

Chapter 2

The Lowdown on the Science of Speech Sounds

In This Chapter

- ▶ Spelling out what phonetics and phonology are
- ▶ Understanding how speech sounds are made
- ▶ Recognizing speech anatomy, up close and personal

Phonetics is centrally concerned with speech, a uniquely human behavior. Animals may bark, squeak, or meow to communicate. Parrots and mynah birds can imitate speech and even follow limited sets of human commands. However, only people naturally use speech to communicate. As the philosopher Bertrand Russell put it, “No matter how eloquently a dog may bark, he cannot tell you that his parents were poor, but honest.”

In this chapter, I introduce you to the basic way in which speech is produced. I explain the source-filter theory of speech production and the key parts of your anatomy responsible for carrying it out. I begin picking up key features that phoneticians use to describe speech sounds, such as voicing, place of articulation, and manner of articulation.



Phoneticians *transcribe* (write down) speech sounds of any language in the world using special symbols that are part of the International Phonetic Alphabet (IPA). Throughout this book, I walk you through more and more of these IPA symbols, until transcription becomes a cinch. For now, I am careful to indicate spelled words in quotes (such as “bee”) and their IPA symbols in slash marks, meaning broad transcription, such as /bi/. (Refer to Chapter 3 for in-depth information on the IPA.)

Defining Phonetics and Phonology

Phonetics is the scientific study of the sounds of language. You may recognize the root *phon-* meaning sound (as in “telephone”). However, phonetics doesn’t refer to just any sort of sound (such as a door slamming). Rather, it deals specifically with the sounds of spoken human language. As such, it’s part of the larger field of linguistics, the scientific study of language. (Check out *Linguistics For Dummies* by Rose-Marie Dechaine PhD, Strang Burton PhD, and Eric Vatikiotis-Bateson PhD [John Wiley & Sons, Inc.] for more information.)

Phonetics is closely related to *phonology*, the study of the sound systems and rules in language. The difference between phonetics and phonology can seem a bit tricky at first, but it’s actually pretty straightforward. Phonetics deals with the sounds themselves. The more complicated part is the rules and systems (phonology). All languages have sound rules. They’re not explicit (such as “Keep off the grass!”), but instead they’re implicit or effortlessly understood.



To get a basic idea of phonological rules, try a simple exercise. Fill in the opposite of these three English words. (I did the first one for you.)

tolerant	<i>intolerant</i>
consistent	_____
possible	_____

You probably answered “*inconsistent*” and “*impossible*,” right? Here’s the issue. The prefix “in” means “not” (or opposite) in English, so why does the “in” change to “im” for “*impossible*?” It does so because of a sound rule. In this case, the phonological rule is known as *assimilation* (one sound becoming more like another). In this example, a key consonant changes from one made with the tongue (the “n” sound) to one made at the lips (the “m” sound) in order to match the “p” sound of “possible,” also produced at the lips. The effect of this phonological rule is to make speech easier to produce. To get a feel for this, try to say “*in-possible*” three times rapidly in succession. Now, try “*impossible*.” You can see that saying “impossible” is easier.

I focus more discussion on phonology in Chapters 8 and 9. Now you just need to know that phonological rules are an important part of all spoken languages. One of the key goals of phonology is to figure out which rules are language-specific (applying only to that language) and which are universal.

Phoneticians specialize in describing and understanding speech sounds. A phonetician typically has a good ear for hearing languages and accents, is skilled in the use of computer programs for speech analysis, can analyze speech movement or physiology, and can transcribe using the IPA.

Because phonetics and phonology are closely allied disciplines, a phonetician typically knows some phonology, and a phonologist is grounded in phonetics, even though their main objects of study are somewhat different.



Phonetics can tell people about what language sounds are, how language sounds are produced, and how to transcribe these sounds for many purposes. Phonetics is important for a wide variety of fields, including computer speech and language processing, speech and language pathology, language instruction, acting, voice-over coaching, dialectology, and forensics.

A big part of a person's identity is how you sound when you speak — phonetics lets you understand this in a whole new way. And it's true what the experts say: Phonetics is definitely helpful for anyone learning a new language.

Sourcing and Filtering: How People Make Speech

Scientists have long wondered exactly how speech is produced. Our current best explanation is called the *source-filter theory*, also known as the *acoustic theory of speech production*. The source-filter theory best explains how speech works.



The idea behind this theory is that speech begins with a breathy exhalation from the lungs, causing raw sound to be generated in the throat. This sound-generating activity is the *source*. The source may consist of buzzing of the *vocal folds* (also known as the *vocal cords*), which sounds like an ordinary voice. The source may also include hissing noise, which sounds like a whisper. The movement of the lips, tongue, and jaw (for oral sounds) and the use of the nose (for nasal sounds) shapes this raw sound and is the part of the system known as the *filter*.

The raw sound is filtered into something recognizable. A filter is anything that can selectively permit some things to pass through and block other things (kind of like what your coffee filter does). In this case, the filter allows some frequencies of sound to pass through, while blocking others.

After raw sound is created by a buzzing larynx and/or hissing noise, the sound is filtered by passing through differently shaped airway channels formed by the movement of the speech *articulators* (tongue, lips, jaw, and velum). This sound-shaping process results in fully formed speech (see Figure 2-1 for what this looks like).

