Phonetics and Linguistics

Fourth year

Faculty of Arts

English language and translation program

2025

In this part . . .

- Understand how consonant and vowel sounds are produced in order to classify the different sounds we use in language. Understanding sound production also helps with pronunciation.
- Differentiate between broad and narrow transcriptions, identify the purpose for each type, and begin to make your own transcriptions.
- Take a closer look at how *phonology* (sound systems and rules in languages) and *phonetics* (the study of the actual speech sounds) are related and see how together they provide a richer description of spoken language.
- Acquaint yourself with some basic phonological rules for the English language so you can make more informed transcriptions.
- Grasp the concepts of juncture, stress, rhythm, intonation, and emotion and what you need to know about them when transcribing.
- Know how to identify *prosody* (language melody) details in speech and applying what you've identified into your transcriptions.

Chapter 6

Sounding Out English Consonants

In This Chapter

- Showcasing stops
- Focusing on fricatives and affricates
- Analyzing the production of approximants
- Describing coarticulation

Producing speech is a tricky business and the exact way in which consonants are made can result in vast differences in how these sounds are heard. In this chapter, I walk you through some different types of consonant manners (stops, fricatives, affricates, and approximants), zeroing in on those mouth and throat details that make big perceptual differences in the English language.

Stopping Your Airflow

Stop consonants (sounds made by completely blocking oral airflow) are part of a larger group called *obstruents*, which are sounds formed by shaping airflow via obstruction (this group also includes fricatives and affricates). *Fricatives* are made when air is blown through a space tight enough to cause friction (or hissiness). *Affricates* are sounds that begin as a stop, then release into a fricative. Refer to Chapters 4 and 5 for more information on these types of sounds. When airflow is completely stopped, several different things can happen:

- Air can be released into the vocal tract in different ways.
- Air can flow into different regions when the sound is released.
- The duration of the closure itself can last for longer or shorter periods.

Some of these puzzling mechanics are revealed in the following sections.

Huffing and puffing: Aspiration when you need it

Aspiration is the airy event that takes place just after the burst of the articulators blasting open and before the voicing of the vowel. Aspirated voiceless stop consonants are made with an audible puff of breath. Aspiration, represented by the raised letter "h" ($[^h]$) occurs for a brief period of time starting just after the beginning of a stop consonant. To see how this works, consider what happens when you produce the word "pie."

1. **The lips close together to make the [p^h].**

   This is referred to as *closure.*
2. **Air pressure increases to start the [p^h] gesture.**

   This step refers to *oral pressure buildup.*
3. **The lips are rapidly blown apart, resulting in a typically "p"-like sound.**

   This step is also referred to as a *burst.*
4. **Because the vocal folds are opening and the pressure conditions are right, a puff of air follows just after the burst.**

   This audible puff of air is called the *aspiration.*
5. **The vocal folds start to buzz for the [aɪ] diphthong.**

TRY IT If you want to feel the aspiration, place your hand just under your bottom lip while you're talking. Do you feel the air pass over your hand? That air is aspiration.

Try this again and say "*p*ot." You should be able to feel the aspiration of the [p^h] as a puff of air hits your hand when you begin the word.

Now, try the same exercise while saying "*t*ot" and "*c*ot." At the beginning of these words, you also produce aspirated stops ([t^h] and [k^h]), but you may not feel much of a puff because the release is taking place farther back in your mouth. Even though you may not always feel aspiration, it's important you be able to hear and transcribe it.

Being able to work with aspiration comes with practice. In English, the voiceless stops [p], [t], and [k] are aspirated at the start of a word and at the beginning of stressed syllables. You transcribe the aspiration by adding the diacritic ([^{h}]), resulting in [p^h], [t^h], and [k^h]. In other contexts, [p], [t], and [k] aren't aspirated.

Table 6-1 shows you a quick overview of the rules of aspiration in English.

Table 6-1 The Rules of Aspiration in English

Context	*Examples*	*Aspiration*	*IPA*
Syllable initial	*p*ot	Strong for most speakers	[p^h]
	*t*ot		[t^h]
	*c*ot		[k^h]
Following an /s/	s*p*ot	None	[p]
	s*t*ock		[t]
	S*c*ott		[k]
Syllable final	ho*p*	None	[p]
	ho*t*		[t]
	ho*ck*		[k]

I use square brackets ([ ]) instead of slash marks (/ /) to mark these sounds in Table 6-1 because doing so shows narrow *phonetic* detail. The aspiration diacritic [^{h}] is included in *narrow* transcriptions of English, not broad. Aspirated stops in English occur as the result of rule–governed processes (also called *allophonic processes*).

Try saying the words in the second column and make sure you can hear the aspiration in the underlined consonants in the first row (but not in the stop consonants in the second and third row).

Declaring victory with voicing

The English voiced stop phonemes (/b/, /d/, and /g/) aren't produced with aspiration, so it may seem simple that they can be distinguished from their voiceless counterparts (/p/, /t/, and /k/). However, if you listen carefully, you should be able to tell that voicing also behaves rather differently in different environments. Take a look at Table 6-2 where you see how the amount of voicing for /b/, /d/, and /g/ changes in different environments.

Table 6-2 The Amount of Voicing in English Stops

Context	*Examples*	*Voicing*	*IPA — Narrow Transcription*
Middle of voiced phrase (VCV)	a *boon*	Strong	[b]
	a *dune*		[d]
	a *goon*		[g]
Sentence initial	*boon*	Weak	[b]
	dune		[d]
	goon		[g]
Following voiceless sound	that *boon*	Weak	[b]
	that *dune*		[d]
	that *goon*		[g]
Syllable final	ta*b*	Weak	[b]
	ta*d*		[d]
	ta*g*		[g]

When a voiced stop occurs between flanking voiced sounds (as shown in the first row of Table 6-2), voicing is usually strongly produced throughout the stop closure. However, in all the other cases, English [b], [d], and [g] actually aren't that strongly voiced.

The reason these weakling voiced stops (in rows 2, 3, and 4 of Table 6-2) are still heard as voiced (that is, as [b], [d], and [g]) is because other information signals listeners that a voiced sound is intended. One of these cues, *voice onset time* (VOT) is discussed in more detail in Chapter 14.

Another interesting way voicing is conveyed in English is by vowel length. To get an idea of how this works, concentrate on how long each word is when you say the word pairs in the following list:

tap	tab
tat	tad
tack	tag

What do you notice? You may hear that the vowel /æ/ is longer before the voiced stops /b/, /d/, and /g/ than the voiceless stops /p/, /t/, and /k/. People hear this change in *vowel* length as the voicing of the final *consonant*. Although physical voicing may be stronger or weaker depending on the context (as shown in Table 6-2), the *feature* of voicing is abstract and perceptual. That is, the feature of voicing is in the ear of the beholder and can be signaled by various types of information.

It's possible to computer-edit versions of these words. Computer editing involves entirely removing the final consonant release and leaving only the vowel length. People still hear the (missing) final consonant voicing difference quite reliably.

Glottal stopping on a dime

If you already read Chapter 2, you discovered information about your glottis. A *glottal stop* is made whenever the vocal folds are pressed together. This process happens easily and naturally, such as whenever you cough. To make a glottal stop on command, just say "u*h*-oh" and hold the "u*h*."

Glottal stops appear in English more than people think. In London, Cockney accents are a key feature, appearing in words such as *bottle* [ˈbɒʔɫ̩] and, yes, *glottal* [ˈglɒʔɫ̩]. In North American English, glottal stops are often produced before a stop or affricate at the end of a syllable, for instance "ra*p*" or "chur*ch*."

To get a sense, try saying "ta*p*" and "ta*b*." You'll notice that the vowel in these words is longer for the second word (ending with the voiced stop, /b/), than the first, containing, /p/. Also, you may notice that you close down your glottis before you get to the final /p/ of "ta*p*," and release no air afterwards. You probably produced [tʰæʔp].

Of course, you can pronounce "tap" in different ways. Try the varieties in Table 6-3.

Table 6-3 Different Ways to Pronounce "Ta*p*"

Pronunciation	***IPA***
With no glottal stop and no final release	[tʰæp]
With no glottal stop and final release	[tʰæpʰ]
With glottal stop and final release	[tʰæʔpʰ]
With glottal stop only	[tʰæʔ]

I give you the audio examples that are linked to each way of making the final consonant at www.utdallas.edu/~wkatz/PFD/tap_examples.html.

Doing the funky plosion: Nasal

In *oral plosion* (or *explosion,* when a sound is made by the articulators forced open under pressure), the articulators separate and a burst of air is released from the oral cavity. This happens for most English stops. However, when a voiced stop and a nasal occur together, as in the word "su*dden,*" something quite different happens: The air pressure built up by the stop is instead released through the nose. This process is called *nasal plosion*, which you accomplish by lowering your soft palate, also called the *velum*. Nasal plosion has the effect of producing less of a vowel-like quality for the release and more of a nasal quality. Refer to Chapter 6 for more information on oral and nasal stop consonants.

Say the word "sudden." How much of an "un" sound do you hear at the end? It shouldn't be much. Next, imagine there was an ancient poet named "*Sud Un*." (Yes, it's a bit far-fetched, but at least it provides a different stress structure!) Say the two, side by side:

sudden Sud Un

You should be able to hear nasal plosion in *sudden*, but not in the "Un" of "Sud Un." The latter should have much more vowel quality because it's pronounced with more stress and no nasal release of the previous stop.

Notice that nasal plosion only occurs for stops that are *homorganic*, sharing the same place of articulation. This table shows the possible *homorganic* combinations of oral and stop consonants for English.

Oral Stops	*Nasal Stop*
/p/, /b/	/m/
/t/, /d/	/n/
/k/, /g/	/ŋ/

To put it another way, /pm/, /bm/, /tn/, /dn/, /kŋ/, and /gŋ/ are the homorganic stop/nasal combinations in English. When you say words having these combinations in English, chances are you'll use nasal plosion (as in "sudden" and "hidden"). However, when stop/nasal combinations aren't homorganic (such as /bn/ or /gn/), nasal plosion doesn't occur. You'll notice this if you say "ribbon" and "dragon," where there is no nasal release because these combinations of stop and nasal aren't homorganic.

Doing the funky plosion: Lateral

Lateral plosion involves a stop being released by lowering the sides of the tongue, instead of making an oral release by the articulator. When lateral plosion occurs, no vowel sound takes place in the syllable involved. Instead, there is more of a pure "l" sound. To get an idea, try saying these utterances, side by side, while listening to the final syllable:

Lateral Plosion	*Without Lateral Plosion*
ladle	lay dull
noodle	new dull

Depending on your accent, you may have slightly different realizations of the vowels in these expressions. There should be more of an "l" ending for the endings of the left column, and a vowel-containing ending (/əl/) for those on the right.

Tongue tapping, tongue flapping

The *tap* [ɾ] is a rapid, voiced alveolar stop used by many speakers to substitute for a /t/ or /d/. It's typically an American (and Canadian) gesture in words such as "Be*tt*y," "ci*t*y," "bu*tt*er," and "be*tt*er." (Refer to Chapter 18 where I discuss American and Canadian dialects.) I call it a *tap,* although some phoneticians refer to it as a *flap*. The difference between a tap and a flap is whether an articulator comes up and hits the articulator surface from one direction and returns (tap), or hits and continues on in the same direction in a continuous flapping motion (flap). I say we call it a tap and be done with it.

Notice that tap is shown in square brackets ([ ]) because it's an allophone in English and can't stand on its own freely to change meaning. That is, you can't say something like "Tomorrow is Fat Tuesday" [ɾəˈmɑɹo ɪz fæɾ ˌɾuzɾeɪ] where tap freely stands in for *any* /t/ or /d/.

Taps are quite important for North American English. Most Americans and Canadians replace medial /t/ and /d/ phonemes with a tap in words such as "latter" and "ladder." Forget about spelling — for spoken American English, these words often sound just the same.

Say these phrases and see how you sound:

It’s the la*tt*er.

It’s the la*dd*er.

If you’re a native speaker of English from somewhere in North America, you may likely tap the *medial alveolar consonant* (/t/ or /d/ in the middle of a word). If you speak British English or other varieties, this isn’t likely.

Having a Hissy Fit

Fricatives are formed by bringing the articulators close enough together that a small slit or passageway is formed and friction, or hissiness, results. The fricatives are copycats of many of the allophonic processes of the stops. For example, just as vowel length acts as a cue to the voicing of the following stop (as in “bi*t*” versus “bi*d*”), a similar process takes place with voiceless and voiced fricatives.

Try pronouncing the pairs in Table 6-4 and convince yourself this is the case. Notice that the /i/ in “gr*ie*ve” is longer than the /i/ in “*gr*ief” (and so forth, for the remaining pairs). This table shows examples of English word pairs having voiceless and voiced fricatives in syllable-final positions. In each case, relatively longer vowels in front cue the voiced members.

Table 6-4 Voicing Contrasts for Fricatives in Syllable-Final Position

Voiceless	*Voiced*	*IPA*	
grief	grieve	/ɡɹif/	/ɡɹiv/
teeth	teethe	/tiθ/	/tið/
race	raze	/ɹes/	/ɹez/
nation	Asian	/ˈneʃən/	/ˈeʒən/

After you've spoken the words in Table 6-4, listen carefully to the fricatives and focus on how long each fricative portion lasts. Here, you should hear another length distinction but one that's going in the opposite direction. Final voiceless fricatives are longer than final voiced fricatives. That is, the /v/ in "grie*v*e" is shorter than the /f/ in "grie*f*." See if you can hear these differences for the rest of the pairs.

This consonant duration difference is also found for stops in final position (such as "bi*t*" versus "bi*d*"). However, because stop consonants are so short, it's difficult to get a sense of this without measuring them acoustically (see Chapter 12 for more information on acoustic phonetics).

Another important point to note about the English fricatives is that four of them are *labialized*, produced with secondary action of the lips. These "lippy" upstarts include /ʃ/ and /ʒ/ (highly labialized), and /s/ and /z/ (partially labialized). For these sounds, the position of the lips helps make the closure of the fricative.

Feel the positions of your lips while pronouncing these words:

- /ʃ/: "pre*ss*ure"
- /ʒ/: "trea*s*ure"
- /s/: "*s*ip"
- /z/: "*z*ip"

For these fricatives, you purse your lips to help make the sound. In contrast, for the fricatives /θ/ and /ð/ (as in "*th*ick" and "*th*is"), the placement of your lips isn't particularly important. Your tongue placed in between your teeth causes the hissiness.

The phoneme /h/ is a lost soul that needs to be given a special place of its own. Although technically classified in the IPA as a voiceless glottal fricative, its occurrence in English can be rather puzzling. (Refer to Chapter 3 for more about the IPA.) Often, it's produced without any glottal friction at all, such as "a*h*ead" or "a*h*oy there." In such cases, a weakening of the flanking vowels signal the /h/. In contrast, there may be strong frication in words such as "h*ue*" (/hju/). To add to the mix, people who produce the voiceless phoneme /ʍ/ in their dialect (as in "whip" pronounced as "hwip") are pronouncing "h" as part of an approximant (again, with no friction). So /h/ is wild and crazy, but I say you give it a home in the fricative category (as long as you remember it may not always stay put).

Going in Half and Half

Affricates are a combination of a stop followed by a fricative. English has two affricate phonemes: /ʧ/ and /ʤ/. In the IPA chart, /ʧ/ and /ʤ/ are listed as *post-alveolar* (produced by placing the tongue front just behind the alveolar ridge) because this place of articulation corresponds to the major part of the sound — namely, the fricative.

In some situations in English, a stop butts up against a *homorganic* (sharing the same place of articulation) fricative, creating situations that may seem "affricate-like.” However, these instances aren’t true affricates. For example, the sound /t/ can sometimes adjoin the sound /s/, as in the phrase “It seems.”

However, to demonstrate that this phrase isn’t a true affricate, you couldn’t get away with new English expressions, such as “*ts*ello,” “*ts*ow are you?” and so on, and expect anyone to think you’re speaking English. This is because /ts/ can’t stand alone as an English phoneme (although in other languages, such as Japanese, a /ts/ affricate phoneme is found, such as in the word “*tsunami*”).

Shaping Your Approximants

Approximants are formed by bringing the articulators together, close enough to shape sound, but not so close that friction is created. The English voiced approximant phonemes are /w/, /ɹ/, /j/, and /l/, as illustrated in the phrase “your whirlies” /jɚr 'wɪɹliz/. In addition to this set, “hw” (written in IPA with the symbol /ʍ/) is produced by some talkers as an alternative to voiced /w/ for some words. Some pronounce “whip” or “whether” with a /w/, and others with a /ʍ/. In most forms of English, the use of /ʍ/ seems to be on the decline.

Voiced approximants partially lose their voicing when they combine with other consonants to form *consonant clusters*, lawful consonant combinations. In Table 6-5, listen as you say the approximants in the middle column, followed by the same sounds contained in consonant clusters (right column):

Table 6-5 Fully and Partially Voiced Approximants in English

Approximant (IPA)	***Fully Voiced***	***Partially Voiced***
/w/	*wheat*	*tweet*
/ɹ/	*ray*	*tray*
/l/	*lay*	*play*
/j/	*you*	*cue*

Focus on the second row in the table. Place your fingers lightly over your Adam's apple and feel the buzzing while you say the "r" in the two words. You should feel less buzzing during the "r" in "tray" than in "ray." This is because the aspiration of the voiceless stop $[t^h]$ in "tray" prevents the approximant from remaining very voiced.

The "r" sound perhaps causes more grief to people learning English as a second language than any other. This is particularly true for speakers of Hindi, German, French, Portuguese, Japanese, Korean, and many other languages that don't include the English /ɹ/ phoneme.

Recent physiological studies show much variety in how native talkers produce this sound, although tongue shapes range between two basic patterns:

- **Bunched:** The anterior tongue body is lowered and drawn inwards, away from the front incisors, with an oral constriction made by humping the tongue body toward the palatal region. This variety is quite common in the United States and Canada.

- **Retroflex:** The tip is raised and curled toward the anterior portion of the palate.

Some clinicians use the hand cues seen in Figure 6-1 to help patients remember the bunched versus retroflex /ɹ/difference.

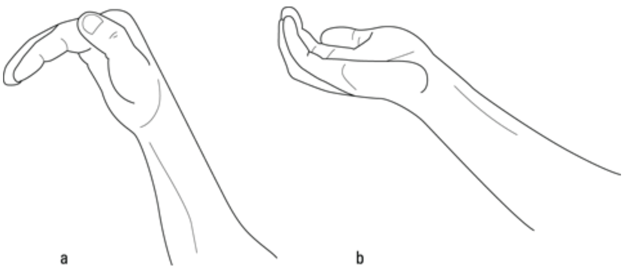

Illustration by Wiley, Composition Services Graphics

Figure 6-1: Bunched (a) versus retroflex (b) /ɹ/ hand signals.

Retroflex /ɹ/ varieties are more common in British English than in North American dialects. Also, most speakers of American and Canadian English make a secondary constriction in the pharyngeal region, as well as lip rounding behavior.

Here are some important points to know about English /ɹ/:

- /ɹ/ is a consonant.
- English also has two rhotic (r-colored) *vowels*, /ɝ/ (in stressed syllables) and /ɚ/, (in unstressed syllables); also as in “further” (/ˈfɝðɚ/).
- /ɹ/ is often called a *liquid* approximant (along with its cousin /l/) for rather odd reasons (dating back to how these sounds were used in Greek syllables).
- /ɹ/ is a relatively late-acquired sound during childhood, commonly achieved between the ages of 3 and 6 years. /

ɹ/, /ɝ/, and /ɚ/ are also error-prone sounds for children, with frequent /w/ substitutions (for example, "Mister Rabbit ([ˈmɪstə ˈwæbɪt]).

Exploring Coarticulation

Speech sounds aren't produced like beads on a string. When you say a word such as "suit," you aren't individually producing /s/, then /u/, and then /t/. Doing so would sound too choppy. Instead, you produce these sounds with *gestural overlap* (overlapping movements from different key parts of your articulatory system). (Chapter 4 provides further discussion.) *Coarticulation* refers to the overlapping of neighboring sound segments. In Figure 6-2, you see an image of what that overlap looks like for the word "suit."

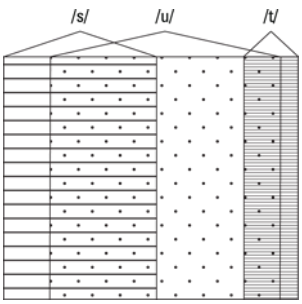

Illustration by Wiley, Composition Services Graphics

Figure 6-2: "Suit" showing sound overlap.

