is, UR.)

Which one out of (a)–(d) did you most quickly and easily find variations on for each of these three sounds? Which one out of (a)–(d) did you find the most variations on for each of these three sounds? (Of course, for voicing we have only a two-way split: a sound is voiced or voiceless. But for place and articulator and manner, there are several possibilities.) In this way, what other group of sounds are these three glides most like—the nasals, the liquids, the fricatives, or the oral stops?

5. Say the English word *onion*. Consider the intervocalic nasal. How is this nasal different from the nasal in *nose*? Compare their place of articulation. The symbol for this sound is [ɲ]. Is [ɲ] an allophone of /n/ or a separate phoneme? Why?

Now consider the following data on Italian nasals. Before each word I have given the sound of the initial consonant:

[n] nomi ‘names’
[ɲ] gnomi ‘gnomes’

Just based on these data, is it possible to tell whether [n] and [ɲ] are separate phonemes or allophones in Italian? Why or why not? If you say it is possible, what are they: separate phonemes or allophones?

6. A typical pattern for languages is to have nasal consonants only at places of articulation where oral stops also occur. That is, nasals are generally homorganic to oral stops. Give a reason for this fact using what you know about the articulatory apparatus.

7. Say *information* slowly and then quickly. Speed should affect your pronunciation. (Hint: Pay particular attention to the sounds in the first two syllables of the word.)

Give two transcriptions of *information* (and you will need to consult the appendix to do this, since you’ll need a symbol that wasn’t introduced in the text):

- (a) one for fast speech,
and (b) one for slow speech.

- (c) Consider the difference in the two transcriptions in (a) and (b) with respect to the

first nasal sound in *information*. Why do you think this difference occurs? (Notice that nothing like this happens to the second nasal sound in *information*.) As you answer this, be sure to think about the places of articulation and the articulators of the two nasals and of the sound(s) around them. How does this confound your answer to 6 above?

8. Many varieties of Spanish do not have [ʃ] but they do have [tʃ]. What does this suggest about this affricate in those varieties of Spanish? Is it a single phoneme or two?

9. English uses an egressive pulmonic airstream for most sounds in speech. However, there are situations in which we use other types of airstreams. Give an example of a situation in which one might use the following types of sound:

- (a) an ingressive pulmonic voiceless labiodental fricative
- (b) a velaric alveolar lateral

State as precisely as you can what is happening to the airstream as you make these sounds.

10. I want you to think about the application of your knowledge of articulatory phonetics to education of the Deaf. Here's the situation. You are a kindergarten teacher and you have a Deaf child in your classroom. You are supposed to try to help this child learn how to speak English. Assume that the child has no hearing whatsoever (although many people who are Deaf have partial hearing). Assume also that others have worked with this child on speaking before, so the idea of speaking is not new to the child. Explain what you would do and why if you encountered the events described here. (Assume you have the ordinary equipment of a kindergarten. So you have a blackboard, but you have no phonetics lab, for example.)

- (a) You are teaching the alphabet. The child says [bi] as the name of the letter *p*.
- (b) You put up a picture of a doe. The child calls it "no."
- (c) The child says, "Ep," when she clearly means *Help*.
- (d) The child says [wɑɪ] when she means *water*.

NOTE: Please do not take this problem as an endorsement for oral programs for the Deaf. I include this problem simply because oral programs exist and because a basic knowledge of phonetics can give minimal aid to a teacher in such situations.

11. There are two major types of ventriloquism. One is known as "near" ventriloquism. That's when you try to give the impression that the voice coming from you is really coming from some nearby object, such as a dummy. The main object is to move your lips as little as possible and simultaneously move the mouth of the dummy so that the illusion is created. To do this, try retracting your tongue root and talking with your lips slack and slightly apart.

The second type of ventriloquism is called "distant" ventriloquism: You try to give the impression that the voice coming from you is really coming from some distant object, perhaps across the room. To do this, you need to do all you did in near ventriloquism plus add on a change of pitch and phonation. Try closing the posterior section of your vocal cords and, thus, talking with a creaky voice, while still keeping the tongue retracted and the lips slack and slightly apart. If that sounds too awful, try raising the larynx and constricting the muscles of the throat, while keeping the tongue root retracted. This should result in a much higher pitched, muffled voice.

If you get good at it, figure out what you're really doing (since I'm only guessing here, from what I've been able to pick up in encyclopedias), and write a description of it. And throw a party at which you perform. Why not?

Problem Set 1.4: Articulatory Phonetics, Syllable Structure, and Vowels

(NOTE: Question 7 on this problem set is more difficult than the others.)