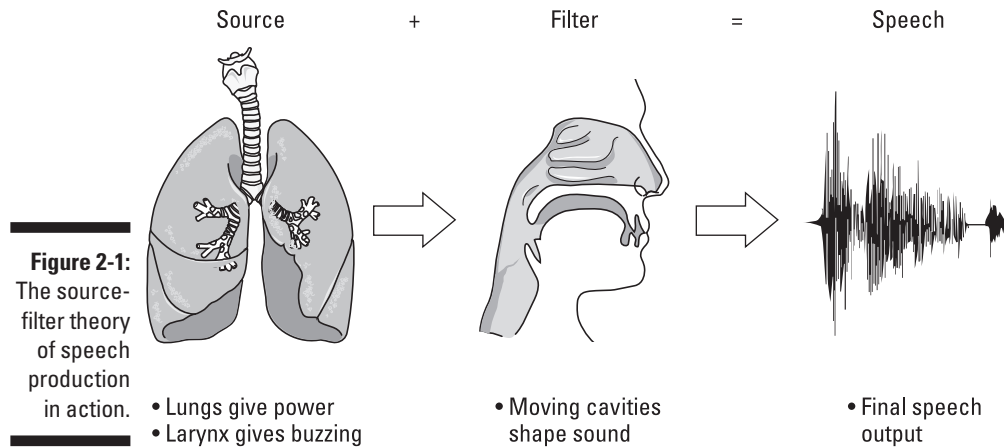


Illustration by Wiley, Composition Services Graphics

Let me give you an analogy to help you understand. The first part of the speaking process is like the mouthpiece of a wind instrument, converting air pressure into sound. The filter is the main part of a wind instrument; no one simply plays a mouthpiece. Some kind of instrument body (such as a saxophone or flute) must form the musical sound. Similarly, you start talking with a vibrating source (your vocal folds). You then shape the sound with the instrument of your moving articulators, as the filter.



Here are a few other important points to remember with the source-filter theory:

- ✓ The source and filter are largely independent of each other. A talker can have problems with one part of the system, while the other part remains intact.
- ✓ The voicing source can be affected by laryngitis (as in a common cold), more serious disease (such as cancers), injuries, or paralysis.
- ✓ An alternative voicing source, such as an external artificial larynx, can provide voicing if the vocal folds are no longer able to function.
- ✓ The sources and filters of men and women differ. Overall, men have lower voices (different source characteristics) and different filter shapes (created by the mouth and throat passageways) than women.

Thankfully, people never have to really think about making these shapes. If so, imagine how people would ever be able to talk. Nevertheless, this theory explains how humans do talk. It's quite different than, say, rubbing a raspy limb across your body (like the katydid) or drumming your feet on the ground (like the prairie vole cricket) to communicate.

Gunnar Fant and the source-filter theory

The source-filter theory of speech production was the brainchild of Gunnar Fant (1919–2009), a pioneering Swedish professor of speech communication. After earning a master's degree in engineering at Stockholm at the end of World War II, Fant began to apply this knowledge to analyze and synthesize speech sounds. His doctoral dissertation, the *Acoustic Theory of Speech Production*, soon became an international standard. Fant's research led to the development of whole new technologies, including

computer speech synthesis, and helped make phonetics more available to a variety of professions. At age 81, while still working actively on phonetics research, Fant wrote in "Half a Century in Phonetics and Speech Research,":

"Mankind is making much progress in mapping the genetic code. We need some of the same patience and persistence in mapping the speech code."

Getting Acquainted with Your Speaking System

Although most people speak all their lives without really thinking about how they do it, phonetics begins with a close analysis of the speaking system. This part of phonetics, called *articulatory phonetics*, deals with the movement and physiology of speech. However, don't fear — you don't need to be a master phonetician to get this part of the field. In fact, the best way is to pay close attention to your own tongue, lips, jaw, and velum when you speak. As you get better acquainted with your speaking system, the basics of articulatory phonetics should become clear.

Figure 2-2 shows the broad divisions of the speaking system. Researchers divide the system into three levels, separated at the larynx. The lungs, responsible for the breathy source, are below the larynx. The next division is the larynx itself. Buzzing at this part of the body causes voiced sounds, such as in the vowel "ah" or "hot" (written in IPA characters as /a/) or the sound /z/ of "zip." Finally, the parts of the body that shape sound (the tongue, lips, jaw, and velum) are located above the larynx and are therefore called *supralaryngeal*.

In the following sections, I delve deeper into the different parts of the speech production system and what those parts do to help in the creation of sound. I also walk you through some exercises so you can see by doing — feeling the motion of the lungs, vocal folds, tongue, lips, jaw, and velum, through speech examples.

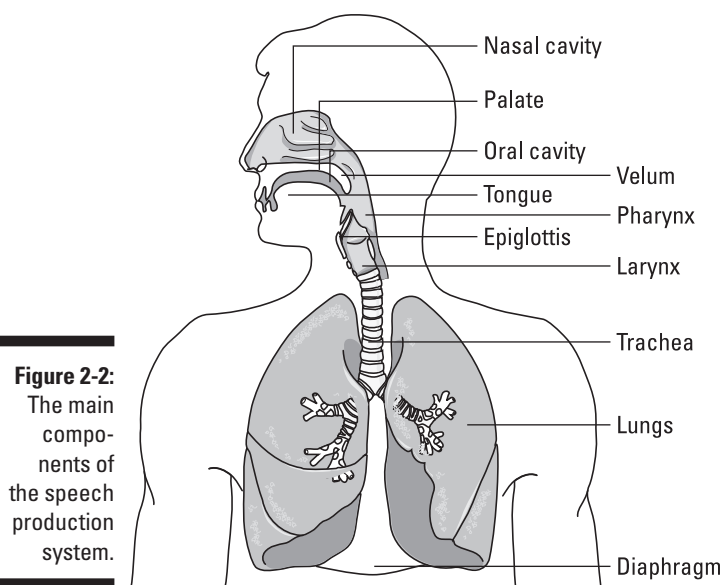


Figure 2-2:
The main components of the speech production system.

Illustration by Wiley, Composition Services Graphics

If you're a shy person, you may want to close the door, because some of these exercises can sound, well, embarrassing. On the other hand, if you're a more outgoing type, you can probably enjoy this opportunity to release your inner phonetician.

Powering up your lungs

Speech begins with your lungs. For anyone who has been asked to speak just after an exhausting physical event (say, a marathon), it should come as no surprise that it can be difficult to get words out.

Lung power is important in terms of studying speech sounds for several reasons: Individuals with weakened lungs have characteristic speech difficulties, which is an important part of the study of speech language pathology. Furthermore, as I discuss in Chapter 10, an important feature of speech called *stress* is controlled in large part by how loud a sound is — this, in turn, relates to how much air is puffed out by the lungs.