While the tongue, lips, and jaw are positioned to produce the frication (hissiness) for /s/, the lips have already become rounded (pursed) for the upcoming rounded vowel, /u/. This section explores some basics of coarticulation and introduces two main types of coarticulation.

Tackling some coarticulation basics

In order to better understand how coarticulation works, you need to master some important attributes. Keep in mind these general principles about coarticulation as you study more phonetics and phonology. These principles can help explain the distribution of allophones. Here are some cool things to know about coarticulation:

- All speech is coarticulated. Without it, humans would sound robot-like.
- The extent (and precision) of coarticulation differs between languages.
- Because many aspects of coarticulation are language-dependent, to some extent coarticulation must be acquired during childhood, and learned in an adult second language.

  However, birds (Japanese quail) have been trained to distinguish coarticulated speech sounds, suggesting that at some coarticulated processes can be accomplished on the basis of general auditory processing alone.
- Psycholinguistic research suggests children acquire coarticulation early in development.
- Coarticulation is thought to break down in certain speech and language disorders, such as apraxia of speech (AOS).

Anticipating: Anticipatory coarticulation

A “look ahead” activity is called *anticipatory* (or right-to-left) *coarticulation*. It is considered a measure of speech planning and as such is of great interest to psychologists (see below).

Try out anticipatory coarticulation for yourself! Say the following two phrases:

“I said suit again.”

“I said seat again.”

Pay special attention to when your lips begin to protrude for the /u/ in the word “s*u*it.” ***Note:*** There’s no such lip protrusion for the /i/ in “s*ea*t”; this is just for comparison. Most people will begin lip rounding for the /u/ by the beginning of the /s/ of “suit,” and some even earlier (for example, by the vowel /ɛ/ in the word “said”). That is, nobody waits until the /s/ is over until they begin to lip-round for the rounded vowel /u/.

These effects are important for optimizing speech speed and efficiency. The average person produces about 12 to 18 phonemes per second when speaking at a normal rate of speed. There would be no way to achieve such a rate if each phoneme’s properties had to switch on and off in an individual manner (such as when using a signaling system like Morse Code). However, when speech properties are overlapped, the system can operate faster and more efficiently.

Preserving: Perseveratory coarticulation

A second type of coarticulation called *perseveratory* (or *left-to-right*) coarticulation, is also known as *carry-over*.

Perseveration means that something continues or hangs on. In this case, it is the lingering of a previous sound on to the next. For instance, in "suit" it would be the hissiness of the /s/ carrying over to the beginning of the vowel /u/, or the rounding of the /u/ continuing on and influencing the final /t/. Perseverative coarticulation is a measure of the mechanical/elastic properties of the speech articulators, instead of planning.

> One way to remember perseverative coarticulation is to think about the role of this property in complex speech, such as *tongue twisters*. A property common to tongue twisters throughout the world is that they have phonemes with similar features in close proximity. The hope is that you'll have carryover effects from a sound you just made as you attempt to produce an upcoming sound with similar properties. Actually, if you begin thinking about it too much, you might then develop anticipatory problems as you desperately thrash around trying to keep the proper phonemes in mind. This is evident in the saying "She *s*old *s*eashells by the *s*ea*sh*ore."

You may end up saying "She *sh*old" as you carry over from the initial /ʃ/ of "she" to the target /s/ of "*s*old."

How far can you stretch it? Look-ahead coarticulation in English and French

English speakers only show anticipatory lip rounding over the course of a syllable or so. For example, in the word "suit," there's reliable anticipation of lip rounding for /u/ during the initial /s/. However, if someone begins to load up the front of the syllable such that more consonants intervene between the /s/ and the /u/, anticipatory lip-rounding gradually diminishes (such as more in "*s*uit" than in "*sp*ool").

What about other languages where lip rounding plays a more important (phonemic) role?

A now classic study asked six French speakers to produce a series of tongue-twister type expressions to see how far lip protrusion may extend. The subjects had a photocell attached to their upper lips as they said things like "une sinistre structure (a sinister structure)." The researchers measured when (and how far and how fast) the upper lip began to protrude for the rounded vowel /y/ in the upcoming word /strykˈtyr/.

The researchers found that notable lip protrusion for the /y/ of /strykˈtyr/ began as early as four to six consonants before.

In a follow-up experiment, listeners were given gated segments of the consonant clusters prepared by a waveform-editing program and asked to detect whether the segment was taken from an utterance that was before a rounded vowel. Listeners could do this at better than 50 percent accuracy by up to four consonants before the rounded vowel, suggesting that long-ranging labial coarticulation can be accurately tracked by listeners.

These experiments provided early evidence that when it comes to coarticulation, one size does not fit all: Languages which emphasize certain sound features (such as lip-rounding) in their sound inventories (such as French) have different coarticulatory features for these sounds than do languages (such as English) that don't.

Chapter 7

Sounding Out English Vowels

In This Chapter

- Searching for (IPA) meaning in all the right places
- Hearing vowels in full and reduced forms
- Switching between British and North American English vowels
- Keeping track of vowel quality over time

Vowels are a favorite subject of phoneticians because they play such an important role in perception, yet they pose so many mysteries about how speech is perceived and produced. Some vowels are quite easy to transcribe; some remain difficult. In this chapter, I highlight the commonalities among English vowels by describing the group's tense and lax characteristics. I also talk about rhoticization (also referred to as r-coloring), which is important for many applications, including the description of various English accents and understanding children's language development.

Cruising through the Vowel Quadrilateral

Making vowels is all about the tongue, lips, and jaw. However, the final product is *acoustic* (sound related), not *articulatory* (mouth related). Phonetics texts typically start out with articulatory instructions to get people started, but it becomes important to transfer this information to the ear — to the world of auditory information.

In articulatory phonetics, vowels are studied using the *vowel quadrilateral*, a trapezoid-like diagram that classifies vowels according to tongue height, advancement (front-back positioning), and lip rounding.

This section focuses on moving your tongue to known target regions and consciously getting used to what these regions sound like. In this way, sound anchors become familiar landmarks as you cruise through the land of vowels.

Looking up vowel sounds

Are you a book person, an Internet person, or both? The tools available for looking up sounds are becoming more extensive and convenient. As your transcription skills improve, you may find yourself wishing to compare pronunciations of certain words now and then. At such times, you may notice quite a bit of difference among various reference sources because vowels can be transcribed in a number of ways, depending on the exact needs of the transcription.

In the following, I compare some dictionary and Internet sources for the English words "cheap" and "chip" (American pronunciation) to see how they're transcribed. This table can give you an idea of how transcription is handled in a range of available sources.

Source	**"cheap"**	**"chip"**
English Pronouncing Dictionary (EPD)	/ʧi:p/	/ʧɪp/

Longman Pronunciation Dictionary (LPD)	/ʧiːp/	/ʧɪp/
Oxford Dictionary of Pronunciation for Current English (2003)	/ʧiːp/	/ʧɪp/
American Heritage Dictionary 5th Edition (2011)	(chēp)	(chĭp)

I selected these sources because they're some of the most authoritative print dictionary sources. The first three sources provide both British and American English pronunciations. Unlike the conventions used by most phonetics teachers (and used in this book), these first three sources are also more detailed in that they show both length and quality features for the "cheap/chip" distinction. That is, there are two related phonetic factors that contribute to the vowel difference:

- **Quality:** The formants are different in the two words, requiring different symbols /i/ versus /ɪ/
- **Length:** The vowel is longer in *cheap* than "chip," requiring a length diacritic for *cheap* /iː/

The Oxford Dictionary uses an even larger set of symbols and gives more details than the first two sources.

A different approach has been taken in the American Heritage, where an alternative to the International Phonetic Alphabet (IPA) was used, presumably to be more user friendly. This, however, isn't as handy or reliable for some people who've already learned the IPA.

This book follows a convention used by many phonetics instructors and transcribes vowel quality, assuming that length can be implied. Therefore, I transcript "cheap" as /ʧip/ and "chip" as /ʧɪp/. As long as you remember that quality and length go together, you should be able to appreciate other transcriptions (such as EPD and LPD) when you use them.

Sounding out front and back

Sound-based descriptions are especially important for vowels. For this reason, phoneticians have long relied on perceptual descriptions of vowels. For instance, front vowels were frequently called *acute* because they're perceptually sharp and high in intensity. These vowels also trigger certain sound changes in language (notably, palatalization) and involve active tongue blade (coronal) participation. In contrast, back vowels were called *grave* because they have dull, low intensity.

TRY IT You have made front vowels, but you have probably not spent that much time attending to what the vowels *sound* like. So here, you tune in to the sounds themselves. Begin by making an /i/ as in "heed," then move to the following, one by one:

- /ɪ/ as in "h*i*d"
- /e/ as in "h*ay*ed"
- /ɛ/ as in "h*ea*d"
- /æ/ as in "h*a*t"

You hold your tongue in a certain position for each vowel (although there is some wiggle room), and the tongue position need not be exact. Also, each vowel position can blend somewhat into the position of the next.

Now, try saying them all together in a sequence, /i ɪ e ɛ æ/. Notice that the vowels are actually in a continuum. Unlike consonants, vowels are made with the tongue relatively free in the articulatory space and the shaping of the whole vocal tract is what determines the acoustic quality of each sound.

Now try the same listening exercise with the back vowel series, beginning with /u/ as in *who'd*, and proceeding to the following:

- /ʊ/ of "h*oo*d"
- /o/ of "h*o*pe"
- /ɔ/ of "l*aw*"
- /ɑ/ of "d*o*g"

In the back vowel series, you pass through the often-confused /ɔ/ and /ɑ/. There are many dialectal differences in the use of these two vowels. For instance, in Southern California (and most Western United States dialects), most talkers pronounce "cot" and "caught" with /ɑ/. In Northern regions, say Toronto, talkers use /ɒ/ for both words. This vowel /ɒ/ is a low back vowel similar to /ɑ/ but produced with slight lip rounding. However, elsewhere in the States (especially in the Mid-Atlantic States) talkers typically produce "cot" with /ɑ/ and "caught" with /ɔ/. You can easily tell the two apart by looking at your lips in a mirror. During /ɑ/, your lips are more spread than in /ɔ/, and in /ɔ/ the lips are slightly puckered. Compare your productions with the drawings in Figure 7-1.

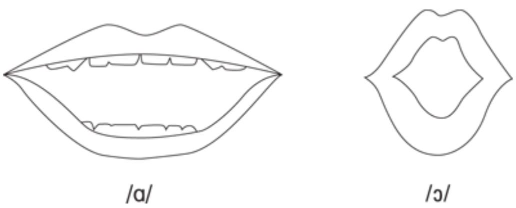

Illustration by Wiley, Composition Services Graphics

Figure 7-1: Lip positions for /ɑ/ versus /ɔ/.

The point here is to be able to *hear* such differences. Try moving from an /o/ as in "hope" to an /ɔ/ in "l*aw*" to an /ɑ/ in "f*a*ther." Now try this while adding lip rounding to the /ɑ/. You should hear its quality change, sounding more like /ɔ/. As you get more fine-grained in your transcriptions, you need to be able to distinguish vowels better, including whether lip rounding occurs as a secondary articulation.

Stressing out when needed

In English, stress refers to a sound being longer, louder, and higher. Stress is a *suprasegmental* property, meaning it affects speech units larger than an individual vowel or consonant. I also discuss stress in Chapters 10 and 11.

In English, the amount of stress a syllable receives influences vowel quality. Stressed syllables tend to have a full vowel realization, while unstressed syllables have a centralized, reduced quality. Sometimes there is a more complicated situation, where a full vowel will appear in fully stressed syllables, but whether a vowel is reduced in unstressed syllables depends on the particular word involved. Take a look at Table 7-1, and try saying the English words.

Table 7-1 Vowels in Different Stress Conditions

Vowel	*Fully Stressed*	*Unstressed, Not Reduced*	*Unstressed, Possibly Reduced?*
/i/	agr*ee*	r*e*active	*re*sourceful
/u/	ref*ute*	ref*u*tation	Herc*u*les
/aɪ/	re*cite*	c*i*tation	*recit*ation

You should have nice, full /i/, /u/, and /aɪ/ vowels in the words of the second column of Table 7-1. This should also be the case for the words in the third column.

The most variable response is the fourth column. The vowels produced here depend on your accent. These words contain unstressed syllables that some speakers produce with a fully realized vowel quality (for example, /ɹi/ for the first syllable of "*re*sourceful") while others use a reduced vowel instead (such as /ɹə/). If your vowel is a bit higher toward /ɪ/, it may qualify for being /ɨ/ (called barred-I in IPA), as is frequently heard in American productions of words such as "dish*es*" and "rich*es*."

By the way: If you find yourself almost forgetting what you normally sound like, please remember these rules:

- **Use a carrier phrase.** A *carrier phrase* is a series of words you place your test word into so that it's pronounced more naturally. For example, "I said ____ again."
- **Have one or two repetitions and then move on.** Natural speech is usually automatic and not consciously fixated on. If a word or phrase is repeated over and over, this natural, automatic quality may be lost.

Coloring with an "r"

Whether or not people produce an "r" quality in words like "*fu*rther," "fath*er*," and "s*ir*" is a huge clue to their English accent. Most speakers of North American English produce these vowels with *rhoticization*. This term, also referred to as *r-coloring,* means that the vowel (not the consonant) has an "r"-like sound. If the vowel is stressed, as in "*fu*rther" or "*sir*," then you use the mid-central stressed vowel /ɝ/ symbol for transcription. For unstressed syllables, such as the "er" of "fath*er*," you instead use the IPA symbol schwar /ɚ/.

R-coloring is a perceptual quality that can be reached in a number of ways. R-coloring demonstrates the property of *compensatory articulation*, that a given acoustic goal can be reached by a number of different mouth positions.

R-coloring can differ substantially among individual speakers. Some make a *retroflex* gesture, putting the tip of the tongue against the rear of the alveolar ridge, while others hump the tongue in the middle of the mouth, sometimes called American *bunched r*. These vowel gestures are very similar to the consonant /ɹ/ in English and are described in detail in Chapter 6.

A useful series of r-colored vowels can be elicited in the context /fVɹ/ where V stands for a vowel. Table 7-2 contains many of these items and some others, including common North American English and British English words. Try these words out and see how much rhoticization (r-coloring) you use.

Table 7-2 Vowel R-Coloring in Three Varieties of English

IPA	Example	American English	Canadian (similar to AE unless indicated)	British English (RP)
/i/	seer	/siɹ/ or /siɚ/	…	/siə/
/ɪ/	fear	/fɪɹ/	…	/fɪə/
/e/	payer	/peɹ/ or /peɚ/	…	/peə/
/ɛ/	fair	/fɛɹ/	/fer/	/fɛə/
/ɝ/	fur	/fɝ/	…	/fɜː/
/ʊ/	poor	/pʊɹ/	…	/pʊə/
/ɔ/	sore	/sɔɹ/	…	/sɔə/
/ɑ/	far	/fɑɹ/	…	/fɑː/
/aɪ/	fire	/faɪɹ/	/fʌɪɹ/	/faɪə/
/aʊ/	flower	/flaʊɹ/	…	/flaʊə/
/ɔɪ/	foyer	/fɔɪɚ/	…	/fɔɪə/

Different transcription systems may be used for non-rhotic forms of English, such as commonly found in parts of the United Kingdom, Ireland, South Africa, and the Caribbean. I give more detail on different accents in Chapter 18.

A symbol used to describe the central nonrhotic (stressed) vowel is /ɜ/ (reversed epsilon). You can find this vowel in

Received Pronunciation (RP) British for words such as "fur" and "bird."

Neutralizing in the right places

The vowels /o/ and /i/ make predictable changes in particular environments. Phoneticians have adopted conventions for transcribing these patterns. For example, take a look at these transcriptions (GAE accent):

sore	/sɔɹ/
selling	/ˈsɛlɪŋ/

Beginning transcribers are often puzzled as to why /ɔ/ is used in "s*o*re" (instead of /o/), and why /ɪ/ is used before /ŋ/ in words that end with -ing, such as "selling." The answer is that vowels are affected by their surrounding consonants. These effects are more pronounced with certain consonants, especially the liquids (/ɹ/ and /l/) and nasals (/m/, /n/, and /ŋ/). This results in *neutralization*, the merger of a contrast that otherwise exists. For example, /o/ and /ɔ/ sound quite distinct in the words "b*oa*t" and "b*ough*t" (at least in GAE). However, before /ɹ/ these vowels often neutralize, as the /ɹ/ has the effect of lowering and fronting the /ɔ/ toward the /o/. Front vowel examples include "t*ie*r" and "p*ie*r" (pronounced with /ɪ/). The same process can take place before /l/. Examples include "pill" and "peel," both produced as /pɪɫ/ in some accents.

Say "runn*ing*." Do you *really* make a tense /i/ as in "b*ea*t" during the final syllable? Probably not. For that matter, you're probably not making a very pure /ɪ/ either. You're neutralizing, making something in-between. To label this sound, phoneticians lean toward the lax member and label it /ɪ/. Thus, /ˈɹʌnɪŋ/.

It's the same principle with "s*o*re." You're probably not using a tense /o/, such as in "b*oa*t." Listen closely! The closest vowel that qualifies is /ɔ/, even though its quality is different when rhotic.

Tensing up, laxing out

This tense versus lax vowel difference is important for a number of applications in language instruction and clinical linguistics. Specifically, the tense-lax difference indicates whether a vowel can stand alone at the end of a stressed syllable (*tense*), or whether the syllable must be closed off by a consonant at the end (*lax*). Many languages (such as Spanish) don't have any of the English lax vowels, and native speakers will therefore have difficulty learning them when studying English as a second language.

TRY IT Take a look at these word pairs and pronounce them, one by one.

"beat" versus "bit"

"bait" versus "bet"

"Luke" versus "look"

Can you hear a systematic change in the sound of each pair? The first member of each pair is tense, and the second member, lax. This distinction was originally thought to result from how the vowels were made, muscularly. However, these differences are now understood as relating to English *phonology* (system of sound rules). Refer to Table 7-3 for examples.

Table 7-3 Distribution of English Tense and Lax Vowels

	Vowel	Stressed Open Syllable	Closed Syllable
Tense	/i/	bee /bi/	beat /bit/
Lax	/ɪ/	bih /bɪ/ (not a real word)	bit /bɪt/

The tense vowel /i/ can appear in a stressed open syllable word such as "b*ee*," or in a syllable closed with a consonant at the end, such as "b*ea*t." If you try to leave a lax vowel in a stressed open syllable (such as the made-up word "bih"), you end up with something very un-English-like. You can pronounce such a word, but it will sound like something from another language. The same is true with /ɛ/, /æ/, /ʊ/, and /ʌ/. You can't really go around saying "That is *veh*. I appreciate your *geh* very much."

Because of this restriction of not being able to appear in stressed open syllables, /ɪ/, /ɛ/, /æ/, /ʊ/ and /ʌ/, as in "h*i*d," "h*ea*d," "h*a*d," "h*oo*d," and "m*u*d" are called the *lax* vowels of English. Most phoneticians consider the vowels /ɑ/, /i/, /u/, /e/, and /o/ to be the *tense* vowels. These vowels are produced more at the edges of the vowel space (less centralized) than their lax counterparts. You can hear the difference between these tense vowels and their corresponding lax member in the pairs /i/ and /ɪ/, /e/ and /ɛ/, and /u/ and /ʊ/. If you say these in pairs, you should be able to hear both a difference in quality and quantity (with the lax member being shorter in duration). The /ɑ/ and /o/ tense vowels don't really have a lax member to pair up with (oh, well — somebody has to stay single!).

RP: Received from *whom*, exactly?

RP, the abbreviation for the *Received Pronunciation*, is a prestigious accent spoken by approximately 2 to 3 percent of people in the United Kingdom. This includes the royal family and some members of the government and the media. The term *RP* is usually credited to phonetician Daniel Jones, although the usage can be traced back earlier. The idea of *received* means approved, like *received wisdom*. RP is often associated with the south of England, but is actually a blend of speech from East Midlands, Middlesex, and Essex. Historically, RP was common at Oxford University. As more British families sent their children to the British public schools during the 19th century (similar to American private schools), this accent took hold as a symbol of access to this world of education and power. Currently, RP is admired in some circles and viewed negatively in others.