1. Make a list of monosyllabic words of English that begin with three consonants. What is the first consonant? Characterize the possible second consonants in terms of a feature bundle. Characterize the possible third consonants in terms of a feature bundle. In light of the sonority hierarchy, which aspects of these consonant clusters are surprising and which are expected?
2. Compare these Italian and Spanish cognates:

<i>Italian</i>	<i>Spanish</i>	
Spagna [spa.ɲa]	España [es.pa.ɲa]	'Spain'
stato [sta.to]	estado [es.ta.do]	'state'
scuola [skʊɔ.la]	escuela [es.kʷe.la]	'school'
signore [si.ɲo.re]	señor [se.ɲor]	'mister'
schiavo [skʲa.vo]	esclavo [es.kla.vo]	'slave'
blasfema [blas.fe.ma]	blasfemia [blas.fe.mia]	'blasphemy'
tragico [tra.ɟi.ko]	tragico [tra.hi.ko]	'tragic'

Assume the above data are representative. Pay attention to consonant clusters. Does Italian organize its syllable structure according to the sonority hierarchy? Does Spanish? How does Spanish manage to do this, in contrast to Italian?

3. Transcribe the following words:

fate fit fat feet fought foot fun phone
 food fete (it rhymes with *met*)

Whether or not your variety of English uses [ɔ], you should have a different vowel in each of these words (and a diphthong in some of them).

4. Out of the twelve vowels of English discussed in this chapter, which two are missing from the words in 3? (Of course, if you do not have [ɔ] in your speech, then you have only eleven vowels—and only one of the eleven vowels you use is missing from 3. Which one?)

Which of the following words contain the vowel(s) you find missing in 3? Transcribe only those words.

book rate pot fainted pawn

5. Transcribe *fight*. Is there a diphthong here? If so, does it contain a rising or falling glide?
6. Write the following in regular English spelling.

- (a) [spɪf sáɪntɪsts hæv dævələpt məʃɪnz fɔɪ ənə́lɪsɪs]
- (b) [ɪɪvɒ nom ʃámski hæz ə bɪθdeɪ]
- (c) [pʰliɪz ɪɪd wətévɪ θóvni mɔ́ɪsən ɪɪts]

7. These words are not words of English:

[bræft] [ltos] [vléɪ.vɪ] [zo.ɪf]
[ɔʒí.məh] [kuʊl]

Which of them do you think could become English words (that is, which could be added to our LEXICON—our mental list of all the words of our language) and which could not? Which of them do you think could not be words of any language? State your reasons for both answers. (Be careful. Words without a period inside them are monosyllabic.)

Problem Set 1.5: Features

1. Make a list of all the features discussed in Chapter 1.

(a) Which ones of them deal directly or indirectly with place of articulation?

(b) Which features entail which other features? For example, if a sound is [+sonorant], is there any other feature it has in the unmarked case? Go through all the features, asking yourself which ones entail which others. (This is a tedious exercise, but what you learn will be helpful to you throughout Chapter 2, particularly when you meet Geometric Phonology.)

(c) Which features are incompatible? For each feature, tell which other features (if any) it is incompatible with.

(d) Give the feature bundle for a palatal voiceless oral stop.

(e) List all the distinctive sounds of English that are [+strident].

(f) What feature(s) do these sounds have in common?

[i] [j] [ʃ] [ɔʒ]

(g) What sound or sounds have these features: [+back, +round]?

(h) Give a feature bundle that will pick out all stops other than the nasals.

2. In Figure 1.10 you see nine signs of American Sign Language, the manual-visual language used by Deaf communities in the United States of America. Only one hand moves in these signs—the dominant hand (right hand for right-handed people; left hand for left-handed people).

These signs show three separate characteristics of signs that distinguish one sign from another. We can think of these as distinctive features of signs. Organize these nine signs into three groups, where each group forms a minimal triplet. That is, for Group 1, hold Feature 1 steady across the three signs and vary Features 2 and 3. For Group 2, hold Feature 2 steady and vary Features 1 and 3. And for Group 3, hold Feature 3 steady and vary Features 1 and 2. Label your groups 1 through 3.

(a) What is Feature 1? Feature 2? Feature 3?

(Clearly you have to come up with your own terms for these features, since they are not any of the distinctive features that we discussed for oral languages. But simply define your terms and then use them.)

Now please try to answer these additional questions. On some of these questions you will be asked to make conjectures. The object, however, is to make intelligent, educated guesses. Please don't forget to consider the suprasegmentals of language.

(b) One way to make a sign emphatic in ASL is to make the movement rapid and

Figure 1.10. Nine signs of ASL, demonstrating three distinctive features.

tense. What would be a phonetic analogue in spoken English? (And, by the way, what would be a graphemic analogue in written English?)

(c) Your hands can make many shapes. Yet only some are actually used in sign languages. In Figure 1.11 you can see four hand shapes. One of these four handshapes does not occur in ASL. Guess which one and explain why you chose it. What would be an analogous fact about spoken English? (I'm asking you here to speculate about why possible handshapes and sounds are not likely to be used as part of a linguistic system. To see the analogy, think of sounds the vocal tract can make, but that aren't used in natural language, and discuss why they aren't used.)