The role of the lungs in breathing and speech

Your lungs clearly aren't designed to serve only speech. They're part of the respiratory system, designed to bring in oxygen and remove carbon dioxide. Breathing typically begins with the nose, where air is filtered, warmed, and

moistened. Air then moves to the *pharynx*, the part of the throat just behind the nose and into the *trachea*, the so-called *windpipe* that lies in front of the esophagus (or the *food tube*). From the trachea, the tubes split into two bronchi (left and right), then into many *bronchioles* (tiny bronchi), and finally ending up in tiny air sacs called *alveoli*. The gas exchange takes place in these sacs.



When you breathe for speaking, you go into a special mode that is very different than when you walk, run, or just sit around. Basically, speech breathing involves taking in a big breath, then holding back or checking the exhalation process so that enough pressure allows for buzzing at the larynx (also known as *voicing*). If you don't have a steady flow of pressure at the level of the larynx, you can't produce the voiced sounds, which include all the vowels and half of the consonants.

Young children take time to get the timing of this speech breathing right; think of how often you may have heard young kids say overly short breath-group phrases, such as this example:

“so like Joey got a . . . got a candy and a . . . nice picture from his uncle”

Here the child talker quite literally runs out of breath before finishing his thought.

Some interesting bits about the lungs can give you some more insight into these powerhouse organs:

- ✓ They're light and spongy, and they can float on water.
- ✓ They contain about two liters (three quarts) of air, fully inflated.
- ✓ Your left and right lungs aren't exactly the same. The left lung is divided into two lobes, and the lung on your right side is divided into three. The left lung is also slightly smaller, allowing room for your heart.



When resting, the average adult breathes around 12 to 20 times a minute, which adds up to a total of about 11,000 liters (or 11,623 quarts) of air every day.

Testing your own lung power

You can test your lung power by producing a sustained vowel. To test your lung power, sit up, take a deep breath, and produce the vowel /a/, as in the word “hot,” holding it as long as you can. The vowel /a/ is part of the IPA, which I discuss in Chapter 3.

How did you do? Most healthy men can sustain a vowel for around 25 to 35 seconds, and women for 15 to 25 seconds. Next, try the same vowel exercise while lying flat on your back (called being *supine*). You probably can't go on

as long as you did when you were sitting up, and the task should be harder. Due to gravity and biomechanics, the lungs are simply more efficient in certain positions than others. The effect of body position on speech breathing is important to many medical fields, such as speech language pathology.

Buzzing with the vocal folds in the larynx

The *larynx*, a cartilaginous structure sometimes called the *voice box*, is the part of the body responsible for making all voiced sounds. The larynx is a series of cartilages held together by various ligaments and membranes, and also interwoven by a series of muscles. The most important muscles are the vocal folds, two muscular flaps that control the miraculous process of voicing.

Figure 2-3 shows a midsection image of the head. In this figure, you can see the positions of the nasal cavity, oral cavity, pharynx, and larynx. Look to see where the vocal folds and glottis are located. The vocal folds (also known as the *vocal cords*) are located in the larynx. You can find the larynx in the figure at the upper part of the air passage.

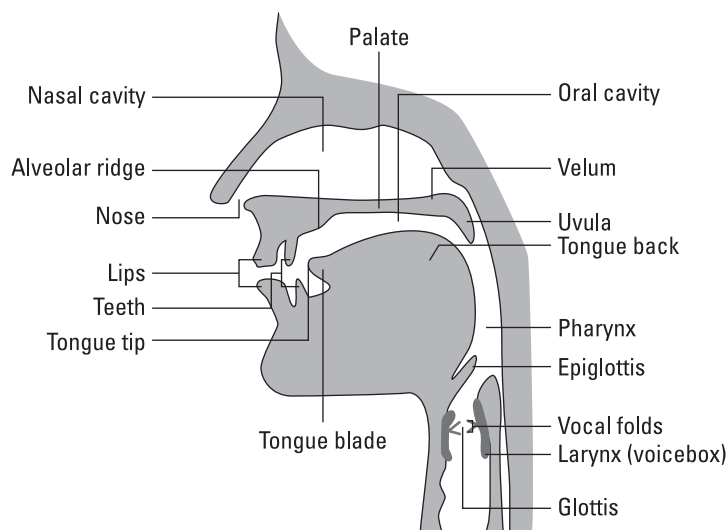


Figure 2-3:
The
midsagittal
view of the
vocal tract.

Illustration by Wiley, Composition Services Graphics

The following sections provide some examples you can do to help you get better acquainted with your larynx and glottis.

Getting a buzz from a different source

A common surgery used for the treatment of laryngeal cancer is *laryngectomy*, which is the complete or partial removal of the larynx and vocal folds. After such a surgery, several methods can be used to help a patient speak. One way is to train patients to use an *electrolarynx*, a mechanical (buzzing) device held against the throat to provide vibrations for speech. For laryngectomy patients, the electro-larynx has the advantage of being simple and accessible pre- and post-operation. However, a disadvantage is the rather mechanical voice that results (see www.youtube.com/watch?v=v55NAjqltEI).

For phonetics students, trying out an electro-larynx is a fun way to really get the idea of the

independence of source and filter. See if you can borrow one from a nearby communication disorders group or clinic. To see how it works, follow these steps:

1. **Place the vibrating membrane against the side of your Adam's apple (laryngeal prominence).**
2. **Turn on the device and silently count to 10.**

If you did it correctly, you'll feel a pleasant buzzing on your neck while the device *voices (phonates)* for you. You may need to try several times to get the coupling just right, so that others can hear you.

Locating your larynx

You can easily find your own larynx. Lightly place your thumb and forefinger on the front of your throat and hold out a vowel. You should feel a buzzing. If you have correctly done it, you're pressing down over the *thyroid cartilage* (refer to the larynx area shown in Figure 2-3) to sense the vibration of the vocal folds while you phonate. If you're male, finding your vocal folds is even more obvious because of your *Adam's apple* (more technically called the *laryngeal prominence*), which is more pronounced in men than women.