It's important to remember that a person can speak grammatically correct English with a Standard English dialect (often called SE, for Britain) without speaking the RP accent. Because the RP accent is well documented and is used in dictionaries, it's frequently referenced here and in other phonetics texts. Chapter 18 gives more on this dialect (and other British varieties).

Most forms of British English have one more lax vowel than American English, /ɒ/ called *turned script a* in IPA. This is an open, back rounded vowel, as in RP "c*o*d" and "c*o*mmon." It can't appear in stressed open syllables and is lax.

Sorting the Yanks from the Brits

Phoneticians focus on the sound-based aspect of language and don't fret about the spelling, syntax (grammar), or vocabulary differences between North American and British varieties. This helps narrow down the issues to the world of phonetics and phonology.

In terms of vowels, you need to consider other issues than just the presence or absence of rhotics. There are quality differences in monophthongs as well as different patterns of diphthongization depending on which side of the pond you live. These sections take a closer look at these differences.

Differentiating vowel sounds

For front vowels (ranging from /i/ to /æ/), both North American English and British English have sounds spaced in fairly equal steps (perceptually). You should be able to hear this spacing as you pronounce the words "heed," "hid," "head," and "had." Try it and see if you agree.

Things get testy, however, with the vowel /e/. English /e/ is transcribed as /eɪ/ by many phoneticians (especially in open syllables) because this vowel is typically realized as a diphthong, beginning with /e/ and ending higher, usually around /ɪ/. This is shown in a traditional vowel quadrilateral (Figure 7-2a). Overall, the amount of diphthongal change for American /eɪ/ is less than that found for the major English diphthongs /aɪ/, /aʊ/, and /ɔɪ/.

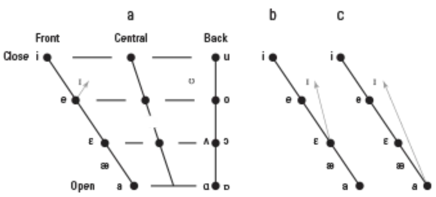

Illustration by Wiley, Composition Services Graphics

Figure 7-2: Vowel quadrilateral showing different offglides used for varieties of GAE (both a and b) and British English (c).

However, talkers vary with respect to where they really start from. Fine-grained studies of American English talkers suggest that many people start from lower vowel positions, producing words like “great” as /gɹɛɪt/. The trajectory of this diphthong is shown in Figure 7-2b. Forms of English spoken in the United Kingdom have different trajectory patterns. The direction of the /e/ diphthong for RP is similar to the direction of the GAE /eɪ/, but extends slightly further (not shown in figure).

Other British dialects have larger diphthong changes, including London accents sometimes called *Estuary English* (see Chapter 18). These upstarts (named for people living around the Thames, not birds), produce /seɪ/ “say” sounding more like /saɪ/. A panel showing the diphthong trajectories of this accent is shown in Figure 7-2c. Not to be outdone, the Scots arrive at a vowel like the Japanese, doing away with a diphthong altogether and instead producing a high monophthong that can be transcribed [e]:

> “Which way (should we go to Lochwinnoch?) [ʍɪtʃ we:] . . .

There are even more North American versus British differences in the mid and back vowels. Starting with the mid vowel /ʌ/, British speakers produce the vowel lower than their North American counterparts. This is likely due to the fact that British talkers distinguish words like “bud” and “bird” by distinguishing between low /ʌ/ and the higher mid-central vowel /ɜ/. However, North American talkers use a rhotic distinction (/ʌ/ versus /ɝ/) and don’t require this height separation.

North American talkers show regional differences among the back vowels, particularly for the notorious pair /ɑ/ and /ɔ/. The tendencies are either to merge the two toward /ɑ/ (Southern California) or closer to /ɔ/ (Northern American dialects). Most speakers of British English have added another vowel to the mix: high back rounded /ɒ/.

Table 7-4 shows some examples of these British back vowel distinctions so you can get grounded in the differences. This may be especially helpful if you're interested in working on accents for acting, singing, or other performance purposes. (I also include URLs where you can listen to audio files.)

Table 7-4 English Back Vowels – British SE

Back Vowel	Examples	IPA	URL to Listen to
/ɑ/	balm	/bɑm/	www.utdallas.edu/~wkatz/PFD/balm.wav
/ɒ/	bomb	/bɒm/	www.utdallas.edu/~wkatz/PFD/bomb.wav
/ɔ/	bought	/bɔt/	www.utdallas.edu/~wkatz/PFD/bought.wav

These differences provide an insight into the challenges facing people trying to master new accents. Namely, it's difficult moving from an accent with fewer distinctions (such as no difference between /ɑ/ and /ɔ/) to an accent with more distinctions. This is not only because the learner must use more sounds but also because the distribution of these sounds isn't always straightforward.

For example, British RP accent uses an /ɑ:/ sound for many words that American English uses an /æ/ for. For instance, "glass" and "laugh" (/glɑ:s/ and /lɑ:f/). However, speakers of RP pronounce "gas" and "lamp" the same as in GAE, with /æ/. Thus, a common mistake for GAE speakers attempting RP is to overdo it, producing "gas" as /gɑ:s/. Actually, there is no easy way to know which RP words take /ɑ:/ and which take /æ/, except to memorize.

Notice that it's not as tricky to go in the opposite direction, from more accent distinctions to less. For example, a British RP speaker trying to imitate a California surfer could simply insert an /ɑ/ vowel for "b*o*mb," "b*al*m," and "b*ou*ght" and probably get away with it. But could that British person actually surf?

English has a diphthongal quality to the tense vowels /e/, /i/, /o/, and /u/, particularly in open syllables. For this reason, these vowels are often transcribed /eɪ/, /ij/, /oʊ/, and /uw/ (see also Chapter 2).

Dropping your "r"s and finding them again

Rhotic and non-rhotic accents are a bit more complicated than is indicated in the "Coloring with an 'r'" section, earlier in the chapter. Many of the nonrhotic accents (they don't pronounce an "r" at the end of a syllable) express an /ɹ/ under certain interesting circumstances.

A *linking-r* occurs if another morpheme beginning with a vowel sound closely follows nonrhotic sounds. This is typical of some British accents, but not American Southern States. Here are a couple examples.

Example Word	***British SE***	***American Southern States***
care	/keə/	/keə/
care about	/ˈkeɚəbaʊt/	/ˈkeəʔ əbaʊt/

A similar-sounding process is *intrusive-r*, the result of sound rules trying to fix things that really aren't broken. For these cases, such as *law-r-and-order,* an "r" is inserted either to fix the emptiness *(hiatus)* between two vowels in a row, or to serve as a *linking*-r that was never really there in the first place (for example, if "tuna oil" is pronounced "tun*er* oil"). Table 7-5 shows some examples. I also include URLs where you can listen to audio files.

Table 7-5 Examples of Linking-r

Phrase	***IPA***	***URL***
Australia or New Zealand	/ɒs ˈtɚɹlɪɚ ɔ:nju: ˈzi:lṇd/	www.utdallas.edu/~wkatz/PFD/Linking_R1.wav
There's a comma after that.	/ðəzə ˈkɒmɚ ɑ:ftə θæt/	www.utdallas.edu/~wkatz/PFD/Linking_R2.wav
Draw all the flowers	/drɔ:ɹ ˈɔ:l ðə flaʊəz/	www.utdallas.edu/~wkatz/PFD/Linking_R3.wav

Noticing offglides and onglides

There are a number of different ways to describe the dynamic movement of sound within a vowel. One way, as I describe in Chapter 2, is to classify vowels as monophthongs, diphthongs, or triphthongs. This description takes into account the number of varying sound qualities within a vowel. Phoneticians also note which part of the diphthongs (the end or the beginning) is the most prominent (or unchanging). This distinction is commonly referred to as offglides and onglides:

- **Offglides:** If the more prominent portion is the first vowel (as in /aɪ/), the second (nonsyllabic) part is the *offglide*. This idea of an offglide also provides a handy way to mark many types of diphthongs that you may find across different accents. For instance, in American Southern States accents, lax /æ/ becomes /ɛə/ or /eə/. That is, they are transcribed including a /ə/ offglide. Table 7-6 shows some examples with URLs to audio files.

Table 7-6 Vowels Produced with an Offglide

Example Word	*IPA*	*URL*
lamp	/leəmp/	www.utdallas.edu/~wkatz/PFD/lamp-offglide.wav
gas	/geəs/	www.utdallas.edu/~wkatz/PFD/gas-offglide.wav

Some phoneticians denote an offglide with a full-sized character (such as /eə/), while others place the offglide symbol in superscript (such as /eə/).

- **Onglides:** An *onglide* is a transitional sound in which the prominent portion is at the end of the syllable. These sounds begin with a constriction and end with a more open, vowel quality.

  An example of an onglide in English would be the /j/ portion of /ju/. Some phoneticians treat this unit as a diphthong, while a more traditional approach is to consider this syllable a combination of an approximant consonant followed by a vowel.

Doubling Down on Diphthongs

American English and British English accents have in common a set of three major diphthongs, /aɪ/, /aʊ/, and /ɔɪ/. These are called *closing diphthongs*, because their second element is higher than the first (the mouth becomes more closed). You can see the three major diphthongs (similar in GAE and British English) in Figure 7-3a, and a minor diphthong (found in British English) in Figure 7-3b. The /aɪ/, /aʊ/, and /ɔɪ/ diphthongs are also called *wide* (instead of narrow) because they involve a large movement between their initial and final elements.

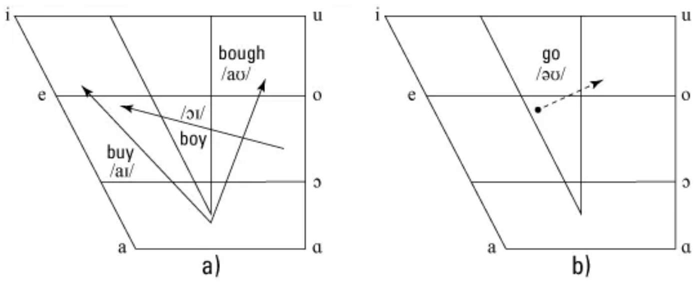

Illustration by Wiley, Composition Services Graphics

Figure 7-3: Diphthongs found in both GAE and British English (a), and in only British English (b).

Considering first /aʊ/, as in "cow," a similar trajectory is seen in BBC broadcaster English as in GAE. The /aʊ/ diphthong is also called a *backing* diphthong because posterior tongue movement is involved when moving from /a/ to /ʊ/. As may be expected, there are many variants on this sound, especially in some of the London accents (which can sound like gliding through /ɛ/, /ʌ/, /u/ or /æ/, /ə/, and /ʊ/).

The /aɪ/ sound is a *fronting* diphthong. An important thing to remember about this sound is that few talkers will reach all the way up to a tense /i/ for the offglide; it's usually /ɪ/. A second fronting diphthong found in British English and American English accents begins in the mid back regions. This is the diphthong /ɔɪ/, as in "b*oy*," "Fl*oy*d," and "*oi*l".

An interesting diphthong found in British accents (but *not* in GAE) is the closing diphthong /əʊ/. Look at the dotted line in Figure 7-3b. This sound is found in place of the GAE tense vowel /o/. Because it doesn't have much of a sound change, it would qualify as a narrow diphthong. Table 7-7 shows some examples. You can also check out the audio files.

Table 7-7 How to Say "o" in GAE and RP

Example Word	*GAE*	*RP*	*URL*
go	/goʊ/	/gəʊ/	www.utdallas.edu/~wkatz/PFD/go-GAE-RP.wav
so	/soʊ/	/səʊ/	www.utdallas.edu/~wkatz/PFD/sew-GAE-RP.wav

Lengthening and Shortening: The Rules

This section concentrates on *vowel length*, namely how a given vowel's length changes as a function of context. Such context-conditioned change is called *allophonic variation* (see Chapter 5 for more information).

REMEMBER If you're an English speaker, you naturally carry out at least three subtle timing changes for vowels when you speak. Here, I note these processes formally as *rules*. This information can come in handy if you teach English as a second language, compare English to other languages, or engage in any work where you need to be able to explain what the English sound system is doing (instead of, say, stamping your feet and saying "because that's just how it is!").

Vowel Inherent Spectral Change (VISC): See you in court?

Monophthongs don't have as much quality change as diphthongs and triphthongs. However, exciting new research by Professors Terrance Nearey, Peter Assmann, and others has shown that in many languages (particularly those with large vowel inventories) monophthongs also demonstrate substantially changing sound patterns. This information is called *vowel inherent spectral change (VISC)*. It seems to be important for human vowel perception, affecting speech development, second language learning, and dialect change.

Research by Professor Geoffrey Stewart Morrison has shown that VISC can provide useful information for forensic voice comparison. Although this work is still in an early stage, the goal would be to boost speaker identification using VISC from recordings of a subject's speech.

Check out each rule and its examples:

- **Rule No. 1:** Vowels are longest in open syllables, shorter in syllables closed by a voiced consonant and shortest when in syllables closed by a voiceless consonant. For example:

  "bay" (/beɪ/)

  "bayed" (/bed/)

  "bait" (/bet/)

- **Rule No. 2:** Vowels are longer in stressed syllables.

  For example:

  "re*peat*" (/ɹəˈpit/

  "to *re*peat" (/ˈɹipit/)

  Here, "peat" (/pit/) should sound longer in the first than the second example.

- **Rule No. 3:** Vowels get shorter as syllables are added to a word (up to three syllable-words). For example:

  "zip" (/zɪp/)

  "zipper" (/ˈzɪpɚ/)

  "zippering" (/ˈzɪpɚɪŋ/)

Chapter 8

Getting Narrow with Phonology

In This Chapter

- Digging into phonology
- Sorting out types of transcription
- Getting a sense of rule ordering and morphophonology

As you study phonetics, many of the IPA symbols and the sounds of English will become warmer and cozier, as you become more familiar with them. You can look at symbols, such as /æ/ and /ʃ/ and know they represent sounds in the words "c*a*t" and "*sh*out." To help you be more comfortable, you need a firm grasp of the relationship between phonetics and phonology, which allows you to move between broad and narrow transcription. This chapter helps clarify how phonetics and phonology are related, which can help you take your transcriptions to the next level.

Phonetics is the study of the sounds of language. Phonetics describes how speech sounds are produced, represented as sound waves, heard, and interpreted. Phonetics works hand-in-hand with *phonology*, the study of the sound systems and rules in language.

Phonologists typically describe the sound processes of language in terms of *phonological rules*, patterns that are *implicit* (naturally understood) by speakers of the language. For example, a speaker of English naturally (implicitly) nasalizes a vowel before a nasal consonant, as in “run” ([ɹʌ̃n]) and “dam” ([dæ̃m]). English speakers nasalize a vowel even for a nonsense words, such as “zint” ([zɪ̃nt]) and “lemp” ([lɛ̃mp]).

If you’re a native English speaker, you’ll also nasalize the vowels in these examples. Go ahead and speak these nonsense words (“zint” and “lemp”). Word meaning has nothing to do with it; it’s a sound thing!

Part of knowing a language entails you understand and use its phonology, processes that can be described so you’ll be able to incorporate information about language sound rules into your transcriptions. The following sections explain the main kinds of transcriptions and how they differ.

Distinguishing Types of Transcription

Phonetic transcription uses symbols to represent speech sounds. However, depending on your need, you can transcribe in many different ways. A transcription can look quite different based on whether you'll use it for theoretical linguistics, language teaching, speech technology, drama, or speech and language pathology. Here are some important distinctions used to classify the main types of transcriptions.

Impressionistic versus systematic

The transcriber's knowledge can play a key role in two main types of transcription classifications. They are:

- **Impressionistic:** An *impressionistic* transcription occurs when you, as the transcriber, have minimal knowledge of the language, dialect, or talker being worked with. As such, you'll use your minimal experience to make judgments about the incoming sounds. An example would be somebody trying a first transcription of a complex African language. In such a situation, the transcriber could only hope to describe the new language in terms of the categories of his or her native language. The results probably wouldn't be very accurate because the transcriber wouldn't know which details would turn out to be important.

- **Systematic:** In contrast, if you, as the transcriber, are well trained in phonetics and had made several passes over the new language, you can note important details. This transcription would be *systematic*, reflecting the structure of the language under description.

Impressionistic and systematic are therefore endpoints on a continuum. The more detailed and accurate the transcription, the more it moves from impressionistic to systematic.

Broad versus narrow

Transcription can also be classified as simple or detailed, as the following explains:

- **Broad**: The simpler your transcription (with the less phonetic detail), the more broad it is. *Broad* transcription has the advantage of keeping the material less complicated. Although a broad transcription is sufficient for many applications and you can complete a broad transcription with less phonetic training, you basically get what you pay for. If you want to later go back to these transcriptions and reproduce the fine details, you'll probably be out of luck.
- **Narrow:** A maximally *narrow* transcription indicates all the phonetic detail that is available and relevant. Completing a narrow transcription requires more training than simply knowing IPA characters: You must know something about the phonology of the language and the diacritics typically used to designate *allophones* (contextually-related sound variants). Narrow transcriptions offer substantial detail, useful for scientific and technical work. Making sure that such

transcriptions don't become needlessly cluttered is important; otherwise, readers may have a nightmare getting through it.

Like the preceding section (on impressionistic and systemic dichotomy), the broad and narrow contrast can be best thought of in terms of a continuum. That is, a transcription can range from broad to narrow.

Capturing Universal Processes

Just as phonetics has a universal slant (to describe the speech sounds of *language* — as in all of the languages of the world), phonology also seeks to describe the sound processes of all the world's languages. This emphasis on universal goals has affected how phonetics and phonology are taught worldwide. For example, whereas phonetics and phonology used to be taught predominantly within the auspices of particular language and literature departments (such as English and the Slavic languages), they're now frequently integrated with linguistic, cognitive, and brain sciences because of the assumption that speech and language are universal human properties.

Getting More Alike: Assimilation

One of the most universal of sound phonological rules in language is *assimilation*, when neighboring sound segments become more similar in their production. They're frequently called *harmony* processes.

At a physiological level, you can describe assimilation as *coarticulation* — the fact that the articulators for one sound are influenced by those of a surrounding sound. Speech is *co-produced* — an upcoming sound can influence an articulator or set of articulators (an *anticipatory coarticulation*), and a given sound often has leftover influences from a sound that was just made (referred to as a *perseverative coarticulation*). The result is the same; sounds next to each other becoming more similar. Chapter 4 gives more information on anticipatory and perspective coarticulation.

Table 8-1 shows some major varieties of assimilation.

Table 8-1 Assimilation in Action

Example	*Realization (IPA)*	*Explanation*	*Assimilation Type*
ba*d* *g*uy	[bæg̚ gaɪ]	Place of an alveolar stop assimilates to the place of the following consonant.	*Regressive* (or right to left, anticipatory coarticulation)
captain	[ˈkʰæpm̩]	Place of preceding consonant has an effect on the place of a following one.	*Progressive* (or left to right, perseverative coarticulation)
pan	[pʰæ̃n]	The vowel becomes nasalized before a nasal consonant.	*Similitude* (the phone still sounds like it's in the same category)
sandwich	[ˈsæ̃mɪtʃ]	The /n/ and /w/ are coproduced and fused into a /m/ gesture.	*Coalescence*, also referred to as *fusion* (two sounds merge to form a new segment)

From this table, notice assimilation can proceed in two directions.

- In the first example, "bad guy," a sound segment [g] modifies an earlier sound, which is called *regressive* (or right-to-left) assimilation. You can see a similar direction in the word, "pan," although the process results in a sound just having a slight change that doesn't alter its phonemic status (referred to as *similitude*).

- In contrast, “captain” goes in the opposite direction. The production of [p] affects the place of articulation of the following nasal, [m], a progressive or left-to-right effect. *Progressive* means that a given sound affects the sound following it.
- Finally, “sandwich” illustrates a fusion of two sounds (/n/ and /w/) to result in /m/. This is called *coalescence* because the result of having two distinct phonemes affect each other is a third, different sound.

These examples come from English where harmony cases are local. However, languages such as Turkish and Hungarian have long distance vowel assimilation because these processes cross more than one segment. Refer to the nearby sidebar for a closer look at Hungarian.