Figure 1.11. Four handshapes, of which one does not occur in ASL.

(d) ASL signs are made in a highly restricted signing space in front of the body, which runs from the top of the head down to the waist, and out to each side as far as the reach of the arms with the elbows bent. What that means is that signs will be made in this signing space and not, for example, in front of the knees or behind the head. Why is this a natural signing space? What is a phonetic analogue in spoken language?

Problem Set 1.6: Acoustic Phonetics without a Laboratory

(NOTE: Question 5 on this problem set is difficult.)

1. Which of the following makes a periodic sound and which makes an aperiodic sound? Explain your answer.

a waterfall a drum

(If you know the difference between pitched drums, like timpani drums, and unpitched drums, like snare drums, you might want to elaborate on your answer. If not, please just give me the answer you know I'm asking for.)

Figure 1.12. Spectrogram for the sounds of a tuning fork hitting a metal table.

2. It is common knowledge that children generally have higher-pitched voices than adult women, who have higher-pitched voices than adult men. In fact, prepubescent voices have a range from 200 to 500 Hz; adult female, from 150 to 300 or even 400 Hz; and adult male, from 80 to 200 Hz. Explain these facts.
3. In Figure 1.12 you have a spectrogram of the sounds that occurred when a tuning fork hit a metal table. We have here the smack of the fork on the table; the pitch of the fork (which is the same fork used in Figure 1.3, so the note here is A); and the pitch of the resonances set up in the table. Label each one and explain how you knew which was which. (The horizontal lines are spaced 1,000 Hz apart. If you are having trouble telling the resonance of the table from the tuning fork, look back at Figure 1.3 and figure out what the Hz of the tuning fork is.)
4. In Figure 1.13 we have a spectrogram of an alveolar nasal followed by a dental nasal. How many (significantly intense) formants do nasals have (assuming [n] is representative of all nasals)? In comparison with vowels, which formant is missing? (If you're having trouble seeing this, compare Fig. 1.13 to Fig. 1.5.) Describe any acoustic differences you see between the alveolar and dental nasals—although these differences are decidedly slight. Would you expect there might be a language which distinctively contrasts dental to alveolar nasals? Why or why not?

Figure 1.13. Spectrogram for an alveolar nasal followed by spectrogram for a dental nasal.

Figure 1.14. Spectrogram of an Italian sentence.

5. In Figure 1.14 you have a spectrogram of a sentence of Italian. Consider the following questions. Tell me which of these questions is theoretically answerable (if you have the experience) from looking at a spectrogram?

- (a) How many vowels are in this utterance?
- (b) Which are the stressed syllables?
- (c) Are there any geminates?
- (d) Are there any syllabic consonants?
- (e) Are there any oral stops?
- (f) Are there any diphthongs?
- (g) Are there any nasals?
- (h) Are there any affricates (stops released as fricatives)?
- (i) Are there any fricatives that are clearly not the release part of a stop—i.e., that are not part of affricates?

For each question that you say is answerable, answer that question and explain what you looked for in the spectrogram to arrive at your answer. Now label the segments on the spectrogram appropriately wherever possible in light of your answers to these questions.

Problem Set 1.7: Acoustic Phonetics with a Laboratory

1. Do a wide band spectrogram of eight minimal pairs. Does the feature of [+strident] correspond to an acoustic reality? It is your job to find an appropriate set of eight words.
2. Reading spectrograms is confounded by the fact that individual speakers bring their own speech peculiarities to utterances. Record a speaker with a very similar linguistic background to your own—in terms of where the speaker grew up, how old that speaker is, the socioeconomic and educational class that speaker belongs to, the ethnic or racial group that speaker belongs to, gender, and anything else you might guess is pertinent. Now record yourself saying the same sentence(s) that speaker said. Do a wide-band spectrogram and point out any individual peculiarities you or the other person has.
3. If consonants are merely ways to begin or end vowels, tell what the effect of a velar consonant is on ending a vowel. You can do this by looking at appropriate spectrograms.

Make sure that you show it is velars only (and not any other kind of consonant) that cause this effect. And make sure that you show velars cause it on every vowel (not just certain kinds of vowels). You can reduce your job by assuming that voiceless and voiced counterparts of a consonant have the same effect—so test only one of them. You can also reduce your job by assuming that front vowels will be affected as a group, as will back vowels. So test only one representative from the front vowels and only one representative from the back vowels.

4. Lip rounding has an effect on formants. What is that effect? Support your answer by presenting and discussing the spectrograms of at least the vowels [i] vs. [y] and [u] vs. [ʊ].

Now look at the spectrograms for [u] and [o]. Which has more lip rounding? How do you know?