Are you happy with your buzzing? Now try saying something else, but this time, whisper. When whispering, switch from a voiced (*phonated*) sound to voiceless. Doing these exercises gives you a good idea of voicing, which is the first of three key features that phoneticians use to classify the speech sounds of the world. (Refer to Chapter 5 for these three key features.)

Voicing is one of the most straightforward features for beginning phonetics students because you can always place your hand up to the throat to determine whether a sound is being produced with a voiced source or not.

Stopping with your glottis

Meanwhile, the *glottis* is the empty space between the two vocal folds when they're held open for breathing or for speech. That is, it's basically an empty hole. Your glottis is probably the most important open space in your body

because it regulates air coming in and out of the lungs. Even if you're otherwise able to breathe just fine, if your glottis is clamped shut, air can't enter the lungs.



Clamping your glottis shut is a dangerous situation, so don't try it for long. Nevertheless, it's fun and instructive to try something called the *glottal stop*, /ʔ/, a temporary closing (also called an *adduction*) of the vocal folds that occurs commonly during speech. Are you ready? Stick to these steps as you try this exercise:

1. Say “uh-oh,” loudly and slowly several times.

Young children like saying this expression as they are about to drop something expensive (say, your new cell phone) on a cement floor.

2. Feel your vocal folds clamp shut at the end of “uh,” and then open again (the technical term is *abduct*) when you begin saying “oh.”

3. Try holding the closing gesture (the adduction) after the “uh.”

You should soon begin feeling uncomfortable and *anoxic* (which means without oxygen) because no air can get to your lungs.

4. Breathe again, please!

I need you alive and healthy to complete these exercises.

5. Practice by saying other sounds, such as “oh-oh,” “ah-ah,” and “eeh-eeh,” each time holding the glottal stop (at will) across the different vowels.

This skill comes in handy when I discuss more about glottal stops used in American English and in different English dialects worldwide in Chapter 18.

Shaping the airflow

Parts of the body filter sound by creating airway shapes above the larynx. Air flowing through differently shaped vessels produces changing speech sounds. Imagine blowing into variously shaped bottles; they don't all sound the same, right? Or consider all the different sizes and shapes of instruments in an orchestra; different shapes lead to different sounds. For this reason, it's important to understand how the movement of your body can shape the air passages in your throat, mouth, and nasal passages in order to produce understandable speech.



Air passages are shaped by the *speech organs*, also known as *articulators*. Phoneticians classify articulators as *movable* (such as the tongue, lips, jaw, and velum) and *fixed* (such as the teeth, alveolar ridge, and hard palate), according to their role in producing sound. Refer to Figures 2-2 and 2-3 to see where the articulators are located.

The movable articulators are as follows. Here you can find some helpful information to understand how each one works:

- ✓ **Tongue:** The tongue is the most important articulator, similar in structure to an elephant's trunk. The tongue is a *muscular hydrostat*, which means it's a muscle with a constant volume. (This characteristic is important in the science of making sound because muscular hydrostats are physiologically complex, requiring muscles to work *antagonistically*, against each other, in order to stretch or bend. Such complexity appears necessary for the motor tasks of speech.) The tongue elongates when it extends and bunches up when it contracts. You never directly see the main part of the tongue (the body and root). You can only view the thinner sections (tip/blade/dorsum) when it's extended for viewing. However, scientists can use imaging technologies such as ultrasound, videofluoroscopy, and magnetic resonance imaging to know what these tongue parts look like and how they behave.
- ✓ **Jaw:** Although classified as a movable speech articulator, the jaw isn't as important as the tongue. The jaw basically serves as a platform to position the tongue.
- ✓ **Lips:** The lips are used mostly to lower vowel sounds through extension. The *lip extension* is also known as *protrusion* or *rounding*. The lips protrude approximately a quarter inch when rounded. English has two rounded vowels, /u/ (as in "boot"), and /ʊ/ (as in "book"). Other languages have more rounded sounds, such as Swedish, French, and German (refer to Chapter 15). These languages require more precise lip rounding than English.

Lips can also flare and spread (widen). This acts like the bell of a brass instrument to brighten up certain sounds (like /i/ in "bead").
- ✓ **Velum:** The *velum*, also known as the *soft palate*, is fleshy, moveable, and made of muscle. The velum regulates the nasality of speech sounds (for example, /d/ versus /n/, as in the words "dice" and "nice"). The velum makes up the rear third of the roof of the mouth and ends with a hanging body called the *uvula*, which means "bundle of grapes," just in front of the throat.

Some parts of the body are more passive or static during sound production. These so-called fixed articulators are as follows:

- ✓ **Teeth:** Your teeth are used to produce the “th” sounds in English, including the voiced consonant /ð/ (as in “those”) and the voiceless consonant /θ/ (as in “thick”). The consonants made here are called *dental*. Your teeth are helpful in making *fricatives*, hissy sounds in which air is forced through a narrow groove, especially /s/, /z/, /f/, and /v/ — like in the words “so,” “zip,” “feel,” and “vote”. Tooth loss can affect other speech sounds, including the affricates /tʃ/ (as in “chop”) and /dʒ/ (as in “Joe”).
- ✓ **Alveolar ridge:** This is a pronounced body ridge located about a quarter of an inch behind your top teeth. Consonants made here are called *alveolar*.



You can easily feel the alveolar ridge with your tongue. Say “na-na” or “da-da,” and feel where your tongue touches on the roof of your mouth.

The alveolar ridge is particularly important for producing consonants, including /t/, /d/, /s/, /z/, /n/, /l/, and /ɹ/, as in the words “time,” “dime,” “sick,” “zoo,” “nice,” “lice,” and “rice.” Many scientists think an exaggerated alveolar ridge has evolved in modern humans to support speech.