Transcribing Hungarian: If you’re into hard-core assimilation

Hungarian has a set of front versus back vowels. When a word root contains a back vowel, it must take a suffix with a back vowel. If the word root ends in a front vowel, the suffix must contain a front vowel. Thus, to form the *dative* (indirect object, as in I gave it *to him*) part of speech, Hungarian would form words like this:

Vowel Type	Root	Dative	Gloss
Back vowel	/fal/	/falnak/	"wall"
Front vowel	/kert/	/kertnek/	"garden"

A front vowel triggers assimilation to a front vowel later in the word, and a back vowel triggers a back vowel later in the word. (Similar processes also take place in Turkish, Finnish, and a number of other languages.) For instance, take a gander at these two Hungarian word endings and see if you can spot the regressive assimilation processes at work:

Vowel Type	Root	Dative	Gloss
Back vowel	/had/	/hadnak/	"army"
Front vowel	/hit/	/hitnek/	"belief"

Getting More Different: Dissimilation

Dissimilation is a process where two close sounds become less alike with respect to some property. In dissimilation, sounds march to a different drummer and become less similar. For instance, if a language requires sounds next to each other that are difficult to produce, dissimilation processes come into play so that the final realizations are bold, clean, and producible.

An example is the word “diphthong,” which should be pronounced [ˈdɪfθɔ̃ŋ], but is frequently mispronounced [ˈdɪpθɔ̃ŋ]. In fact, many people end up misspelling it as “di*p*thong” for this (mispronunciation) reason.

Because producing a /f/ followed by a /θ/ is difficult (go ahead and try it), two fricatives in a row change to a stop followed by a fricative. Dissimilation isn’t quite as common among languages as assimilation.

Putting Stuff In and Out

Processes of *insertion* (also called *epenthesis*) cause a segment not present at the phonemic level to be added. In other words, an unwanted sound gets added to a word.

A common example in English is the insertion of a voiceless stop between a nasal stop and voiceless fricative. Here are some examples:

Example	***Phonemic Level***	***Phonetic Level***
strength	/stɹɛŋθ/	[stɹɛ̃ŋkθ]
hamster	/ˈhæmstɚ/	[ˈhæ̃mpstɚ]

Another form of insertion sometimes noted in the language classroom occurs with consonant clusters. Native speakers of languages such as Japanese or Mandarin who don’t have consonant clusters (such as pl-, kl-, spr-, or -lk) sometimes insert a vowel between the consonants to make the sounds more like their native phonology. Thus, a Japanese speaker learning English may pronounce the following English words with these epenthetic vowels inserted (in *italics*):

Example	***Phonemic Level***	***Phonetic Level***
spoon	/spun/	[suˈpu:nu]
ski slope	/ski.sloʊp/	[suˈki:suˈɹo:pu]

Deletion rules eliminate a sound. An example in English is called h-dropping (or /h/*-deletion*). Try and say this sentence, “I sat on his horse.” Which of the following two work?

✓ [aɪ ˈsæɾɔ̃n ɪs ɔrs]

✓ [aɪ ˈsæt ɔ̃n hɪz hɔrs]

Probably the first is more natural, where /h/ is deleted from “his” and “horse.”

Moving Things Around: Metathesis

In *metathesis,* a speaker changes the order of sounds. Basically, one sound is swapped for another. Check out these examples:

Example	*Phonemic Level*	*Phonetic Level*	*Found*
asked	/æskt/	[ækst]	Southern States' dialects and African American Vernacular English (AAVE)
animal	/ˈænɪməl/	[ˈæ̃mĩnəɫ]	Child language

Putting the Rules Together

Some phonological rules depend on others and either set up another rule to operate or deprive them of their chance. The rules in this chapter can all be represented with a basic format:

A → B/C __ (D)

A becomes B, in the environment after C or before D.

With this format, the following clarify what each letter stands for:

- **A:** The letter on the left side of the arrow is called the *structural description*. This is the sound (at the phonemic level) before anything happens to it.
- **B:** The letter to the right of the arrow is the *structural change*. It's the result of a sound change occurring in a certain environment.
- **C** and **D:** They represent that environment where the sound change occurs.

From the earlier section, "Getting More Alike: Assimilation," I now show the examples in phonological rule format here:

Example	*Structural Description (A)*	*Structural Change (B)*	*Phonological Rule*
ba*d* guy	/bæd gaɪ/	[bæg gaɪ]	Alveolar → + velar/_ velar
pan	/pæn/	[pæ̃n]	Vowel→ + nasal/_ nasal

REMEMBER Instead of writing out lengthy prose, you can use rules to represent phonological processes. For example, for "pan" the rule is that a vowel becomes nasalized in the environment before a nasal.

Consider for a moment how the (tricky) English plurals are pronounced in most words. Although the plural marker -s or -es is used in spelling, it doesn't always result in an [s] pronunciation. Rather, a plural is sometimes pronounced as [s], sometimes as [z], and sometimes as [ɪz], depending on the final sound of the root, as the following examples demonstrate:

Singular	*IPA of Plural Form*	*Suffix*
ra*t*	/ɹæts/	[s]
da*d*	/dædz/	[z]
di*sh*	/ˈdɪʃɪz/	[ɪz]

REMEMBER The plural "s" is a kind of *morpheme*, the smallest meaningful units in language. The study of how *morphemes* show regular sound change is called *morpho-phonology*. To make the English plural system work, phonologists make two assumptions:

- /z/ is the underlying form of the plural marker.
- Two rules must apply and apply in the correct order.

Table 8-2 specifies these two rules.

Table 8-2 Two Rules of Morphophonology

Rule	*Formula*	*Translation*
Rule No. 1	Insertion: /0/→[ɪ] / [+ sibilant] __ [+ sibilant]	[ɪ] is inserted between two sibilants.
Rule No. 2	Assimilation: /z/ → [-voiced]/[-voiced, + cons] ____#	[z] becomes devoiced after a voiceless consonant at the end of the word.

In Rule No. 2, the hashmark (#) is an abbreviation for boundary at the end of a word.

TIP You must apply the rules in order for the system to work. If you don't, it bombs. Consider the word "dishes" that ends with [ɪz]:

Singular: [dɪʃ]

Plural: [ˈdɪʃɪz]

Table 8-3 shows correct rules applied in the order. However, the reverse order with Rule No. 2 first doesn't give the right answer. Assimilation changes the /z/ to /s/, then insertion changes the /s/ to /ɪs/, yielding [ˈdɪʃɪs] (incorrect).

Table 8-3 Insertion and Assimilation in Action

First Order (Correct)		*Second Order (Incorrect)*	
Underlying representation	/dɪʃ/ +/z/	Underlying representation	/dɪʃ/ +/z/
Rule No. 1: Insertion	/dɪʃ/ +/ɪz/	Rule No. 2: Assimilation	/dɪʃ /+/s/
Rule No. 2: Assimilation	N/A	Rule No. 1: Insertion	/dɪʃ/ + /ɪs/
Phonetic realization	[ˈdɪʃɪz]	Phonetic realization	[ˈdɪʃ + ɪs]*

Chapter 9

Perusing the Phonological Rules of English

In This Chapter

- Narrowing in on consonant allophones
- Recognizing principled change in vowels
- Getting rule application *just right!*

Phonological rules describe sound processes in language that are naturally understood by speakers and listeners. In order to transcribe well, particularly when completing narrow transcription, it's important to understand these sound processes and describe their output using the correct symbols in the International Phonetic Alphabet (IPA).

Phonological rules take the following form:

Structural description → Structural change /___ (in some environment)

The *structural description* is the condition that the rule applies to. The *structural change* is the result of the rule, occurring in a specific phonetic context. The arrow shows that a given input sound (the structural description) changes or becomes modified in some environment.

A phonological rule can be described in a short description or in a formula. To keep things simple, in this chapter I focus on descriptions upfront to help you understand. I also include a few technical formulas as secondary information. Make sure to check out Chapter 8 for more background about phonology and phonological rules.

There is no set number of phonological rules for any given language. In this chapter, I use 13 phonological rules to capture some of the most important regularities of English phonology. These rules describe *implicit* (naturally understood) processes of a language. The exact numbering doesn't really matter: I group these rules into sections to make them easier to memorize.

As a transcriber, use these rules as a guide for what talkers likely do, but let your ear be the final judge for what you end up transcribing.

Rule No. 1: Stop Consonant Aspiration

A traditional first rule in phonetics is that English voiceless stops, which are /p/, /t/, and /k/, become aspirated when stressed and syllable initial (at the beginning of a syllable). This rule captures that fact that the phoneme /t/ is represented by the aspirated allophone [t^h] under these specific conditions.

Each phonological rule usually has an IPA diacritic or symbol involved. As a result, I list relevant diacritics and symbols following each rule. I also provide some examples, and I encourage you to generate your own. The diacritic for Rule No. 1 is [^{h}]. Here are some examples:

*p*eace [p^{h}is]

at*t*ire [əˈt^{h}aɪɹ]

*k*iss [k^{h}ɪs]

This rule captures one of the essential properties of English phonology. Try saying each word while holding your hand under your mouth (near your bottom lip) and you should feel a puff of air that is the aspiration of the [p^h], [t^h], and [k^h].

Monosyllabic words, those words that have just one syllable, such as “peace” and “kiss,” are easy. However, in *polysyllabic words,* words with multiple syllables, things get a bit more complicated. Aspiration is stronger in stressed syllables than unstressed (see Chapter 6 for further discussion), which means in polysyllabic words the aspiration rule applies chiefly to stressed syllables. Otherwise, the /p/, /t/, and /k/ consonants are released, but not stressed. Here are some examples of polysyllabic words:

catapult [ˈk^{h}æɾəpəɫt]

repulsive [ɹəˈp^{h}ʌɫsɪv]

Try the aspiration test, feeling for an air puff, when saying “catapult” and “repulsive.” Aspiration is on the initial stop in “catapult” because the [k^h] is syllable-initial and stressed. However, even though the [p] in “cata*p*ult” is syllable initial, it isn’t aspirated. It’s only released. In “re*p*ulsive,” the [p] is aspirated because it’s syllable-initial and stressed (even though it’s not word initial).

Aspiration for English /p/, /t/, and /k/ generally isn't as strong when word-initial than, for example, when following another word. *Word initial* means at the beginning of a word, so the [p^{h}] in "*p*ie" generally has less aspiration than the [p^{h}] in "the *p*ie." For this reason, you may see different conventions used by phoneticians when marking aspiration in narrow transcriptions at the beginning of words. Some mark it and others don't. In this book, I mark aspiration at the beginning of a word, according to Rule No. 1.

Table 9-1 includes some practice items containing /p/, /t/, and /k/. Mark the aspiration using narrow transcription in column three. I have done the first one for you. Ready?

Table 9-1 Stop Consonant Aspiration Practice

Example Word	*Broad*	*Narrow*
appear	/əˈpɪɹ/	[əˈp^{h}ɪɹ]
khaki	/ˈkæki/	
uncouth	/ənˈkuθ/	

The answers are as follows:

- **khaki:** You should only have marked the initial [k^{h}] of "khaki" as aspirated because that "k" is stressed and syllable initial. The second [k] is released but not aspirated.
- **uncouth:** The [k] is aspirated because it's stressed and syllable initial, even though it's the final syllable in the word.

If you're more into formulas, you can write Rule No. 1 as:

C [+stop, –voice] → [+aspiration]/ #____ [+ syllable, + stress], (where # = boundary)

Here's how to read this formula. "A consonant (that is a stop and is voiceless) becomes aspirated in the environment at the beginning of a stressed syllable." Or more simply, stop consonants are aspirated in stressed syllable-initial position.

Rule No. 2: Aspiration Blocked by /s/

Another rule of phonetics is voiceless stops become unaspirated after /s/ at the beginning of a syllable. Because English has many consonant clusters (groups of consonants in a row, such as [spɹ] and [sk]), some phonologists consider this an important rule to remember. Others note that it overlaps with Rule No. 1. I emphasize this rule because it shows the importance of rule interaction. ***Note:*** There really is no diacritic or symbol for this rule because a feature is being blocked, not added.

English syllable-initial, s-containing consonant clusters (sp-, st-, and sk-) all share something in common: the production of the /s/ blocks the following stop consonant from having much aspiration. Try the following examples of minimal pairs, putting your hand near your mouth for the aspiration test.

*p*ill [p^hɪɫ]	*sp*ill [spɪɫ]
*t*ill [t^hɪɫ]	*st*ill [stɪɫ]

*k*ale [kʰeɫ] *sc*ale [skeɫ]

Notice that this rule would *not* apply in words such as "wa*sp*," "wri*st,*" or "fla*sk*," where the s-clusters occur at the end of a syllable. In such cases, the structural description isn't met and the rule isn't relevant. In words such as "whi*sp*er" (s-cluster in the medial position), the rule does apply because the stop comes after /s/ and at the beginning of a syllable. Try the aspiration test for "whisper" and see for yourself! No aspiration should be notable on the [p].

If you enjoy formulas, you can write phonological Rule No. 2 as follows:

C [+stop, -voice] → [-aspiration]/ #s__ [+ syllable]

A rough description of this rule would be: "Consonants that are stops and voiceless don't become aspirated when following an /s/ at the beginning of a syllable." Or more simply, voiceless stops become unaspirated after syllable-initial /s/.

Rule No. 3: Approximant Partial Devoicing

Devoicing rules are a rather depressing thing for phonetics teachers to talk about because it reminds them that life can get really complicated. When someone first starts to study phonetics, voicing is a comfortable, solid binary feature. A *phone* (speech sound) is defined as voiced or voiceless, end of story. However, then a dirty little secret comes out: Under certain conditions, some sounds may become partially *devoiced* (spoken with less buzzing of the vocal folds) because of biomechanical and timing reasons.

If you're making an aspirated stop such as [pʰ] in "*p*ay," the aspiration will affect the following approximant, such as if you say "p*r*ay" or "*pl*ay." For "pray" or "play" the vocal folds won't have time to fully buzz for the [ɹ] and [l], resulting in partial devoicing.

The diacritic for partial devoicing is a small circle placed under the sound, [̥]. Here are some examples for Rule No. 3:

*pr*ay [pʰɹ̥e]

*cl*ass [kʰl̥æs]

*tw*ice [tʰw̥aɪs]

*cu*te [kʰj̊ut]

You can get a good sense of how this works by placing your hand over your voice box to feel the buzzing while you say the following examples of minimal pairs.

*r*ay [ɹe] — p*r*ay [pʰɹ̥e]

*l*ass [læs] — c*l*ass [kʰl̥æs]

*w*eak [wik] — t*w*eak [tʰw̥ik]

you [ju] — *cu*e [kʰj̊u]

You should feel a longer period of buzzing for the (italicized) approximants in the list when they aren't preceded by a [pʰ], [tʰ], or [kʰ].

You place a small circle beneath the symbol as a diacritic indicating partial devoicing (unless the font is too cluttered by a downward-going symbol, such as /j/, in which case the symbol can be placed above the character).

If you're more into formulas, you can write phonological Rule No. 3 as:

C [+approximant] → [–voice]/ C [+ stop, +aspiration]

___

This formula reads "consonants that are approximants become partially devoiced in the environment following consonants that are stops and are aspirated." Or more practically, "approximants become (partially) devoiced after aspirated stops."

Rule No. 4: Stops Are Unreleased before Stops

A *release burst* occurs when a stop consonant closure is opened, producing a sudden impulse that is usually audible. In aspirated stops at the beginning of a syllable (like [p^h] in "*p*et" [p^{h}ɛt]), the vocal folds are apart, and there's aspiration (breathy, voiceless airflow) after the release of the stop. Try it and you can feel the aspiration on your hand or watch a candle blow out. English syllable-initial voiced stops (as in *b*et [bɛt]), also have a burst, but without aspiration and with a shorter voice onset time

(VOT, see Chapter 15 for more info). This release burst energy is weaker but is usually audible.

Release bursts may not be audible in other situations. These situations are referred to as *no audible release*. In syllable-final position, stop consonants can be optionally released. Try saying “tap/tab” in two speaking conditions:

- Quickly and casually: In casual speech, people usually produce no audible release for a syllable-final stop.
- Carefully, as if you were addressing a large audience that could barely make out what you were saying: More formal speech can override this no audible release condition. In formal speech, release characteristics are often emphasized for clarity or style.

When a person produces two stops in a row, the release characteristics become poorly audible. For instance, when saying “risked” [ɹɪsk˺t], just as the vocal tract is being configured for the release burst of the /k/, the tongue is also making closure for /t/, effectively cancelling out any sound of a released /k/. The diacritic for this rule is [˺]. Some examples are

risked [ɹɪsk˺t]

bumped [bʌm˺pt]

To see how hard it is to produce the word “risk” (with release) followed by a /t/, try these four steps:

1. Produce “risk” casually with no audible release.

risk [ɹɪsk]: No special diacritic is needed to mark lack of audible release.

2. Add the final /t/ to “risk.”

risked [ɹɪsk˺t]: This is the normal output of Rule 4.

3. Produce "risk" with a full release.

risk [ɹɪskʰ]: The aspiration diacritic is used only if the final release is strong enough to warrant it.

4. Add the final /t/ again.

risked [ɹɪskʰt]: Argh! This won't sound natural.

Rule No. 5: Glottal Stopping at Word Beginning

A rather surprising use of the glottal stop in English occurs before vowels at the beginning of a word or phrase. Unless you ease into an utterance (making some kind of ultra-calm announcement to zoned-out meditators at a Yoga retreat), you probably precede a vowel with a glottal stop. The IPA character that you need to remember for this rule is the glottal stop ([ʔ]).

Here are some examples. Try them and pay attention to whether your glottis is open or closed.

eye [ʔaɪ]

eaten [ˈʔiʔn̩]

Some phoneticians consider this rule in transcriptions and some don't. I use word-initial glottal stopping in the optional transcriptions listed in the audiovisual materials located at www.dummies.com/go/phoneticsfd.

Rule No. 6: Glottal Stopping at Word End

Voiceless stops are preceded by glottal stops after a vowel and at the end of a word. This rule also applies to word-final voiceless affricates. The IPA symbol involved in this rule is the glottal stop [ʔ]. Some examples include

steep [stiʔp]

pitch [pʰɪʔtʃ]

This rule is a use of a glottal stop that many English speakers don't believe at first, but eventually they'll accept. Before syllable-final /p/, /t/, /k/, or /ʧ/, many speakers of English restrict the flow of air at the glottis before getting to the stop itself (or at the same time as realizing the stop). Such timing doesn't occur if the final stop is voiced. Try these following words and see if you pronounce the voiceless stops in such a manner:

ri*p* [ɹɪʔp]

ri*ch* [ɹɪʔʧ]

rib [ɹɪb]

ridge [ɹɪʤ]

Whether glottal stop in this pre-consonantal position is transcribed or not is generally up to the discretion of the phonetician. Some capture this detail in narrow transcription and others don't. I provide this detail (as alternate transcriptions) in the audiovisual materials (located at www.dummies.com/go/phoneticsfd).

Rule No. 7: Glottal Stopping before Nasals

Here is another rule that describes the distribution of glottal stop: "Voiceless alveolar stops become glottal stops before a nasal in the same word." In other words, this rule captures the fact that /t/ and /d/ become [ʔ] in certain environments.

The symbol for this rule is the glottal stop [ʔ]. Say these words and think about what they all have in common:

ea*t*en [ˈiʔn̩]

wri*tt*en [ˈɹɪʔn̩]

bi*tt*en[ˈbɪʔn̩]

ro*tt*en[ˈɹɑʔn̩]

ki*tt*en[ˈkɪʔn̩]

glu*tt*on[ˈglʌʔn̩]

If you speak North American English, you'll almost certainly pronounce the medial /t/ phoneme as glottal stop [ʔ], followed by a syllabic nasal, indicated by placing a small line below the [n̩], described in Rule No. 9, explained later in this chapter.

Notice that none of these word examples involve an aspirated medial /t/ phoneme ([t^h]). Also, the stress pattern is *trochaic* (which means the syllable's stress is strong, then weak, sounding *loud*-soft (as in "*ri*fle" "*dou*ble", and "*ti*ger").

Rule No. 8: Tapping Your Alveolars

Alveolar stops (/t/ or /d/) become a voiced tap between a stressed vowel and an unstressed vowel. A *tap* (also called *flap* by some phoneticians, see Chapter 6) is a rapid articulation in which one articulator makes contact with another. Unlike a stop, there's not enough time to build up a release burst.

This rule involves the IPA symbol [ɾ], an English *allophone*. That is, a tap can't stand by itself anywhere in the language to change meaning. In English, a tap only occurs in certain environments, as specified by phonological rules. Here are some examples:

glottal [ˈglɑɾl̩]

Betty [ˈbɛɾi]

daddy [ˈdæɾi]

The stress patterns of the words involved are trochaic, like the cases in Rule No. 7. If there were someone named "Beh *Tee*," for example, this tapping rule wouldn't work! In such a case, the alveolar stop would instead be aspirated: [bɛˈtʰi]. Some speakers of North American English may produce medial /d/ as more of a voiced stop than a tap, thus pronouncing "daddy" as [ˈdædi].

Rule No. 9: Nasals Becoming Syllabic

This rule states that nasals become syllabic at the end of a word and after an *obstruent* (such as fricatives, stops, and affricates). In broad transcription, words ending with (spelled) "-en" and "-em" are represented using the IPA symbols /ən/ and /əm/. However, broad transcription doesn't capture all the possibilities for these sounds. The diacritic for this rule is a small vertical line placed under the nasal consonant [ˌ].

For instance, in the word "button," you usually don't include much [ə] vowel quality in the final syllable. Instead, you make a nasal release by lowering the soft palate, rather than the tongue, which results in a pure "n" that stands by itself as a syllable. Here are the broad narrow transcriptions for "button."

Broad: /ˈbʌtən/

Narrow: [ˈbʌʔn̩]

Try just holding out a long "n". You can do this without any [ə] vowel quality. To transcribe a syllabic nasal "n," you place a small vertical line under the character, like this: [n̩]. You transcribe syllabic "m" like this: [m̩].

Here are some examples in a GAE accent, narrowly transcribed:

written [ˈɹɪʔn̩]

bottom [ˈbɑɾm̩]

Rule No. 10: Liquids Become Syllabic

This rule is similar to Rule No. 9; however, it applies to sounds that are typically spelled with “-er” and “-el”. In certain environments, sounds that are broadly transcribed /ɚ/ or /əl/ are in fact produced syllabically, [ɹ̩] and [l̩]. This rule has the same diacritic as Rule No. 9, a combining small vertical bar under the consonant [̩].

The following examples compare broad and narrow transcriptions for words containing liquid consonants (/ɹ/ and /l/):

Word Example	***Broad IPA***	***Narrow IPA***
couple	/ˈkʌpəl/	[ˈkʰʌpl̩]
writer	/ˈɹaɪtɚ/	[ˈɹaɪɾɹ̩]

The word “couple” has a lateral release of the plosive. Say the word and pay attention to the final syllable; you’ll probably find not much [ə] vowel quality.

The case for (spelled) “-er” is more ambiguous: Some phoneticians use syllabic “r” ([ɹ̩]) in narrow transcription for words like “writer.” Others point out that syllabic “r” is equivalent to [ɚ] in most cases, and tend to use this syllabic diacritic less. I use syllabic “r” in narrow transcription, following Rule No. 8.

In these words, like the nasal examples in Rule No. 9, the syllabified liquids occur in the unstressed syllable of trochaic (*loud*-soft) word patterns.

Rule No. 11: Alveolars Become Dentalized

before Dentals

This is an *assimilation* rule, where one sound becomes more like its neighbor. The main influencing sounds are the interdentals /θ/ and /ð/, which can influence a number of alveolars (/n/, /l/, /t/, /d/, /s/, /ɹ/, and /z/). The dental fricatives in English (/θ/ and /ð/) are also called *interdentals* because they involve airflow between the upper and lower teeth.

The diacritic associated with Rule No. 11 is a small square bracket, that looks like a staple, placed under the consonant: [̪].

When an alveolar consonant is produced before a *dental* (sound produced against the teeth), the alveolar is produced more forward than usual. This is called being *dentalized* because the affected sound is now made closer to the teeth.

Try these minimal pair examples, paying attention to where your tongue tip is at the end of each alveolar (italicized) sound.

ten [tʰɛn] te*n*th [tʰɛn̪θ]

fill [fɪɫ] fi*l*th [fɪɫ̪θ]

no*r*[ˈnɔɹ] north [ˈnɔɹ̪θ]

Rule No. 12: Laterals Become Velarized

This rule refers to the English lateral (“l” consonant) becoming *dark* (velarized) in certain environments, otherwise remaining *light* (clear, or alveolar). Specifically, laterals become velarized after a vowel and before a consonant or at the end of a word.

If you sing *la-la-la,* you can remember that light (or clear) "l" comes at the beginning of syllables, while dark "l" is at the end. Another way of distinguishing dark from light "l" is to use your ear: The sound of "l" at the end of the word "little" (syllable final) sounds much lower than the "l" at the beginning (syllable-initial).

The diacritic used to denote velarization is a tilde placed in the middle of an IPA character. For instance, velarized "l" is written as [ɫ]. A couple examples of this rule are

waffle [ˈwɑfɫ̩]

silk [sɪɫk]

Rule No. 13: Vowels Become Nasalized before Nasals

If you happen to be a speaker of Portuguese, you'll have fairly precise control of nasality in vowels because this serves meaning in your language. This is because nasality is phonemic in Portuguese; it matters to the listener. However, in English nasality spreads from a consonant onto the vowel in front of it. As such, there is much variation from talker to talker: Some people partially nasalize the vowel and others nasalize it entirely. The amount doesn't matter that much to the listener.

The diacritic for nasalization is a tilde placed over a vowel symbol, [̃]. Some examples of this rule are

seem [sĩm]

soon [sũn]

As a transcriber, your job on this one is easy. Every time you see a vowel in front of a nasal, that vowel is nasalized. This isn't indicated in broad transcription, but is usually marked in narrow. Table 9-2 gives you some examples.

Table 9-2 Examples of Nasalized Vowels

Example Word	*Broad*	*Narrow*
banana	/bəˈnænə/	[bə̃ˈnæ̃nə]
incomplete	/ɪnkəmˈplit/	[ɪ̃nˈkə̃m˺ˈpʰl̥iʔt]
camping	/ˈkæmpɪŋ/	[ˈkʰæ̃m˺pɪ̃ŋ]

In addition to noting how the nasality rule (Rule No. 13) operated on these words, can you also see how a consonant glottalizing rule (Rule No. 5) and a stop release rule (Rule No. 4) applied? How about aspiration (Rule No. 1) and approximant partial devoicing (Rule No. 3)?

Applying the Rules

It's one thing to know these rules in this chapter; it's another to apply them. Beginning transcribers sometimes have trouble using the rules of English phonology to complete narrow transcriptions. In this section, I show you the most common errors made and provide a quiz to get you started on the right track.

Are you a phonological rule over-applier, under-applier, or "just right"? Take this simple test. Some people are phonological *rule under appliers*. Due to extreme caution (or perhaps just due to confusion) these folks tend to not apply rules where needed. In contrast, others take a *What the heck!* approach and plaster rules all over the place, even when such rules could not conceivable apply. For instance, aspiration may be placed on fricatives, nasalization over stops, and so forth.

Table 9-3 shows some examples of these three types of transcribers. Look to see where you fall.

Table 9-3 Three Degrees of Transcribers

Example Word		*Under-applied*	*Over-applied*	*"Just right"*
pants	/pænts/	[pænts]	[pʰæ̃n˺ʔs̩]	[pʰæ̃n˺ts]
pack rats	/ˈpækɹæts/	/pækɹæts/	[ˈpʰækʰɹætʰs]	[ˈpʰækɹæts]

In "pants" (American English accent), the syllable-initial /p/ would ordinarily be aspirated and the nasal vowel would be nasalized (as shown in the "just right" column). Here, an under-applier might note nothing, while the over-applier throws in a gratuitous syllabic symbol under the [s], which would make "pants" a two-syllable word.

In "pack rats," the under-applier again misses all rules. In this case, stress assignment is also missed. The over-applier liberally sprinkles aspirations everywhere, even when they don't apply. Just because voiceless stops can be aspirated doesn't mean they *are* (the rule notes this occurs only in syllable initial position).

TIP To avoid such boo-boos, remember that you don't need to use all diacritics all the time. Only use them if they're absolutely needed. Every diacritic counts. Phoneticians are picky. Ready for a quick quiz?

Which of the following narrow transcriptions would apply to the following broad transcription of "crunch," (/kɹʌnʧ/) as produced by someone from North America?

a. [kʰə̃ɹ̥ʌ̃ʔn˺ʧ]

b. [kʰɹ̥ʌ̃nʧ]

c. [kɹʌnʧ]

d. All of the above are correct

The correct answer is *b.*

If you answered *a,* you over-applied the rules. If you answered *c*, you under-applied. Answer *d* is incorrect, because *a* and *c* are highly unlikely narrow transcriptions of /kɹʌntʃ/.

Chapter 10

Grasping the Melody of Language

In This Chapter

- Using juncture for different speaking styles and rates
- Exploring the syllable and stress assignment
- Patching with sonority and prominence measures

Transcribing is more than just getting the vowels and consonants down on paper. You need that extra zest! For instance, you should be able to describe how phonemes and syllables join together, a property called *juncture*. A phonetician must be able to hear and describe the melody of language, focusing on patterns meaningful for language. This important sound aspect, called *prosody,* gives speech its zing and is described with a number of specialized terms. This chapter gives you the tools to handle bigger chunks of language, so that you can master description of the melody of language.

Joining Words with Juncture

Unless you're a lifeless android (or have simply had a very bad night), you probably don't say things such as "Hel-lo-how-are-you-to-day?" That is, people don't often speak one word (or syllable) at a time. Instead, speech sounds naturally flow together. *Juncture* is the degree to which words and syllables are connected in a language. These sections explain some characteristics of juncture and help you transcribe it.

Knowing what affects juncture

A number of factors can affect juncture, including the following:

- **Some factors are language-specific.** Some languages (such as Hawaiian) break things up and have relatively little carryover between syllables, while other languages (such as French) allow sounds to be run together. In French, the process of sounds blending into each other is called *liaison,* in which sounds change across word boundaries. Check out these two examples:

Language	***Spelling***	***IPA***	***Translation***
Hawaiian	humu humu nuku nuku apu a u a	/hu.muˌhu.muˌnu.kuˌnu.kuˌaː.puˈa.ʔa/	Small reef triggerfish
French	les amis	/le‿zaˈmi/	the friends

In these examples, the syllables of Hawaiian have little effect on each other, whereas the French has *resyllabification* (the shift of a syllable boundary) and a voicing of an underlying /s/ sound — a clear example of adjacent sounds affecting each other.

- **Other factors are more personal.** They include speaking formality and rate. Think about how your speech changes when you formally address a group versus talking casually with your friends. In a formal setting, you usually use more polite forms of address (*sir* and *madam*), fancier terms for things (*restroom* or *public convenience* instead of *john* or *loo*), and frillier sentence constructions (*Would you kindly pass the hors d'oeuvre please?* instead of *Yo. The cheese, please?*).

  In informal speech, talkers usually have less precise boundaries than in formal speech. This register change often interacts with *rate,* because rapid speech often causes people to *undershoot* articulatory positions (not reach full articulatory positions). The result can be *vowel centralization* (sounds taking on more of an [ə]-like quality), *de-diphthongization* (diphthongs becoming monophthongs), changes in *consonant quality* (such as the tongue moving less completely to make speech sounds), and changes in *juncture boundaries* (including one boundary shifting into another).

Check out these examples from American and British English:

Language	*Example*	*Formal (Slow Rate)*	*Informal (Fast Rate)*
American English	How are you?	/haʊ ˈɑɹ ju/	[həˈwɑɹjə]
British Standard English	Nice, isn't it?	/naɪs ˈɪznt ɪt/	[ˈnaɪsɪnʔɪ]

REMEMBER Changes in register and style clearly affect juncture (how speech sounds are connected in terms of pauses or gaps). Some phoneticians refer to juncture as *oral punctuation* because it acts somewhat like the commas and periods in written language.

Transcribing juncture

You can transcribe juncture in a couple different ways. They are as follows:

- **Close juncture:** This default way of transcribing shows that sounds are close together by placing IPA symbols close together in transcription from phoneme to phoneme. An example is "Have a nice day!" /hævə naɪs ˈdeɪ/.
- **Open juncture:** You use *open juncture* (also referred to as *plus juncture*) symbols when you need to emphasize gaps separating sounds. Consider these two expressions:

  "Have a nice day!" /ˈhævə + naɪs ˈdeɪ/

  "Have an ice day!" /ˈhævən + aɪs deɪ/

  Many speakers would probably produce this second example ("Have an ice day") with a glottal stop before the vowel of ice, as a way of marking the gap between the words "an" and "ice." To distinguish these two expressions, the exact placement of the gap between the

/ə/ and /n/ is critical. Therefore, open juncture symbols are helpful.

Phoneticians use different conventions for juncture between words. Depending on the speaking style, some phoneticians place a content word (such as the verb "have" in the preceding examples) next to an adjacent function word (such as the determiner "a"), resulting in /ˈhævə/. Doing so tells the reader there is no pause between these sounds. Other transcribers indicate such juncture with a tie-bar at the bottom of the two words: (/ˈhæv‿ə/).

The flow of spoken language doesn't necessarily follow the grammatical patterns you learned in English class. Talkers can run-on or hesitate during speech for many reasons. Consider the sentence, "I went to the store." This sentence can be produced with many different juncture patterns, such as

I . . . went to the store.

I went . . . to the store.

I went to . . . the store.

I went to the . . . store.

And so on. You get the idea. Transcribing all the potential variations in the exact same way wouldn't make sense. What's important is showing where all the gaps take place. Many phoneticians use the IPA pipe symbol ([|]), which technically indicates a *minor foot,* a prosodic unit that acts like a comma (I describe it in greater detail in Chapter 11). However, many transcribers also use this symbol to represent a short pause, whereas they use a double bar ([‖]) to represent a long pause, such as at the end of a sentence. Here are some examples:

/aɪ |wɛn tə ðə stɔɹ‖/

/aɪ wɛnt | tə ðə stɔɹ‖/

If you use these symbols in this manner, be sure to indicate it in notes to your transcription. A good general principle to follow is to employ juncture and timing information only when needed. For instance, the hash mark (#) is a linguistic symbol that means a boundary, such as the end of a word. I have seen older phonetic transcriptions with a hash mark placed between every word. These ended up looking as if a psychotic chicken used the transcription to practice the Rhumba. Keep your transcriptions tailored to your needs, with just the amount of detail your applications require.

Emphasizing Your Syllables

A syllable is something everyone knows intuitively, but can drive phoneticians nuts trying to pin down precisely. By definition, a *syllable* is a unit of spoken language consisting of a single uninterrupted sound formed by a vowel, diphthong, or syllabic consonant, with other sounds preceding or following it. Phoneticians don't see the definition so cut and dry.

Phoneticians consider a *syllable* an essential unit of speech production. It's a unit with a center having a louder portion (made with more air flow) and optional ends having quieter portions (made with less air flow). Phoneticians agree on descriptive components of an English syllable, as shown in Figure 10-1.

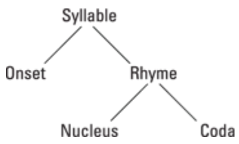

Illustration by Wiley, Composition Services Graphics

Figure 10-1: Parts of an English syllable.

From Figure 10-1, you can see that an English syllable (often represented by the symbol *sigma* [σ]), consists of an optional *onset* (beginning) and a *rhyme* (main part). The rhyming part consists of the vowel and any consonants that come after it. The vowels in a rhyme sound alike. At a finer level of description, the rhyme is divided into the *nucleus* (the vowel part) and the *coda* (tail or end) where the final consonants are. From this figure, you can take a word like "cat" and identify the different parts of the syllable. For "cat" (/kæt/), the /k/ is the onset, /æ/ is the nucleus, and the /t/ is the coda.

This is why this type of poem rhymes:

Roses are red, violets are bl*ue*. . . .

blah blah blah *blah,* blah *blah* blah blah . . . *you*.

Languages vary considerably with which kinds of onsets and codas are allowed. Table 10-1 shows some samples of syllable types permissible for English.

Table 10-1 Sample Syllable Types in English

Example	*IPA*	*Syllable Type*
eye	/aɪ/	V
hi	/haɪ/	CV

height	/haɪt/	CVC
slight	/slaɪt/	CCVC
sliced	/slaɪst/	CCVCC
sprints	/spɹɪnts/	CCCVCCC

The last column lists a common abbreviation for each syllable type, where “C” represents a consonant and “V” represents a vowel or diphthong. For instance, “eye” is a single diphthong and thus has the syllable structure “V.” At the bottom of the table, “sprints” consists of a vowel preceded and followed by three consonants, having the structure “CCCVCCC.”

Strings of consonants next to each other are called *consonant clusters* (or *blends*). Each language has its own rules for consonant cluster formation. The permissible types of consonants clusters in English are, well, rather odd. Figure 10-2 shows some of the English initial consonant clusters in a chart created by the famous Danish linguist, Eli Fisher-Jørgensen.

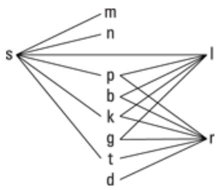

Illustration by Wiley, Composition Services Graphics

Figure 10-2: Some English syllable-initial consonant clusters.

Notice the *phonotactic* (permissible sound combination) constraints at work in Figure 10-2. It’s possible to have *sm*- and *sn*- word beginnings, but not *sd*-, *sb*-, or *sg*-. There can be an *spl*- cluster, but not a *ps*- or *psl*- cluster.

Stressing Stress

Nothing makes a person stand out as a foreign speaker more than placing stress on the wrong syl*lab*le. In order to effectively teach English as a second language, transcribe patient notes for speech language pathology purposes, or work with foreign accent reduction, you need to know how and where English stress is assigned. This, in turn, requires an understanding of phonetic stress at the physiologic and acoustic levels.

Stress is a property of English that's signaled by a syllable being louder, longer, and higher than its neighbors. It's a *suprasegmental* property (which means that it extends beyond the individual consonant or vowel). *Louder, longer and higher* are perceptual properties, that is, in the ear of the beholder. For a syllable to be perceived as stressed, physical attributes must be physically changed. For now, this table describes what a talker does to produce each of these speech properties (*articulatory*), what the acoustic property is called (*acoustic change*), and how it's heard (*perceptual impression*). Check out Chapter 12 for more in-depth information.

To understand Table 10-2 and get a sense of how louder, longer, and higher works, say a polysyllabic word correctly and then say it incorrectly. Say "*syl*lable" correctly, with stress on the initial syllable. Next, incorrectly place the stress on the second to last syllable (also called the *penultimate*, or *penult*), as in "syl*la*ble." Finally, place stress on the final syllable, or ultima, "sylla*ble*."

Table 10-2 Physical, Acoustic, and Perceptual Markers of Stress in English

Articulatory	*Acoustic Change*	*Perceptual Impression*
Increased airflow, greater intensity of vocal fold vibration	The amplitude increases	Louder
Increased duration of vocal and consonantal gestures	The duration increases	Longer sound ("length")
Higher rate of vocal fold vibration	The fundamental frequency increases	Higher pitch

In each case (whether you're correctly or incorrectly pronouncing it), the stressed syllable should sound as if someone cranked up the volume. The following sections tell you more about how stress operates at the word, phrase, and sentence level in English.

Eyeing the predictable cases

Stress serves four important roles in English. They are as follows:

- **Lexical (word level):** When you learn an English word, you learn its stress. This is because stress plays a *lexical* (word specific) role in English: it's assigned as part of the English vocabulary. For example, syllable is pronounced /ˈsɪlebəl/, not /sɪˈlʌbəl/ or /sɪləˈbʌl/.
- **Noun/verb pairs:** In English, stress also describes different functions of words. Try saying these noun-verb pairs, and listen how stress alteration makes a difference (the stressed syllables are italicized):

Spelling	*Part of Speech*	*IPA*
(to) re*cord*	Verb	[ɹəˈkʰɔɹd]

(a) *rec*ord	Noun	[ˈɹɛkɚd]
(to) re*bel*	Verb	[ɹəˈbɛɫ]
(a) *re*bel	Noun	[ˈɹɛbɫ̩]

These stress contrasts are common in stress-timed languages, such as English and Dutch (whereas tone languages, such as Vietnamese, may distinguish word meaning by contrasts in pitch level or pitch contour on a given syllable).

- **Compounding:** With compounding, two or more words come together to form a new meaning, and more stress is given to the first than the second. For example, the words "black" and "board" create "blackboard" /ˈblækbɔɹd/.

  Also, the juncture is closer than a corresponding adjective + noun construction. For example, if you pronounce the following pairs, you'll notice a longer pause between the words in the first example (the English column) than between the words in the second example (the IPA column).