- ✓ **Hard palate:** It continues just behind the alveolar ridge and makes up the first two-thirds of the roof of your mouth. It’s fixed and immovable because it’s backed by bone. Consonants made here are called *palatal*. The English consonant /j/ (as in “yellow”) is produced at the hard palate.

Producing Consonants

A *consonant* is a sound made by partially or totally blocking the vocal tract during speech production. Consonants are classified based on *where* they’re made in the articulatory system (place of articulation), *how* they are produced (manner of articulation), and whether they’re *voiced* (made with buzzing of the larynx) or not. These sections discuss the different ways English consonants are made. Remember, each language has its own set of consonants. So English, for example, doesn’t have the “rolled r” found in Spanish, and Spanish doesn’t have the consonant /dʒ/ as in “judge”.

Getting to the right place

Basically consonant sounds use different parts of the tongue and the lips. Figure 2-4 shows a midsagittal view of the head, including the lips, tongue, and the consonantal places of articulation.

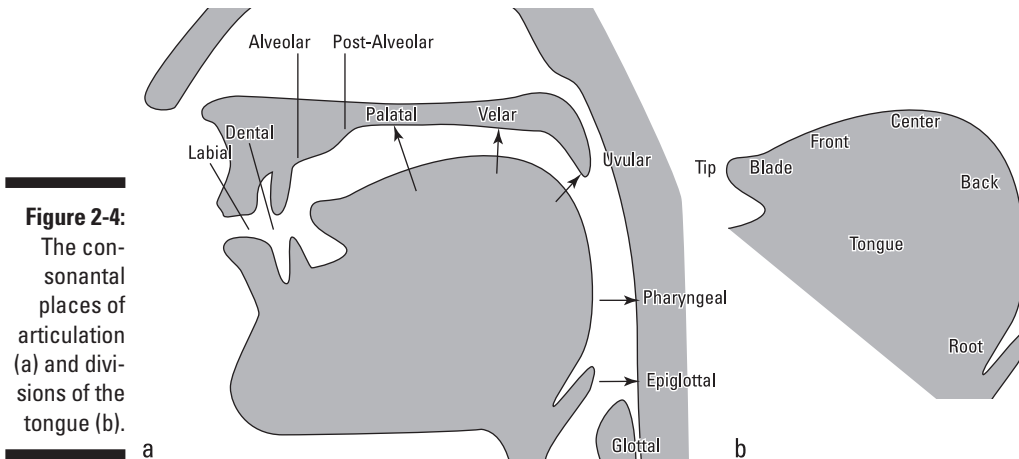


Figure 2-4:
The coronantal places of articulation (a) and divisions of the tongue (b).

Illustration by Wiley, Composition Services Graphics

Notice that these regions are relative; there is clearly no “dotted line” separating the front from the back or marking off the tip from the blade (unless you happen to have a disturbing tattoo there, which I doubt). However, these regions play different functional roles in speech. The tip and blade are the most flexible tongue regions. The different parts of the tongue control the sound in the following ways:

- ✓ **Coronal:** Speech sounds made using either the tip or blade are called *coronal* (crown-like) sounds.
- ✓ **Dorsal:** Speech sounds made using the rear of the tongue are called *dorsal* (back) articulations.



To get an overall feel of what happens when you speak with your lips, tongue, and jaw, slowly say the word “*batik*,” paying attention to where your articulators are as you do so. At the beginning of the word you should sense the separation of the lips for the /b/ (a labial gesture), then the lowering of the tongue and jaw as you pronounce the first syllable. Next, the front of your tongue will rise to make (coronal) contact for the /t/ of “*tik*.” When you reach the end of “*tik*,” you should be able to detect the back (dorsum) of your tongue making (velar) contact with the roof of your mouth for the final /k/ sound.



However, phoneticians typically need to know more detail about where sounds are made than just which parts of the tongue are involved. The following list details the English places of articulation for consonants:

- ✓ **Bilabial:** Also called *labial*, sounds made with a constriction at the lips are very common in the languages of the world. Say “*pat*,” “*bat*,” and “*mat*” to get a good feel for these sounds. Because the lips are a visible part of a

person's body, young children usually use these bilabial sounds in some of their first spoken words ("Momma" or "Poppa"). Think of the baby word terms for mother and father in other languages you may know; they probably contain bilabial consonants.

✓ **Labiodental:** Your top teeth touch your bottom lip to form these sounds. Say "fat" and "vat" to sample a voiceless and voiced pair produced at the labiodental place. A person could logically flip things around and try to make a consonant by touching the bottom teeth to the top lip. I can't take any legal responsibility for any spluttering behavior from such an ill-advised anatomical experiment.

✓ **Dental:** A closure produced at the teeth with contact of the tongue tip and/or blade makes these consonants. For American English, this refers to the "th" sounds, as in "thick" and "this." The first sound is voiceless and is transcribed with the IPA symbol /θ/, *theta*. The second is voiced and is transcribed with the IPA symbol /ð/, *ethe*. Beginning phonetics students frequently mix up /θ/ and /ð/, probably due to the dreadful problem of fixating about spelling. Remember to use your ear and the IPA, and you'll be fine.

Phonetics is a discipline where (for once) you really don't have to worry about how to spell. In fact, an overreliance on spelling can trip you up in many ways. When you hear a word and wish to transcribe it, concentrate on the sounds and don't worry about how it's spelled. Instead, go directly to the IPA characters. If you remain hung up on spelling, a good way to break this habit is to transcribe *nonsense words* also known as *nonce words* because you can't possibly know how they're spelled correctly.

✓ **Alveolar:** As I discuss in the earlier section, "Shaping the airflow," this important bony ridge on your hard palate makes the sounds /t/, /d/, /s/, /z/, /n/, /l/, and /ɹ/. The tongue tip makes some of these sounds, while the tongue blade makes others.

✓ **Retroflex:** This name literally means *flexed backwards*. Placing the tongue tip to the rear of the alveolar ridge makes these sounds. Although (as I show you in Chapter 16) such sounds are common in the English accents of India and Pakistan, they're less common in American or British English.