Grammatical Role	***English***	***IPA***
Adjective + noun	a black board	/ə blæk ˈbɔɹd/
Compound noun	a blackboard	/ə ˈblækbɔɹd/

- **Emphasis in phrases and sentences:** Also known as *focus,* this is a pointer-like function that draws attention to a part of a phrase or sentence. By making a certain syllable's stress louder, longer, and higher, the talker subtly changes the meaning. It's as if the utterance answers a different question. For example:

Dylan sings better than Caruso. (Who sings better than Caruso?)

Dylan *sings* better than Caruso. (What does Dylan do better than Caruso?)

Dylan sings better than *Caruso*. (Who does Dylan sing better than?)

People handle this kind of subtlety every day without much problem. However, just think how difficult it is to get computers to understand this type of complexity.

Identifying the shifty cases

For the most part, English stress remains fairly consistent. However, some cases realign and readjust. You may think of it as a musical score having to be switched around here and there to keep with the rhythm. These adjustments, called *stress-shift*, are a quirky part of English phonology.

REMEMBER Stress realigns itself in a manner to preserve the up-and-down (rhythmic) patterns of English. If syllables happen to combine such that two stressed syllables butt up against each other, one flips away so that there is some breathing room. Think of it like two magnets with positive and negative ends: put two positives together and one flips around so that it's *positive/negative/positive/negative* again.

TRY IT Some English words take primary stress on different syllables, based on the context. For example, you can pronounce the word "clarinet" with initial stress, such as /ˈklɛɹɪnɛt/ or with final stress, as in /klɛɹɪˈnɛt/, depending on the stress of the word that comes next. Try this test:

1. Say "Clari*net mu*sic" three times.

Doing so sounds a bit awkward, right? It should have been more difficult because two stressed syllables had to butt up against each other.

2. Say "*Clar*inet *mu*sic" three times.

You should notice that this second pattern flows more naturally because it permits the usual English stress patterns (strong/weak/strong/weak) to persist.

Sticking to the Rhythm

Another way an English speaker can show adeptness with the language is having the ability to use English *sentence rhythm patterns*, where greater stresses occur at rhythmic intervals, depending on talking speed. To get a sense of these layered rhythms, consider these initially stressed polysyllabic words: "really," "loony," "poodle," "swallowed," "fifty," "plastic," and "noodles."

When you put them together in a sentence, they form:

The really loony poodle swallowed fifty plastic noodles.

Although speaking this sentence is possible in many fashions, a typical way people produce it is something like this:

The *rea*lly loony *poo*dle swallowed *fif*ty plastic *noo*dles.

That is, regularly spaced, strongly stressed syllables (italicized) are interspersed with words that still retain their primary stress (such as "loony"), yet they're relatively deemphasized in sentential context. This kind of timing is rhythmic and can reach high levels in art forms like vocal jazz (or perhaps, rap). Chapter 11 discusses ways you can transcribe this kind of information.

Tuning Up with Intonation

In phonetics, *sentence-level intonation* refers to the melodic patterns over a phrase or sentence that can change meaning. For instance, rising or falling melodic patterns that change a statement to a question, or vice-versa. Intonation is quite different from *tone*, which is the phoneme-level pitch differences that affect word meaning in languages such as Mandarin, Hausa, or Vietnamese (see Chapter 18). English really has no tone. The following sections take a closer look at the three patterns of sentence-level intonation that you find in English.

Making simple declaratives

A basic pattern of English intonation is the simple *declarative* sentence, which is a statement used to convey information. A couple examples are "The sky is blue" or "I have a red pencil box."

> Think of this pattern as the plain gray sweater of the phonetic wardrobe. A bit dull, perhaps, but it's necessary. When you're simply stating something, the chances are your intonation is falling. That is, you start high and end low.

Falling intonation seems to be a universal pattern, perhaps due to the fact that it takes energy to sustain the thoracic pressure needed to keep the voice box (larynx) buzzing. As a person talks, the air pressure drops and the amount of buzzing tends to drop, causing the perceived pitch to fall, as well.

Answering yes-no questions

The second pattern of sentences is called the "yes/no question." When you're asking a question that has a yes or no answer, you probably have rising intonation. This means you start low and end high.

> Try producing the same sentences that I introduce in the previous section, but instead of falling in pitch as you speak, have your voice rise from low to high.

You probably noticed these English statements ("The sky is blue?") have now turned into questions. Specifically, they're questions that can be answered with *yes* or *no* answers. This rising pitch pattern for questions is fairly common among the world's languages. For instance, French forms most questions in this manner. ***Note:*** Some languages don't use intonation at all to form a question. For instance, Japanese forms questions by simply sticking the particle /ka/ at the end of a sentence.

Focusing on "Wh" questions

The third pattern of sentences include English questions with the *Wh questions*, including "who," "what," "when," "where," "why," and "how," (which are produced with falling pitch, rather than rising). Try a few, while determining whether your voice goes up or down:

Who told you that?

What did he say?

When did he tell you?

Where will they take you?

Why are you going?

How much will it cost?

REMEMBER Your intonation likely goes down over the course of these utterances. Try this for yourself. Say the preceding sentences to see whether your intonation goes down.

Showing Your Emotion in Speech

When someone talks, part of the melody serves a language purpose, and part serves an emotional purpose. When you're transcribing speech, you need to understand emotional prosody because it can interact in complex ways with the linguistic functions of prosody. In fact, people can show many emotions in speech, including joy, disgust, anger, fear, sadness, boredom, and anxiety.

Studies have shown that people speak happiness (joy) and fear at higher frequency ranges (heard as pitch) than emotions such as sadness. Anger seems to be an emotion that can go in two directions, phonetically:

- **Hot anger:** When people go up high with the voice and show much variability.
- **Cold anger:** When people are brooding with low pitch range, high intensity, and fast *attack times* (sudden rise in amplitude) at voice onset.

Emotional patterns in speech (also known as *affective prosody*) don't directly affect sentence meaning. However, these patterns can interact with linguistic prosody to affect listeners' understanding. For instance, adults with cerebral right hemisphere damage (RHD) can have difficulty understanding, producing, and mimicking the emotional components of speech. The speech of such individuals can often be *monotonic* (flat). It can sometimes be challenging for clinicians to sort out which aspects of these speech presentations are due to emotion or to linguistic deficits.

Fine-Tuning Speech Melodies

Phoneticians can be sticklers for detail. They just don't like messy bits left over. In addition to the different types of stress, intonation, focus, and emotional prosody, certain aspects of speech melody still require measures to account for them. These sections examine two such measures.

Sonority: A general measure of sound

Sono- means sounds, and *sonority* is therefore a measure of the relative amount of sound something has. Technically, *sonority* refers to a sound's loudness relative to those of other sounds having the same length, stress, and pitch. This measure of sound is particularly handy for working with tone languages, such as Vietnamese, where decisions about tone structure are important.

To get a clearer sense of this jargon, try saying the sound "a" (/ɑ/) followed by the sound "t." Assuming you spoke them at the same rate and loudness, the vowel /ɑ/ should be much more sonorous (have more sound) than the voiceless stop, "t."

The concept of sonority is relative, which means phoneticians often refer to sonority hierarchies or scales. In a *sonority hierarchy*, classes of sounds are grouped by their degree of relative loudness. Check out www-01.sil.org/linguistics/GlossaryOfLinguisticTerms/WhatIsTheSonorityScale.htm for an example of one.

A sonority scale expresses more fine-grained details. For instance, according to phonologist Elizabeth Selkirk, English sounds show the following ranking:

([ɑ] > [e=o] > [i=u] > [r] > [l] > [m=n] > [z=v=ð] > [s=f=θ] > [b=d=g] > [p=t=k])

If you try out some points on this scale, you'll hear, for example, that [ɑ] is more sonorous than [i] and [u].

Sonority is an important principle regulating many phonological processes in language, including *phonotactics* (permissible combinations of phonemes) syllable structure, and stress assignment.

Prominence: Sticking out in unexpected ways

When all is said and done, some problem cases of prosody can still challenge phoneticians. One such problem is exactly how stress is assigned to syllables in words. For instance, some English words can be produced with different amounts of syllables. Consider the words "frightening" and "maddening."

Do you say them with two syllables, such as /ˈfɹaɪtnɪŋ/ and /ˈmædnɪŋ/? Or do you use three syllables, such as /ˈfɹaɪtənɪŋ/ and /ˈmædənɪŋ/? Or sometimes with two and sometimes with three?

Other English words may change meaning based on whether they are pronounced with two or three syllables. For instance:

"lightning" (such as in a storm) /ˈlaɪtnɪŋ/

"lightening" (such as, getting brighter) /ˈlaɪtənɪŋ/

A proposed solution for the more difficult cases of stress patterns is to rely on a feature called *prominence*, consisting of a combination of sonority, length, stress, and pitch. According to this view, prominence peaks are heard in words to define syllables, not solely sonority values.

Prominence remains a rather complex and controversial notion. It's an important concept in *metrical phonology* (a theory concerned with organizing segments into groups of relative prominence), where it's often supported with data from speech experiments. However, other phoneticians have suggested different approaches may be more beneficial in addressing the problems of syllabicity in English (such as the application of speech technology algorithms, rather than linguistic descriptions).

Chapter 11

Marking Melody in Your Transcription

In This Chapter

- Sampling choices for prosodic transcribing
- Defining the tonic syllable and intonational phrase
- Becoming proficient at a three-step process
- Rising and tagging

Imagine you're sitting in a busy restaurant in a big city hearing many different foreign languages spoken. It's noisy, but you want to impress your friends with your (amazing) ability to tell which language is which. One important clue to help you is *language melody,* which includes *stress* (when a syllable is louder, longer, and higher because the talker uses extra breath) and *intonation* (a changing tune during a phrase or sentence).

For instance, someone speaking Spanish has a very different melody than someone speaking Mandarin, and you can hear it if you know what you're listening for. However, capturing these details in written transcription is much more difficult, particularly if you need to compare healthy and disordered speech.

In this chapter, I show you some practical ways to incorporate melodic detail in your transcriptions. I begin with a tried-and-true method useful for clinical notes or field transcription. I also include some examples of a more systemized method, the Tone and Break Indices (ToBI) that linguists and many people in the speech science community use.

Focusing on Stress

When transcribing many languages, being able to identify a stressed syllable is essential. Knowing these characteristics of a stressed syllable can help you identify it. An English *stressed syllable* is louder, longer (in duration), and higher in pitch. In English, stress plays a number of important roles:

- At the vocabulary level, *polysyllabic* words (with more than one syllable) have specified stress that a native speaker must correctly produce to sound appropriate. Thus, "*syl*lable" is okay, but "syl*la*ble" sounds weird.
- For word function, stress makes a difference between nouns such as "*re*bel" and verbs, such as "to re*bel*."
- In phrases and sentences, stress changes *focus*, or emphasis. For example, although these two sentences contain the exact same words, stressing different words gives a different emphasis:

  "*She* never wears Spandex!" (He does!)

  "She never *wears* Spandex!" (She sells it, instead.)

Stress also plays a special role in English when it serves as the *tonic syllable* (a syllable that stands out because it carries the major pitch change of a phrase or sentence).

The following sections describe some of the complexities involved in speech that can make the job of transcribing language melody a challenge.

Recognizing factors that make connected speech hard to transcribe

Understanding the role that stress plays in English is important, and a futher challenge is to be able to accurately complete a prosodic transcription of connected speech.

Transcribing *prosody* (the melody of language) can be challenging, for a number of reasons:

- Several types of prosodic information are present in a person's speech. This information includes *linguistic prosody* where melody and timing specifically affect language, as well as *emotional prosody*, reflecting the speaker's mood and attitude toward what the speaker is discussing.
- People don't usually speak in complete sentences. Nor do they always cleanly break at word or phrase boundaries. For example, here is some everyday talk from teenagers in Dallas, Texas: "So, like, I was gonna see this movie at North Park? But then Alex was there? So . . . yeah, and . . . then it's like . . . Awkward!" (This is an example of the Valley Girl social dialect; refer to Chapter 18 for more information.)

TRY IT Listen to yourself talk sometime, and tune in to the grammatical structures that you use and the precision with which you articulate. Speaking in different *registers* (level of language used for a particular setting) in different settings is natural. When presenting professionally (such as in class, work, or clinic), people are usually on their best behavior and tend to use complete sentences, full grammatical constructions, and more fully achieved articulatory targets (called *hyperspeech*). In contrast, when talking casually with friends, people naturally relax and use more informal constructions, centralized vowels, and reduced articulatory precision, referred to as *hypospeech*.

Finding intonational phrases

The IPA doesn't recommend any one system for capturing language melody (prosody). Instead, various phoneticians have applied rules and theories in the best ways they see fit. Fortunately, many methods are available. One time-honored method begins with defining an intonational phrase. Based on these building blocks, you, as a transcriber, can achieve different degrees of success.

An *intonational phrase,* sometimes called a *tone unit, tonic phrase,* or *tone group,* is a pattern of pitch changes that matches up in a meaningful way with a part of a sentence. Although the exact definition varies between phoneticians, certain key characteristics of an intonational phrase are as follows:

- A part of connected speech containing one tonic syllable.
- Similar to a *breath-group* (sequence of sounds spoken in a single exhalation), a single, continuous airstream supports it.
- Similar to a phrase, a clause, or a non-complex sentence.

- Similar to breaks signaled by written punctuation (commas, periods, or dashes.)
- Intonational phrases aren't syntactic units, but they can frequently match up to them in a practical sense.

Check out these examples:

Example Words	***Number of Intonational Phrases***
|ˈYep!|	1
|The ˈdog.|	1
|Although he ignored the ˈcat,|the boy fed the ˈdog.|	2
|The boy fed the ˈdog,|but ignored the ˈcat.|	2
|The boy fed the ˈdog, | gave it a ˈmeatball,|but ignored the ˈcat.|	3

In these examples, the boundaries of intonational phrases are divided using vertical lines ([|]). The words that typically receive stress have a primary stress mark ([ˈ]) before them. Single words (such as "Yep") and fragments ("The dog") can be intonational phrases. If the speaker is communicating too much in a single breath-group, the utterance is often broken down into separate, shorter tone units (such as phrases, clauses, or shorter bits of choppier speech) containing between one to three intonational phrases (shown here). It's common for a spoken sentence to have one to two intonational phrases, but there could be more, depending on how a person is talking.

Zeroing in on the tonic syllable

Each intonational phrase will have one (and only one) tonic syllable (also called the *nuclear syllable*), the syllable that carries the most pitch change. The tonic syllable is an important idea for many theories of prosody.

A *tonic syllable* is the key part of an intonational phrase because it's a starting point for the melody of that phrase. Together, the concepts of a tonic syllable and intonational phrase allow a thorough description of language melody. If this theory sounds circular to you, it is. However, it's by intention. Here is how it works: An intonational phrase consists of a nucleus and an optional pre-head, head, and tail. The following figure shows an example for an intonational phrase.

Taken separately or in combination these components can describe English melody. This tonic syllable/intonational phrase system is often used for teaching students of English as a second language. It's particularly well suited for British English, especially the Received Pronunciation (RP) accent.

Seeing how phoneticians have reached these conclusions

Phoneticians have come up with these explanations of English melody by considering several factors, including the rhythm of English (called *meter*, described in units called *feet*). Phoneticians also note that intonation corresponds with different types of meaning, such as statements and questions (refer to the next section for more information).

It's beyond the scope of introductory phonetics to explain these theories of prosody. However, you should be able to form an intuitive sense of what an intonational phrase is. For instance, if you review the examples in the previous section, "Finding intonational phrases," you can hear that "The dog" receives a lot of stress, whereas other parts of the sentences (such as "although") don't receive much stress.

Consider the sentence "The boy fed the dog." You can pronounce this sentence in many different ways, depending on what you're emphasizing (for example, "The *boy* fed the dog," "The boy *fed* the dog," and so on). In these cases (which some phoneticians call a *dislocated tonic*), emphasis or focus has shifted the position of the tonic syllable. However, in most cases, the tonic syllable of an intonational phrase is the last stressed syllable that conveys new information, such as:

> "The boy fed the *dog*" /ðə bɔɪ fɛd ðə ˈdɔg/

Here, "dog" is the tonic syllable, carrying the most prosodic information.

Sometimes a person doesn't produce intonational phrases in the usual manner. In actual transcriptions, you may encounter speech like this:

> "The boy . . . fed . . . the dog." /ðə ˈbɔɪ|ˈfɛd | ðə ˈdɔg/

This type of speech (*hesitant speech*) would have more intonational phrases. This particular example uses three intonational phrases instead of one. The tonic syllables would be "boy," "fed," and "dog."

Consider cases where stress is changed due to emphasis. If someone is excited about the fact that the dog was *fed*, rather than *washed*, he or she would probably say the following:

"The boy *fed* the dog" /ðə bɔɪ ˈfɛd ðə dɔg/

This time, "fed" is the tonic syllable of a single intonational phrase.

Applying Intonational Phrase Analysis to Your Transcriptions

Being able to apply intonational phrase analysis can give you a better idea of how phoneticians handle the challenge of transcribing intonation and prosody. This method demonstrates an accurate and easy-to-complete method of prosodic transcription. Although this method has its limitations of describing prosody because it doesn't provide fine-grained details such as the subcomponents of a tonic phrase (pre-head, head, nucleus, and tail), together with *narrow transcription* (recording details about phonetic variations and allophones), it does provide an easy way of denoting the melody of connected speech.

Here I walk you through these three steps and use this example to explain this process:

"The earliest phoneticians were the Indian grammarians."

/ðɪ ɝlɪɛst fɔnətɪʃənz wɚðə ɪndɪən ɡɹəmɛɹɪənz/ (The broad transcription, with no details yet filled in.)

If you want to listen to the sound file, check it out at www.utdallas.edu/~wkatz/PFD/the_earliest_phonetician_WK.wav. This is a recording of me reading a passage in a matter-of-fact manner.

1. **Locate prosodic breaks corresponding with the breath groups.**

To find them, listen for clear gaps during speech. After you locate them, place a vertical bar ([|]) for minor phrase breaks and a double-bar ([‖]) for major phrase-breaks.

For this example, your work should look like this:

/ðɪ ɝlɪɛst fɔnətɪʃənz | wɚðə ɪndɪən ɡɹəmɛɹɪənz‖/

2. **Mark the tonic syllable in each tone unit (intonational phrase) as the primary stressed syllable and denote the stress in other polysyllabic words by marking them with secondary stress.**

Mark the stress of "ˈ*ear*liest" and "graˈ*mmar*ian" with a primary stress mark. In this case, the stress mark further indicates the tonic syllable of an intonational phrase.

At this stage, your transcription should look like this:

/ðɪ ˈɝlɪɛst fɔnətɪʃənz |wɚðə ɪndɪən ɡɹəˈmɛɹɪənz‖/

Continue to mark stress on the other polysyllabic words ("phone*ti*cians" and "*In*dian") to produce the following:

/ðɪ ˈɝlɪɛst fɔnəˌtɪʃənz |wɚðə ˌɪndɪən ɡɹəˈmɛɹɪənz‖/

3. Draw an estimate of the fundamental frequency contour *(pitch plot)* above the IPA transcription.

This is the best part. Use your ear (and hand) to draw the shape of the intonation contour above the transcription. This task is rather like *follow the bouncing ball*. Refer to Figure 11-1 for an example.

In this figure, a hand-drawn pitch contour marks sounds going up and sounds going down. These plots are helpful for transcribing the intonational phrases of connected speech.

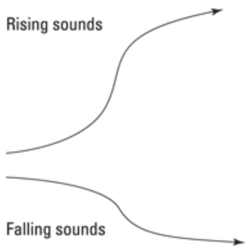

Illustration by Wiley, Composition Services Graphics

Figure 11-1: Rising sounds go up, falling sounds go down.

If you're musically or artistically challenged, you can build your confidence for intonation contour sketching in these ways:

- **Practice.** The old saying is true: Practice makes perfect.
- **Use a speech analysis program.** These types of programs, such as WaveSurfer or Praat (Dutch for "Speech"), can help you analyze the fundamental frequency patterns of the utterances you want to transcribe and compare your freehand attempts with the instrumental results. You're probably better than you think you are.

The intonation contour you arrive at should look something like this:

Don’t worry whether you’ve smoothly connected the pitch contour or whether you make less connected straight forms. The main point is that your figure rises when the pitch rises (such as during the word “earliest”) and falls appropriately (as in the end of the phrase). The goal of using this three-step method is to uncover the melody of the original utterance.

When you finish, your transcription should look like this:

/ ðɪˈɝlɪɛst fɔnəˌtɪʃənz | wɚ ðɪ ˌɪndɪən gɹəˈmɛɹɪənz||/

For more practice, go in the opposite direction. Look up www.utdallas.edu/~wkatz/PFD/the_earliest_phonetician_WK2.wav for a link to a sound file of the phrase, "The earliest phoneticians were the Indian grammarians" read by yours truly in a slightly different manner. This time, I use an impatient tone of voice. Your job is to produce a narrow transcription marked with intonational phrases, tonic syllables, and a prosodic contour (follow the three-step method to do so). Good luck!

You can check your work by going to www.utdallas.edu/~wkatz/PFD/the_earliest_phoneticians_answer.gif.

Tracing Contours: Continuation Rises and Tag Questions

Chapter 10 discusses the three main patterns for English sentence-level intonation. Two other common intonation patterns exist. They can differ slightly as a function of dialect (American versus British English) as well as the mood and attitude of the speaker. These sections take a closer look specifically at continuation rises and tag questions to show you where they occur and how to transcribe them.

Continuing phrases with a rise

A *continuation rise* is a conspicuous lack of a falling pattern on the tonic syllable at the end of that phrase. It occurs when one intonational phrase follows another. For instance, contrast the falling pattern on the tonic syllable ("crazy") in the first example sentence and the continuation rise for that same word in the second example.

- "Eileen is really *crazy*." (www.utdallas.edu/~wkatz/PFD/eileen_crazy1.wav)
- "Eileen is really crazy, but she's my best friend." (www.utdallas.edu/~wkatz/PFD/eileen_crazy2.wav)

Continuation rise patterns are common in English lists. Here is a (ridiculously healthy) shopping list: "She bought peaches, apples, and kiwis." Most North American English speakers pronounce it something like what appears in Figure 11-2.

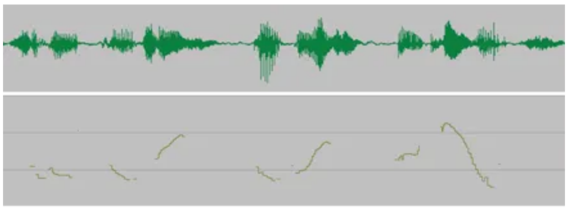

Illustration by Wiley, Composition Services Graphics

Figure 11-2: The speech waveform (above) and the intonation contour (below) of "peaches, apples, and kiwis."

In this figure, notice that the words "peaches" and "apples" rise during the continuation of the sentence, but the word "kiwis" falls at the end. If you were to flip this order and use falling prosody during the production (for "peaches" and "apples"), while rising at the end (for "kiwis"), you would sound, frankly, bizarre.

You have a talker who coughs and says "um" and "er" during the transcription. What do you do? You may wonder if you have to put those utterances in your transcription. For most narrow transcriptions, the answer is yes. Filled pauses such as [əm] and [ɚ] are common in speech. Talkers with foreign accents may use very different filled pauses than native speakers of English (for instance, [ɛm] for Hebrew speakers and [eː] or [n̩] for Japanese speakers). You can indicate nonlinguistic vocalizations (such as coughing, sneezing, laughter) in parentheses.

Tagging along

English *tag questions* (statements made into a question by adding a fragment at the end) have their own characteristic patterns. Tag questions can be either rising or falling. Their patterns depend somewhat on the dialect used (for instance, British or American), but mainly they depend on the exact use of the tag.

Rising patterns are found when a tag question turns a statement into a question, such as these examples:

- "You're kidding, *aren't you?*"

- "It's a real Rolex, *isn't it?*"

Falling patterns are used to emphasize a statement that was just made:

- "He sold you a fake Rolex, *didn't he?*"
- "That's really awful, *isn't it?*"

Testing Out ToBI

The *tone and break indices system (ToBI)* is a set of conventions used for working with speech prosody. ToBI is mainly designed for English, although conventions are being developed for other languages (including German, Korean, Japanese, and Greek). Researchers primarily use this system, and it's a bit more advanced for most clinical and educational settings. However, having a basic grasp of ToBI is important because you'll probably encounter some literature referring to it.

Here are some characteristics concerning ToBI that are important to grasp:

- Instead of drawing a pitch contour over a transcription, ToBI describes pitch peaks and valleys in terms of high and low *target tones,* which are represented as combinations of the letters H (for high) and L (for low). A target tone can be simple, such as *high* (written as H*, called *H star*), or it can have gliding properties, such as L+H*. In this case, the tone starts low (L), and then glides up to high (H*).

  For example, if someone asked "Well?", the target tone would be L+H*.
- Target tones are typically written on a line of text (called a *tier*). A second tier represents break indices, showing pause and gap durations. ToBI uses a range of boundary strength levels, from 0 to 4, representing shortest to longest. A break of 0 represents no break (such as in the contraction between we and are –– *we're*), a 1 represents most breaks between words,

and breaks of 3 and 4 indicate intentional breaks in a phrase and at a sentence ending.

- In ToBI, the last pitch accent of a phrase is called the *nuclear pitch accent*, which is similar to the idea of a tonic syllable. For instance, in a straightforward reading of "The boy fed the dog," the nuclear pitch accent is the word "dog."
- ToBI also allows each phrase to receive another marking after nuclear pitch, called a *phrase accent*. These markers (L- and H-) permit additional prosodic refinement.
- A boundary tone (L% or H%) acts as a kind of marker for sentence-level intonation. When the sentence has an overall falling intonation, as in a simple declarative pattern, the boundary tone is L%. When pitch rises, as in a yes-no question or continuation rise, the sentence has an H% boundary tone. Boundary tones are placed at phrase edges.

Check out this example that compares tonic syllable analysis (that I discuss in this chapter) with ToBI for the sentence, "The boy fed the *dog*."

- **Tonic syllable analysis:** /|ðə bɔɪ fɛd ðə ˈdɔg‖/

  Here, the word "dog" is the tonic syllable of an intonational phrase.

- **ToBI analysis:**

  Break index [1 1 1 1 4]

  Tone tier [H* H*L-L%]

  Segmental tier [ðə bɔɪ fɛd ðə ˈdɔg]

The ToBI analysis specifies level 1 breaks (those used for most middle-of-phrase boundaries) between the words of the sentence and a phrase level break (4) at the end. The tone tier shows high nuclear tones throughout, followed by a nuclear accent marked with a low phrase accent (L-) and a low boundary tone (L%), indicating phrase final fall.

The other H* marker ("boy") is an optional pre-nuclear pitch accent. In this case, it indicates that the entire utterance is produced with a high, flat intonation (with a final fall). However, more gradual pitch declination, called *downdrift*, can be indicated by adding *downstepping* symbols, !H*.

For instance, a ToBI representation of a more declining pattern could be represented like this:

Break index [1 1 1 1 4]

Tone tier [H* !H* !H*L-L%]

Segmental tier [ðə bɔɪ fɛd ðə ˈdɔg]

If you're using a program such as Praat or WaveSurfer, you can integrate ToBi labeling directly into your graphics. Refer to www.cs.columbia.edu/~agus/tobi/ for more information on incorporating ToBI into waveform editing packages.

Part III

Having a Blast: Sound, Waveforms, and Speech Movement

"My husband's uvula is shaped like a duck call. When he sneezes, a flock of mallards usually shows up."

Visit www.dummies.com/extras/phonetics for more great Dummies content online.

In this part . . .

- Comprehend what causes sound and know why this information is essential for understanding how people talk and listen.
- Grasp how to describe sound physically, in terms of frequency, amplitude, and duration.
- Be able to relate physical aspects of sound to people's subjective listening patterns.
- Know how to decode the information in sound spectrograms.
- Gather the basics of current models of human speech perception.

Chapter 12

Making Waves: An Overview of Sound

In This Chapter

- Working with sound waves
- Getting grounded in the physics needed to understand speech
- Relating sound production to your speech articulators

One of the great things about phonetics is that it's a bridge to fields like acoustics, music, and physics. To understand speech sounds, you must explore the world of sound itself, including waves, vibration, and resonance. Many phoneticians seem to be musicians (either at the professional level or as spirited amateurs), and it's normal to find phoneticians hanging around meetings of the Acoustical Society of America. Just trying to talk about this accent or that isn't good enough; if you want to practice good phonetics, you need to know something about acoustics.

This chapter introduces you to the world of sound and describes some basic math and physics needed to better understand speech. It also explains some essential concepts useful for analyzing speech with a computer.

Defining Sound

Sound refers to energy that travels through the air or another medium and can be heard when it reaches the ear. Physically, sound is a longitudinal wave (also known as a *compression wave*). Such a wave is caused when something displaces matter (like somebody's voice yelling, "Look out for that ice cream truck!") and that vibration moves back and forth through the air, causing compression and *rarefaction* (a loss of density, the opposite of compression). When this pressure pattern reaches the ear of the listener, the person will hear it.

REMEMBER When a person shouts, the longitudinal wave hitting another person's ear demonstrates compression and rarefaction. The air particles themselves don't actually move relative to their starting point. They're simply the medium that the sound moves through. None of the air expelled from the person shouting about the truck actually reaches the ear, just the energy itself. It's similar to throwing a rock in the middle of a pond. Waves from the impact will eventually hit the shore, but this isn't the water from the center of the pond, just the energy from the rock's impact.

The speed of sound isn't constant; it varies depending on the stuff it travels through. In air, depending on the purity, temperature, and so forth, sound travels at approximately 740 to 741.5 miles per hour. Sound travels faster through water than through air because water is denser than air (the denser the medium, the faster sound can travel through it). The problem is, humans aren't built to interpret this faster signal in their two ears, and they can't properly pinpoint the signal. For this reason, scuba diving instructors train student divers not to trust their sense of sound localization underwater (for sources such as the dive boat motor). It is just too risky. You can shout at someone underwater and be heard, although the person may not be able to tell where you are.

Cruising with Waves

The universe couldn't exist without waves. Most people have a basic idea of waves, perhaps from watching the ocean or other bodies of water. However, to better understand speech sounds, allow me to further define waves and their properties.

REMEMBER Here are some basic facts about sound and waves:

- Sound is energy transmitted in longitudinal waves.
- Because it needs a medium, sound can't travel in a vacuum.

- Sounds waves travel through media (such as air and water) at different speeds.
- *Sine* (also known as *sinusoid*) waves are simple waves having a single peak and trough structure and a single (fundamental) frequency. The fundamental frequency is the basic vibrating frequency of an entire object, not of its fluttering at higher harmonics.
- People speak in complex waves, not sine waves.
- Complex waves can be considered a series of many sine waves added together.
- Fourier analysis breaks down complex waves into sine waves (refer to the sidebar later in this chapter for more information).
- Complex waves can be *periodic* (as in voiced sounds) or *aperiodic* (as in noisy sounds). Check out the "Sine waves" and "Complex waves" sections for more on periodic and aperiodic waves.

These sections give examples of simple and complex waves, including the relation between the two types of waveforms. I also describe some real-world applications.

Sine waves

The first wave to remember is the *sine wave* (or sinusoid), also called a *simple wave*. Sine is a trigonomic function relating the opposite side of a right-angled triangle to the hypotenuse.

TIP There are some good ways to remember sine waves. Here is a handy list:

- Sine waves are the basic building blocks of the wave world.

- All waveforms can be broken down into a series of sine waves.
- Many things in nature create sine waves — basically anything that sets up a simple oscillation. Figure 12-1 shows a sine wave being created as a piece of paper is pulled under a pendulum that's swinging back and forth.
- In western Texas, if you're lucky, you may see a beautiful sine wave in the sand left by a sidewinder rattlesnake.
- When sound waves are sine waves, they're called *pure tones* and sound cool or cold, like a tuning fork or a flute (not a human voice or a trumpet). This is because the physics of sine wave production involve emphasizing one frequency, either by forcing sound through a hole (as in a flute or whistle) or by generating sound with very precisely machined arms (which reinforce each other as they vibrate), in the case of the tuning fork).

Sine waves are used in clinical audiology for an important test known as *pure-tone audiometry*. Yes, those spooky tones you sometimes can barely hear during an audiology exam are sine waves designed to probe your threshold of hearing. This allows the clinician to rule out different types of hearing loss.

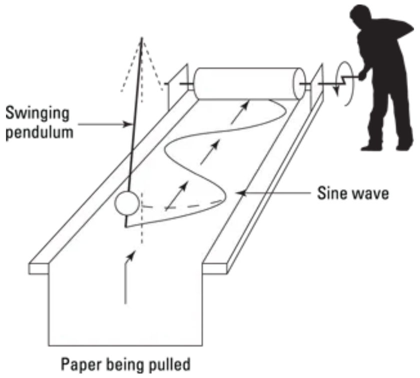

Illustration by Wiley, Composition Services Graphics

Figure 12-1: A pendulum creating sine waves on a piece of paper being pulled by an enthusiastic phonetician.

Complex waves

Everyone knows the world can be pretty complex. Waves are no exception. Unless you're whistling, you don't produce simple waves — all your speech, yelling, humming, whispering, or singing otherwise consists of complex wave production.

A *complex wave* is like a combination of sine waves all piled together. To put it another way, complex waves have more than one simple component — they reflect several frequencies made not by a simple, single vibrating movement (one pendulum motion) but by a number of interrelated motions. It's similar to the way that white light is complex because it's actually a mixture of frequencies of pure light representing the individual colors of the rainbow.

Getting into the formula of sine

If you like formulas, sine waves are created by the sine function:

$$y(t) = A\sin(2\pi ft + \phi)$$

In this formula:

- ***A:*** The *amplitude* is the peak deviation of the function from zero.
- ***f:*** The *frequency* is the number of oscillations (cycles) that occur each second of time.
- **ϕ:** The *phase* specifies (in radians) where in its cycle the oscillation is at $t = 0$.

If you aren't a math fan, no worries!

Measuring Waves

Every wave can be described in terms of its *frequency, amplitude,* and *duration*. But when two or more waves combine, *phase* comes into play. In this section, you discover each of these terms and what they mean to sound.

Frequency

Frequency is the number of times something happens, divided by time. For instance, if you go to the dentist twice a year, your frequency of dental visits is two times per year. But sound waves repeat faster and therefore have a higher frequency.

> Frequency is a very important measure in acoustic phonetics. The number of cycles per second is called *hertz* (Hz) after the famous German physicist Heinrich Hertz. Another commonly used metric is kilohertz (abbreviated kHz), meaning 1,000 Hertz. Thus, 2 kHz = 2,000 Hz = 2,000 cycles per second.

The range of human hearing is roughly 20 to 20,000 cycles per second, which means that the rate of repetition for something to cause such sound is 20 to 20,000 occurrences per second. A bullfrog croaks in the low range (fundamental frequency of approximately 100 Hz), and songbirds sing in the high range (the house sparrow ranges from 675 to 18,000 Hz).

Figure 12-2 shows a sample of frequency demonstrated with a simple example so that you can count the number of oscillations and compute the frequency for yourself. In

Figure 12-2a (periodic wave), you can see that the *waveform* (the curve showing the shape of the wave over time) repeats once in one second (shown on the *x*-axis). Therefore, the frequency is one cycle per second, or 1 Hz. If this were sound, you couldn't hear it because it's under the 20 to 20,000 Hz range that people normally hear.

Seeing is believing! An oscillation can be counted from peak to peak, valley to valley, or zero-crossing to zero-crossing.

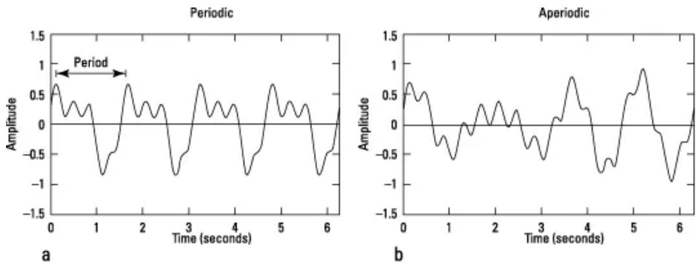

Illustration by Wiley, Composition Services Graphics

Figure 12-2: A sample periodic wave (a) and an aperiodic wave (b).

Period is a useful term related to frequency — it's a measure of the time between two oscillations and the inverse of frequency. If your frequency of dental visits is two times per year, your period of dental visits is every six months.

Waves produced by irregular vibration are said to be *periodic*. These waves sound musical. Sine waves are periodic, and most musical instruments create periodic complex waves. However, waves with cycles of different lengths are *aperiodic* — these sound more like noise. An example would be clapping your hands or hearing a hissing radiator. Figure 12-2b shows an aperiodic wave.

You can also talk about the length of the wave itself. You can sometimes read about the wavelength of light, for example. But did you ever hear about the wavelength of sound? Probably not. This is because wavelengths for sound audible to humans are relatively long, from 17 millimeters to 17 meters, and are therefore rather cumbersome to work with. On the other hand, sound wavelength measurements can be handy for scientists handling higher frequencies, such as ultrasound, which uses much higher frequencies (and therefore much shorter wavelengths).

One frequency that will come in very handy is the *fundamental frequency*, which is the basic frequency of a vibrating body. It's abbreviated F_O and is often called *F-zero* or *F-nought*. A sound's fundamental frequency is the main information telling your ear how low or high a sound is. That is, F_O gives you information about *pitch* (see the section "Relating the physical to the psychological" in this chapter).

Amplitude

Amplitude refers to how forceful a wave is. If there is a weak, wimpy oscillation, there will be a tiny change in the wave's amplitude, reflected on the vertical axis. Such a wave will generally sound quiet. Figure 12-3 shows two waves with the same frequency, where one (shown in the solid line) has twice the amplitude as the other (shown in the dotted line).

Sound amplitude is typically expressed in terms of the air pressure of the wave. The greater the energy behind your yell, the more air pressure and the higher the amplitude of the speech sound. Sound amplitude is also frequently described in decibels (dB). Decibel scales are important and used in many fields including electronics and optics, so it's worth taking a moment to introduce them.

REMEMBER In the following list, I give you the most important things about dB you need to know:

- One dB = one-tenth of a bel.

  The bel was named after Alexander Graham Bell, father of the telephone, which was originally intended as a talking device for the deaf.

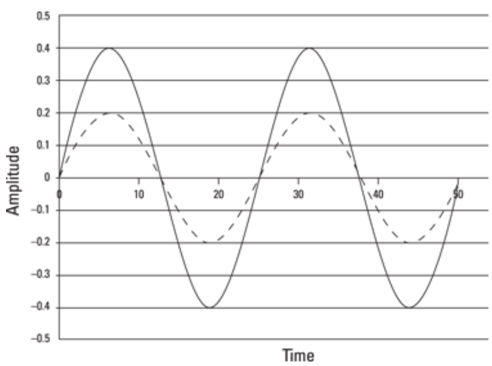

Illustration by Wiley, Composition Services Graphics

Figure 12-3: Two waveforms with the same frequency and different amplitude.

- dB is a logarithmic scale, so an increase of 10 dB represents a ten-fold increase in sound level and causes a doubling of perceived loudness.

  In other words, if the sound of one lawnmower measures 80 dB, then 90 dB would be the equivalent sound of ten lawnmowers. You would hear them twice as loud as one lawnmower.

- Sound levels are often adjusted (weighted) to match the hearing abilities of a given critter. Sound levels adjusted for human hearing are expressed as dB(A) (read as "dee bee A").

  The dBA scale is based on a predefined threshold of hearing reference value for a sine wave at 1000 Hz —

the point at which people can barely hear.

- Conversational speech is typically held at about 60 dBA.
- Too much amplitude can hurt the ears. Noise-induced hearing damage can result from sustained exposure to loud sounds (85 dB and up).

A property associated with amplitude is *damping,* the gradual loss of energy in a waveform. Most vibrating systems don't last forever; they peter out. This shows up in the waveform with gradually reduced amplitude, as shown in Figure 12-4.

Duration

Duration is a measure of how long or short a sound lasts. For speech, duration is usually measured in seconds (for longer units such as words, phrases, and sentences) and *milliseconds* (ms) for individual vowels and consonants.