✓ **Palato-alveolar:** This region is also known as the *post-alveolar*. You make these sounds when you place the tongue blade just behind the alveolar ridge. Constriction is made at the palatal region, as in the sound "sh" of "ship," transcribed with the IPA character /ʃ/, known as "esh." The voiced equivalent, "zh," as in "pleasure" or "leisure," is transcribed in the IPA as /ʒ/, "long z" or "yogh." English has many /ʃ/ sounds, but far fewer /ʒ/ sounds (especially because many /ʒ/-containing words are of French or Hungarian origin, thank you, Zsa Zsa Gabor).

✓ **Palatal:** You make this sound by placing the front of the tongue on the hard palate. It's the loneliest place of articulation in English. Although some languages have many consonants produced here, English has only the gliding sound "y" of "yes," transcribed incidentally, with /j/. Repeat "you young yappy yodelers" if you really want a palatal workout.



✓ **Velar:** For these sounds, you're placing the back of your tongue on the soft palate. That's the pliant, yucky part of the back of your mouth with no underlying bone to make it hard, just cartilage. Try saying "kick" and "gag" to get a mouthful of stop consonants made here. You can also make nasal consonants here, such as the sound at the end of the words "sing, sang, sung" — transcribed with the IPA symbol /ŋ/, "eng" or "long n."

Note that /ŋ/ isn't the same nasal consonant as the alveolar /n/, such as in "sin." Velar nasals have a much more "back of the mouth" sound than alveolars. Also, people speaking English can't start a word with velar nasals — they occur only at the end of syllables. So, if someone says to you "have a gnice /ŋaɪs/ day!," you should suspect something has gone terribly, terribly wrong.

Beginning transcribers may sometimes be confused by "ing" words, such as "thing" (/ɪŋ/ in IPA) or "sang" (/sæŋ/ in IPA). A typical question is "where is the 'g'?" This is a spelling illusion. Although some speakers may possibly be able to produce a "hard g" (made with a full occlusion) for these examples (for example, "sing"), most talkers don't realize a final stop. They simply end with a velar nasal. Try it and see what you do. On the other hand, if you listen carefully to words, such as "singular," "linguistics," or "wrangle," there indeed should probably be a /g/ placed in the IPA transcription because this sound is produced. I provide more help on problem areas for beginning transcribers in Chapter 20.



Nosing around when you need to

Although it may sound disturbing, people actually talk through their noses at times. The oral airway is connected to the nasal passages — you may have unfortunately discovered this connection if you've unluckily burst out laughing at a funny joke while trying to swallow a sip of soda.

Air usually passes from the lungs through the mouth during speech because during most speech the soft palate raises to close off the passage of air through the nose. However, in the case of nasal consonants, the velum lowers roughly at the same time as the consonantal obstruction in the mouth, resulting in air also flowing out through the nose. People do this miraculous process of shunting air from the oral cavity to the nasal cavity (and back again) automatically, thousands of times each day.



Here is a nifty way to detect nasal airflow during speech. Ladies, get your makeup mirrors! Guys, borrow one from a friend. If the mirror is cool to the touch, you're good to go. If not, place it in a refrigerator for an hour or so, and you'll be ready to try a classic phonetician's trick. Hold the mirror directly under your nose and say "dice" three times. Because the beginning of "dice" has an oral consonant, you should observe, well, *nothing* on the mirror. That is, most air escapes through your mouth for this sound. Next, try saying "nice" three times. This time, you should notice some clear fog marks under each

nostril where your outgoing air during nasal release for /n/ made contact with the mirror. You may now try this with other places of (nasal) articulation, such as the words “mime” and “hang.”

Minding your manners

Blocking the vocal tract forms consonants. Forming consonants can happen in different ways: by making a complete closure for a short or long time, by letting air escape in different fashions, or by having the articulators approach each other for a while, resulting in vocal tract shapes that modify airflow. The following list includes some of the main manners of articulation in English. I discuss more details on manner of articulator, including examples for other languages, in Chapters 5 and 16.

- ✓ **Stop:** When air is completely blocked during speech, this is called a *stop* consonant. English stops include voiceless consonants /p/, /t/, and /k/ and voiced consonants /b/, /d/, and /g/, as in the words “pat,” “tat,” and “cat” and “ball,” “doll,” and “gall.” You make these consonants by blocking airflow in different regions of the mouth. *Nasal stops* (sometimes called just *nasals*, for short) also involve blocking air in the oral cavity, but they’re coordinated with a lowering of the velum to allow air to escape through the nose.
- ✓ **Fricative:** These consonants all involve producing *friction*, or hissing sound, by bringing two articulators very close to each other and blowing air through. When air passes through a narrow groove or slit, a hiss results (think of opening your car window just a crack while driving down the freeway at a high speed). You hiss with your articulators when you make sounds, such as /f/, /v/, /s/, or /ð/ (as in “fat,” “vat,” “sat,” and “that”). Chapter 6 provides more information on English fricatives.
- ✓ **Affricate:** This type of consonant may be thought of as a combination of stop and fricative. That is, an *affricate* starts off sharply with a complete blockage of sound and then transitions into a hiss. As such, the symbols for affricates tend to involve double letters, such as the two affricates found in English, the voiceless /tʃ/ for “chip” or “which,” and the voiced affricate /dʒ/, as in “wedge” or “Jeff.” Note that some authors tie the affricate symbols together with a tie or bar, such as /tʃ̄/ or /dʒ̄/. I use more recent conventions and don’t do so.
- ✓ **Approximant:** In these consonants, two articulators approach or *approximate* each other. As a result, the vocal tract briefly assumes an interesting shape that forms sound without creating any hissing or complete blockage. These sounds tend to have a fluid or “wa-wa”-like quality, and include the English consonants /ɹ/, /l/, /j/, and /w/, as in the words “rake,” “lake,” “yell,” and “well.”



A good way to remember the English approximants is to think of the phrase “your whirlies,” because it contains them all: /j/, /ɹ/, /w/, and /l/.

Note that the American English “r” is properly transcribed upside down, /ɹ/, in IPA. Many varieties of “r” sounds exist in the world, and the IPA has reserved the “right side up” symbol, /r/, for the rolling (trilled) “r,” for instance in Spanish. I go over more information on IPA characters in Chapter 3.

- ✓ **Tap:** For this consonant, sometimes called a *flap*, the tongue makes a single hit against the alveolar ridge. It’s a brief voiced event, common in the middle of words such as “city” in American English. A tap is transcribed as /ɾ/ in the IPA.

Producing Vowels

Vowels are produced with relatively little obstruction of air in the vocal tract, which is different than consonants. Phoneticians describe the way in which people produce vowels in different terms than for consonants. Because vowels are made by the tongue being held in rather complicated shapes in various positions, phoneticians settle for rather general expressions such as “high, mid, low” and “front, center, back” to describe vowel place of articulation. Thus, a sound made with the tongue held with the main point of constriction toward the top front of the mouth is called a *high-front vowel*, while a vowel produced with the tongue pretty much in the center of the mouth is called a *mid-central vowel*. The positions of the lips (rounded or not) are also important.

As I describe in Chapters 12 and 13, many phoneticians believe a better description of vowels can be given acoustically, such as what a sound spectrograph measures. Nevertheless, the best way to understand how vowels are formed is to produce them, from the front to the back, and from top to the bottom.

To the front

The *front* vowels are produced with the tongue tip just a bit behind your teeth. Start with the sound “ee” as in “heed,” transcribed in the IPA as /i/. Say this sound three times. This is a high-front vowel because you make it at the very front of your mouth with the tongue pulled as high up as possible. Next, try the words “hid,” “hayed,” “head,” and “had” — in this order. You’ve just made the front vowel series of American English. In IPA symbols, you transcribe these vowels as /ɪ/, /e/, /ɛ/, and /æ/.

As you speak this series, notice your tongue stays at the front of your mouth, but your tongue and jaw drop because the vowels become progressively lower. By the time you get to “had,” you’re making a low-front vowel.

To the back

You form the back vowels at, where else, the back of your mouth (big surprise!). Start with “boot” to make /u/, a high-back vowel. Next, please say “book” and “boat.” You should feel your tongue lowering in the mouth, with the major constriction still being located at the back. Phoneticians transcribe these vowels of American English as /ʊ/ and /o/.

The next two (low-back) sounds are some of the most difficult to tell apart, so don’t panic if you can’t immediately decipher them. Say “law” and “father.” In most dialects of American English, these words contain the vowels “open-o” (/ɔ/) and /ɑ/, respectively. Most students (and even many phoneticians) have difficulty differentiating between them. These vowels also are merging in many English dialects, making consistent examples difficult to list. For example, some American talkers contrast /ɔ/ and /ɑ/ for “caught-cot”, although most don’t. Nonetheless, with practice you can get better at sorting out these notorious two vowel sounds at the low-back region of the vowel space!

In the middle: Mid-central vowels

A time-honored method of many phonetics teachers is to save teaching the English central vowels for last because the basics of mid-central vowels are easy, but processing all the details can get a bit involved. For now, let me break them into these two classifications.

“Uh” vowels

The “uh” vowels include the symbols /ə/ “schwa” and /ʌ/ “wedge”, as in the words “the” and “mud.” Don’t be surprised if these two vowels (/ə/ and /ʌ/) sound pretty much the same to you (they do to me) — the difference here has to do with linguistic stress — because words with linguistic content such as nouns, verbs, and adjectives (for example, “mud” and “cut”) are produced with greater *linguistic stress* (see Chapter 7 for more details). They’re produced with a slightly more open quality and are assigned the symbol /ʌ/. Refer to the later section, “Putting Sounds Together (Suprasegmentals)” for more about linguistic stress. In contrast, English *articles*, such as “the” and “a” (as well as weak syllables in polysyllabic words, such as the “re” in “reply”) tend to be produced quietly, that is with less stress. This results in a relatively more closed mouth position for the “uh” sounds, transcribed as the vowel /ə/.



I dislike the names “schwa” and “wedge” because these character names don’t represent their intended sounds well. Therefore, I suggest you secretly do what my students do and rename them something like “schwuh” and “wudge.” Doing so can help you remember that these symbols represent an “uh” quality.

“Er” vowels

English has /ə/ (“r-colored schwa”) and /ɜ:/ (right-hook reversed epsilon) for “er” mid-central vowels. Notice that both of these characters have a small part on the right (a right hook, not to be confused with the prizefighting gesture) that indicates *rhoticization*, also referred to as *r-coloring*. For most North American accents, you can find the vowels /ə/ and /ɜ:/ in the words “her” and “shirt.”

The good news is that similar stress principles apply with the “er” series as the “uh” series. Pronouns such as “her” or endings such as the “er” in “father” typically don’t attract stress and thus are written with an r-colored schwa, /ə/. On the other hand, you transcribe a verb, such as “hurt” or an adjective such as “first,” with the vowel /ɜ:/ (right-hook reversed epsilon).

Embarrassing ‘phthongs’?

The vowels in the preceding section are called *monophthongs*, literally “single sound” (in Greek). These vowels have only one sound quality. Try saying “the fat cat on the flat mat.” The main words here contain a monophthongal vowel called “ash,” written in the IPA as /æ/. Notice how /æ/ vowels have one basic quality — they are, if you will, flat.



Next, try saying the famous phrase “How now brown cow?” Pronounce the phrases slowly and notice that each vowel will seem to slide from an “ah” to an “oo” (or in the IPA, from an /a/ to an /u/). For this reason, these words are each said to each contain a *diphthong*, or a vowel containing two qualities. For /au/, English speakers transition from a low-front to a high-back vowel quality. In addition to /au/, English has two other diphthongs, /aɪ/ (as in “white” or “size”) and /ɔɪ/ (as in “boy” or “loiter”).

Are diphthongs really embarrassing? They shouldn’t be, unless you produce them in an exaggerated manner (such as in the previous exercise). However, if you feel shy about producing diphthongs, you may wish to think twice about studying a language, such as the Bern dialect of Swiss German, which has diphthongs and even triphthongs aplenty. Yes, you guessed correctly — in a *triphthong*, one would swing through three different vowel qualities within one vowel-like sound. Check it out with the locals the next time you are in Bern (and don’t really worry about being embarrassed).

Putting sounds together (suprasegmentals)

Consonants and vowels are called *segmental* units of speech. When people refer to the consonants and vowels of a language, they’re dealing with individual (and logically separable) divisions of speech. This part is an important aspect of phonetics, but surely not the only part. To start with, consonants and vowels combine into *syllables*, an absolutely essential part of language. Without syllables,

you couldn't even speak your own name (and would, I suppose, be left only with your initials). Therefore, you need to consider larger chunks of language, called *suprasegmentals*, or sections larger than the segment.

Suprasegmentals refer to those features that apply to syllables and larger chunks of language, such as the phrase or sentence. They include changes in *stress* (the relative degree of prominence that a syllable has) and *pitch* (how high or low the sound is), which the following sections explain in greater detail.

Emphasizing a syllable: Linguistic stress

When phoneticians refer to stress, they don't mean emotional stress. For English, *linguistic stress* deals with making a syllable louder, longer, and higher in pitch (that is, making it stand out) compared to others. Stress can serve two different functions in language:

- ✓ *Lexical* (or word level)
- ✓ *Focus* (or contrastive emphasis)



Part of knowing English is realizing when stress is placed on the correct *syllable* (here at the beginning of the word), and not on a *wrong syllable* (such as here, in the middle of the word). Words that are *polysyllabic* (containing more than one syllable) have a correct spot for main stress (also called *primary stress*). Therefore, getting the stress right is an important part of our word learning.

In addition, some English word pairs show regular contrast between nouns and verbs with respect to stress placement. Say these words to yourself:

Noun	Verb
record	(to) record
(his) conduct	(to) conduct
(the) permit	(to) permit

You can tell that stress falls on the first syllable of the nouns, and the last syllable of the verbs, right? For some English word pairs stress assignment serves a grammatical role, helping indicate which words are nouns and which are verbs.

Stress can also be used to draw attention (*focus*) to a certain aspect of an utterance, while downplaying others. Repeat these three sentences, stressing the bolded word in each case:

- Sonya** plays piano.
- Sonya **plays** piano.
- Sonya plays **piano**.

Does your stressing these italicized words differently change the meaning of any of these sentences? Each sentence contains the same words — thus, logically, they should all mean the same thing, right? As you probably guessed, they don't. When people stress a certain word in a phrase or sentence, they do shift the emphasis or meaning. These three sentences all seem to answer three different questions:

Who plays piano?	(<i>Sonya</i> does!)
Does Sonya listen to piano or play piano?	(She <i>plays</i> !)
Does Sonya play the bagpipes?	(No, she plays <i>piano</i> .)

Using stress allows people to convey very different emphasis even when using the same words. Correctly using stress in this way is quite a challenge for computers, by the way. Think of how computer speech often sounds or how the stress in your speech may be misunderstood by computerized telephone answering systems.



A good way to practice finding the primary stress of a word is to say it while rapping out the rhythm with your knuckles on a table. For instance, try this with “refrigerator.” You should get something like:

knock *knock* knock knock knock

That is, the stress falls on the second syllable (“fridge”).

Next, try the word “tendency.” You should have:

knock knock knock

Here, stress falls on the first (or initial) syllable.

This method seems to work well for most beginning phonetics students. I think the only time students have difficulty with stress assignment is if they overthink it. Remember, it is a sound thing and really quite simple after you get the hang of it.

Changing how low or high the sound is

Pitch is a suprasegmental feature that results from changes in the rate of buzzing of the larynx. The faster the buzzing, the higher pitched the sound; the slower the buzzing, the lower the sound.

Men and women buzz the larynx at generally different rates. If you're an adult male, on average your larynx buzzes about 120 times per second when you speak. Women and children (having higher voices) buzz at typically about twice that rate, around 220 times per second. This difference is due to the fact that men have larger laryngeal cartilages (Adam's apple) and vocal folds.

Phoneticians call the rate of this buzzing *frequency*, the number of times something completes a cycle over time. In this case, it's the number of times that air pulses from the larynx (resulting from the opening and closing of the vocal folds) per second.



Pitch refers to the way in which frequency is heard. When phoneticians talk about pitch, they aren't referring to the physical means of producing a speech sound, but the way in which a listener is able to place that sound as being higher or lower than another. For example, when people listen to music, they can usually tell when one note is higher or lower than another, although they may not know much else about the music (such as what those notes are or what instruments produced them). Detecting this auditory property of high and low is very important in speech and language.

English uses pitch patterns known as *sentence-level intonation*, which means the way in which pitch changes over a phrase- or sentence-length utterance to affect meaning. Try these two sentences, and listen carefully to the melody as you say each one:

- ✓ “I am at the supermarket.” This type of simple factual statement is usually produced with a *falling intonation contour*. This means the pitch drops over the course of the sentence, with the word “I” being higher than the word “supermarket.” Many phoneticians think this basic type of pitch pattern may be universal (found across the world's languages). People blow off air when they exhale for speech, providing less energy for increased pitch by the end of an utterance, compared to the beginning.
- ✓ “Are you eating that egg roll?” In this question, you probably noticed your melody going in the opposite direction, that is — from low to high. In English, people usually form this kind of “yes/no question” (a question that can be answered with a yes or no answer) with a rising intonation pattern. Indeed, if you were to restate the factual sentence “I am eating an egg roll” and change your intonation so that the pitch went from low to high, it would turn into a question or expression of astonishment.

These examples show how a simple switch in intonation contour can change the meaning of words from a statement to a question. In Chapter 10, I discuss more about the power of intonation in English speech.