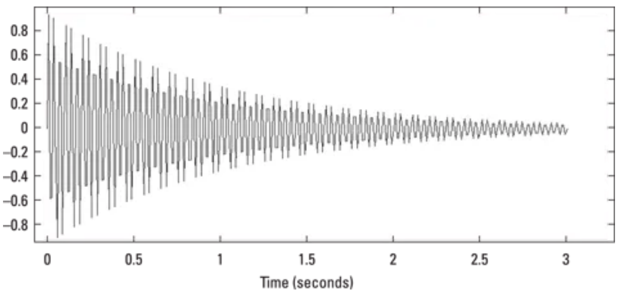

Illustration by Wiley, Composition Services Graphics

Figure 12-4: Damping happens when there is a loss of vibration due to friction.

Phase

Phase is a measure of the time (or angle) between two similar events that run at roughly the same time. Phase can't be measured with a single sound — you need two (waves) to tango. Take a look at Figure 12-5 to get the idea of how it works:

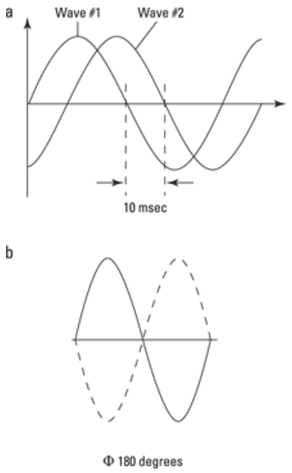

Illustration by Wiley, Composition Services Graphics

Figure 12-5: Two examples of phase differences — by time (a) and by angle (b).

In the top example of Figure 12-5, when wave #1 starts out, wave #2 lags by approximately 10 msec. That is, wave #2 follows the same pattern but is 10 msec behind. This is phase described by time.

The bottom example in Figure 12-5 shows phase described by angle. Two waves are 180 degrees out of phase. This example is described by *phase angle,* thinking of a circle, where the whole is 360 degrees and the half is 180 degrees. To be *180 degrees out of phase* means that when one wave is at its peak, the other is at its valley. It's kind of like a horse race. If one horse is a quarter of a track behind the other horse, you could describe him as being so many yards, or 90 degrees, or a quarter-track behind.

Relating the physical to the psychological

In a perfect world, what you see is what you get. The interesting thing about being an (imperfect) human is that the physical world doesn't relate in a one-to-one fashion with the way people perceive it. That is, just because something vibrates with such and such more energy doesn't mean you necessarily hear it as that much louder. Settings in your perceptual system make certain sounds seem louder than others and can even set up auditory illusions (similar to optical illusions in vision).

This makes sense if you consider how animals are tuned to their environment. Dogs hear high-pitched sounds, elephants are tuned to low frequencies (infrasound) for long-distance communication, and different creatures have different perceptual settings in which trade-offs between frequency, amplitude, and duration play a role in perception. Scientists are so intrigued by this kind of thing that they have made it into its own field of study — *psychophysics,* which is the relationship between physical stimuli and the sensations and perceptions they cause.

Pitch

The psychological impression of fundamental frequency is called *pitch*. High-frequency vibrations sound like high notes, and low-frequency vibrations sound like low notes. The ordinary person can hear between 20 to 20,000 Hz. About 30 to 35 percent of people between 65 and 75 years of age may lose some hearing of higher-pitched sounds, a condition called *presbycusis* (literally "aged hearing").

Loudness

People hear amplitude as *loudness,* a subjective measure that ranges from quiet to loud. Although many measures of sound strength may attempt to adjust to human loudness values, to really measure loudness values is a complex process — it requires human listeners.

REMEMBER Different sounds with the very same amplitude won't have the same loudness, depending on the frequency.

> If two sounds have the same amplitude and their frequencies lie between about 600 and 2,000 Hz, they'll be perceived to be about the same loudness. Otherwise, things get weird! For sounds near 3,000 to 4,000 Hz, the ear is extra-sensitive; these sounds are perceived as being louder than a 1,000 Hz sound of the same amplitude. At frequencies lower than 300 Hz, the ear becomes less sensitive; sounds here are perceived as being less loud than they (logically) "should" be.

This means I can freak you out with the following test. I can play you a 300 Hz tone, a 1,000 Hz tone, and a 4,000 Hz tone, all at *exactly* the same amplitude. I can even show you on a sound-level meter that they are exactly the same. However, although you know they are all the same, you'll *hear* the three as loud, louder, and loudest. Welcome to psychophysics.

Length

The psychological take on duration is *length*. The greater the duration of a speech sound, the longer that signal generally sounds. Again, however, it's not quite as simple as it may seem. Some languages have sounds that listeners hear as double or twin consonants. (***Note:*** Although English spelling has double "n," "t," and so forth, it doesn't always pronounce these sounds for twice as long.) Doubled consonant sounds are called *geminates* (twins). It turns out that geminates are usually about twice the duration as nongeminates. However, it depends on the language. In Japanese, for example, geminates are produced about two to three times as long as nongeminates. An example is /hato/ "dove" versus /hatto/ "hat."

Sound localization

Humans and other creatures use phase for *sound localization,* which allows them to tell where a sound is coming from. A great way to test whether you can do this is to sit in a chair, shut your eyes, and have a friend stand about 3 feet behind you. Have her snap her fingers randomly around the back and sides of your head. Your job is to point to the snap, based only on sound, each time.

Most people do really well at this exercise. Your auditory system uses several types of information for this kind of task, including the time-level difference between the snap waveform hitting your left and right ears, that is — phase. After more than a century of work on this issue, researchers still have a lot to learn about how humans localize sound. There are many important practical applications for this question, including the need to produce better hearing aids and communication systems (military and commercial) that preserve localization information in noisy environments.

Just a phase I'm going through?

The Wright-Patterson Air Force Base in Dayton, Ohio, has an amazing sound localization laboratory containing a geodesic sphere, nearly 10 feet across, holding 277 Bose loudspeakers. Listeners zapped by various sounds from all angles indicate where the sounds came from by pointing on a small globe with a special electromagnetic pen. It allows researchers to conduct experiments designed to determine how people can pinpoint sound source location with such stunning accuracy.

A promising new avenue of development for sound localization technology is the microphone array, where systems for extracting voice can be built by setting up a series of closely spaced microphones that pick up different phase patterns. This allows the system to provide better spatial audio and in some cases reconstruct “virtual” microphones to accept or reject certain sounds. In this way, voicing input in noisy environments can sometimes be boosted — a big problem for people with hearing aids.

Harmonizing with harmonics

The basic opening and closing gestures of your vocal folds produce the fundamental frequency (F_O) of phonation. If you were bionic and made of titanium, this is all you would produce. In such a case, your voice would have only a fundamental frequency, and you would sound, well, kind of creepy, like a tuning fork. Fortunately, your fleshy and muscular vocal folds produce more than just a fundamental frequency — they also produce *harmonics,* which are additional flutters timed with the fundamental frequency at numbered intervals. Harmonics are regions of energy at integer multiples of the fundamental frequency. They’re properties of the voicing source, not the filter.

REMEMBER Harmonics result whenever an imperfect body — like a rubber band, guitar string, clarinet reed, or vocal fold — vibrates. If you could look at one such cycle, slowed down, with Superman’s eyes, you’d see that there’s not only a basic (or fundamental) vibration, but also there’s a whole series of smaller flutters that are timed with the basic vibration. These vibrations are smaller in amplitude, and (here is the amazing thing) they're spaced in frequency by whole numbers. So, if you're a guy and your fundamental frequency is 130 Hz (also known as your first harmonic), then your second harmonic would be 260 Hz, your third harmonic 390, and so forth. For a female with a higher fundamental frequency, say at 240 Hz, the second harmonic would be 480 Hz, the third 720, and so on. Harmonics are found throughout the speech frequency range (20 to 20,000 Hz). However, there's more energy in the lower frequencies than in the higher because of a 12 dB per octave cutoff.

Extreme harmonics: Phonetics at the edge

A favorite classroom demonstration of mine is to take an enormous strip of rubber from a tire inner tube and stretch it across a phonetics class. Somebody grabs the middle of the inner tube strip and pulls it across to one side of the classroom, everybody ducks, and then the strip is released. As the strip zings back and forth, a few things visibly happen:

- Students can clearly see the fundamental frequency (F_0) of the band as it flies back and forth.
- The band wobbles, showing the harmonics — sub-periodic oscillations that occur at whole number multiples of the fundamental frequency.
- Everyone begins to laugh nervously because (after all) they haven't been hit by the giant, dangerous piece of rubber.
- A few students discreetly call their parents or attorneys.

This is the way of nature — you set up a *simple harmonic series*. Each harmonic series includes a fundamental frequency (or first harmonic) and an array of harmonics that have the relations times 2, times 3, times 4, and so on. Figure 12-6 shows these relations on a vibrating string.

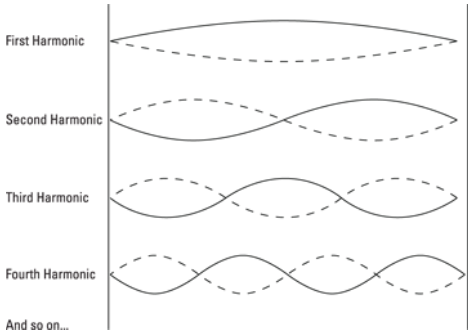

Illustration by Wiley, Composition Services Graphics

Figure 12-6: Harmonic series on a vibrating string.

This spectrum of fundamental frequency plus harmonics gives much of the warmth and richness to the human voice, something in the music world that makes up *timbre* (tone color or tone quality).

Resonating (Ommmm)

Producing voicing is half the story. After you've created a voiced source, you need to shape it. Acoustically, this shaping creates a condition called *resonance,* strengthening of certain aspects of sound and weakening of others. Resonance occurs when a sound source is passed through a structure.

Think about honking your car horn in a tunnel — the sound will carry because the shape of a tunnel boosts it. This kind of resonance occurs as a natural property of physical bodies. Big structures boost low sounds, small structures boost high sounds, and complex-shaped structures may produce different sound qualities.

Think of the shapes of musical instruments in a symphony — most of what you see has to do with resonance. The tube of a saxophone and the bell of a trumpet exist to shape sound, as does the body of a cello.

> The parts of your body above the vocal folds (the lips, tongue, jaw, velum, nose, and throat) are able to form complicated passageway shapes that change with time. These shape changes have a cookie-cutter effect on your spectral source, allowing certain frequencies to be boosted and others to be dampened or suppressed. Figure 12-7 shows how this works acoustically during the production of three vowels, /i/, /ɑ/, and /u/.

Imagine that a crazed phonetician somehow places a microphone down at the level of your larynx just as you make each vowel. There would be only a neutral vibratory source (sounding something like an /ə/) for all three. The result would be a spectrum like the one at the bottom of Figure 12-7. Notice that this spectrum has a fundamental frequency and harmonics, as you might expect. When the vocal tract is positioned into different shapes for the three vowels (shown in the middle row of the figure), this has the effect of strengthening certain frequency areas and weakening others. This is resonance. By the time speech finally comes out the mouth, the acoustic picture is complex (as shown in the top of Figure 12-7). You can still see the fundamental frequency and harmonics of the source; however, there are also broad peaks. These are *formants,* labeled F1, F2, and F3.

18. True or False: English speakers typically release syllable-final stops in casual speech.
19. True or False: The vowel in "button" is typically pronounced with a clear [ə].
20. True or False: Phonemic distinctions are crucial for understanding meaning in a language.
21. True or False: In English, the vowel in "seen" is always nasalized.
22. True or False: The coda of a syllable can only contain consonants.
23. True or False: A syllable consists of a nucleus and an optional coda.
24. True or False: The stress of syllables is always marked in broad transcription.
25. True or False: English has more rules for syllable structure than many other languages.
26. True or False: The IPA symbol for a voiced bilabial stop is [b].
27. True or False: "Tap" and "flap" are different terms for the same phonetic phenomenon.
28. True or False: All English speakers produce glottal stops in the same environments.
29. True or False: Syllabic consonants can act as the nucleus of a syllable.
30. True or False: The term "sonority" refers to the loudness of a sound.
31. True or False: The schwa sound [ə] is always stressed in English.
32. True or False: Aspiration is a feature of all consonants in English.
33. True or False: The main function of stress in English is to indicate the end of a word.
34. True or False: Voiceless stops cannot be aspirated in English.
35. True or False: Syllables can be classified based on their structure.
36. True or False: English allows for consonant clusters at the beginning of syllables.
37. True or False: The term "prosody" refers to the rhythm and melody of speech.
38. True or False: Phonetic rules apply uniformly across all dialects of English.
39. True or False: The vowel in "cat" is a diphthong.
40. True or False: The IPA symbol for a voiceless velar stop is [k].
41. True or False: Stress can change the meaning of a word in English.
42. True or False: The glottal stop can occur at the beginning of a word in English.
43. True or False: The diacritic [˺] indicates that a stop is unreleased.
44. True or False: Syllables can only contain one vowel sound.
45. True or False: Stress is a suprasegmental feature of language.
46. True or False: The IPA provides a one-to-one correspondence between sounds and symbols.
47. True or False: The word "little" contains a dark "l" in its final syllable.

48. True or False: Liquids include sounds like [l] and [r].
49. True or False: The symbol for a nasal consonant is [n].
50. True or False: Vowel centralization occurs in casual speech.
51. True or False: Stress is always placed on the first syllable of English words.
52. True or False: Phonological rules can vary based on speaking context.
53. True or False: The IPA symbol for a voiced velar stop is [g].
54. True or False: Juncture can affect the meaning of phrases in spoken language.
55. True or False: All nasalized vowels are fully nasalized in English.
56. True or False: The relationship between stress and meaning is complex in English.
57. True or False: The IPA uses the same symbols for all languages.
58. True or False: The syllable structure of a language can affect its rhythm.
59. True or False: The IPA character for a voiceless labiodental fricative is [f].
60. True or False: The word "written" often contains a glottal stop in American English.
61. True or False: The nucleus of a syllable can be a consonant.
62. True or False: The diacritic for aspiration is a superscript "h."
63. True or False: The vowel sound in "mop" is a monophthong.
64. True or False: Stress typically falls on the second syllable in compound nouns.
65. True or False: Phonetic transcription is used only in academic settings.
66. True or False: English is a stress-timed language.
67. True or False: The IPA character for a voiceless glottal fricative is [h].
68. True or False: Consonants can only appear at the beginning or end of a syllable.
69. True or False: The rule of stress applies consistently across all English dialects.
70. True or False: The vowel in "moon" is a diphthong.
71. True or False: Phonetic rules are not relevant for understanding spoken language.
72. True or False: The coda of the syllable "dog" is [g].
73. True or False: The term "liquid" refers to a type of vowel.
74. True or False: A syllable can exist without a nucleus.
75. True or False: The IPA symbol for a voiced palatal fricative is [ʒ].
76. True or False: The pronunciation of "bat" and "pat" illustrates a minimal pair.
77. True or False: The IPA allows for detailed representation of speech sounds.

78. True or False: Vowels can be nasalized in English in certain phonetic contexts.
79. True or False: Stress patterns in English can be predicted based on word structure.
80. True or False: The IPA character for a voiced uvular fricative is [χ].
81. True or False: The word "singer" contains a syllabic nasal.
82. True or False: The diacritic [̪] indicates that a sound is dentalized.
83. True or False: Phonological rules apply only to consonants.
84. True or False: The IPA uses the same symbols regardless of context.
85. True or False: The concept of juncture is irrelevant to linguistics.
86. True or False: The word "candle" features a light "l."
87. True or False: The vowel sound in "hat" can be characterized as a short vowel.
88. True or False: Phonetic transcription is an essential tool for language learning.
89. True or False: The IPA character for a voiced dental fricative is [ð].
90. True or False: Syllables can be classified into open and closed types.
91. True or False: The term "phoneme" refers to the smallest unit of sound that can change meaning.
92. True or False: The IPA character for a voiceless postalveolar fricative is [ʃ].
93. True or False: The word "teacher" contains a stress on the second syllable.
94. True or False: The symbol for a nasalized vowel is [ã].
95. True or False: The IPA provides a systematic way to transcribe sounds from any language.
96. True or False: The coda in the word "cat" is a vowel sound.
97. True or False: The concept of sonority is important in phonological rules.
98. True or False: The IPA character for a voiceless glottal stop is [ʔ].
99. True or False: Speech sounds can be analyzed for their articulatory properties.
100. True or False: The IPA helps in understanding dialectal variations in pronunciation.

TEST YOUR SELF

(91-107)

1. What is the primary feature of a stressed syllable in English?
   - A) Shorter duration
   - B) Lower pitch
   - C) Louder and longer
   - D) Voiceless

2. The word "frightening" can be pronounced with:
   - A) One syllable
   - B) Two or three syllables
   - C) Only three syllables
   - D) None of the above

3. What term describes the combination of sonority, length, stress, and pitch?
   - A) Intonation
   - B) Prominence
   - C) Elision
   - D) Assimilation

4. Which of the following words is an example of a two-syllable pronunciation?
   - A) Lightning
   - B) Maddening
   - C) Frightening
   - D) All of the above

5. In the phrase "She never wears Spandex!" what changes the emphasis?
   - A) The pitch of the speaker
   - B) The length of the sentence
   - C) The choice of words
   - D) The stress on different words

6. What is a tonic syllable?
   - A) A syllable that is always stressed
   - B) A syllable that carries the major pitch change
   - C) A syllable at the beginning of a word
   - D) A syllable with no vowel

7. Which of the following is true about connected speech?
   - A) It is always clear and distinct
   - B) It may include hesitations and filled pauses
   - C) It does not require prosodic transcription

- D) It is always grammatically correct

8. An intonational phrase typically contains:
   - A) Multiple tonic syllables
   - B) One tonic syllable
   - C) No stressed syllables
   - D) Only consonant sounds

9. What is the function of pitch changes in speech?
   - A) To indicate the speaker's mood
   - B) To differentiate between sounds
   - C) To mark sentence boundaries
   - D) All of the above

10. The diacritic for a syllabic consonant is:
    - A) [ː]
    - B) [̩]
    - C) [ʔ]
    - D) [ˤ]

11. What does "elision" refer to in phonology?
    - A) The addition of sounds
    - B) The substitution of sounds
    - C) The omission of sounds
    - D) The alteration of sounds

12. Which symbol represents a voiced velar stop?
    - A) [k]
    - B) [b]
    - C) [g]
    - D) [d]

13. How is stress typically marked in a transcription?
    - A) With an asterisk (*)
    - B) With a dot (•)
    - C) With a primary stress mark ([ˈ])
    - D) With a secondary stress mark ([ˌ])

14. What is the main purpose of using ToBI in transcription?
    - A) To indicate vowel length
    - B) To analyze pitch and prosody
    - C) To mark consonant sounds
    - D) To simplify transcriptions

15. A continuation rise in intonation indicates:
  - A) A change in topic
  - B) A question
  - C) A lack of finality in the phrase
  - D) A complete thought

16. In a narrow transcription, which detail is included?
  - A) Only the phonemes
  - B) Allophones and phonetic variations
  - C) Only stress patterns
  - D) Syllable structure only

17. Which of the following is an example of a filled pause?
  - A) "Um"
  - B) "Like"
  - C) "Uh"
  - D) All of the above

18. What is the role of stress in distinguishing between nouns and verbs?
  - A) It has no role
  - B) Stress can change the meaning
  - C) Stress is always the same
  - D) It only applies to adjectives

19. Which phrase contains a tonic syllable?
  - A) "The dog"
  - B) "Eileen is really crazy"
  - C) "She likes apples"
  - D) "The sky is blue"

20. What happens to stops before nasals according to phonological rules?
  - A) They become voiced
  - B) They are released
  - C) They become glottal stops
  - D) They are elongated

21. An intonational phrase is often marked by:
  - A) Commas
  - B) Vertical lines
  - C) Dashes
  - D) All of the above

22. The IPA symbol [ʃ] represents which sound?
- A) Voiced alveolar fricative
- B) Voiceless postalveolar fricative
- C) Voiceless alveolar fricative
- D) Voiced postalveolar fricative

23. In the sentence "The boy fed the dog," which word carries the tonic syllable?
- A) The
- B) Boy
- C) Fed
- D) Dog

24. Which rule states that stops become glottal stops before nasals?
- A) Rule No. 5
- B) Rule No. 6
- C) Rule No. 7
- D) Rule No. 8

25. Stress in phrases can change the:
- A) Meaning
- B) Tone
- C) Length
- D) Word order

26. What is one characteristic of emotional prosody?
- A) It does not affect meaning
- B) It reflects the speaker's mood
- C) It is always consistent
- D) It only occurs in written language

27. A phrase accent in ToBI is used to:
- A) Mark sentence boundaries
- B) Indicate pitch changes
- C) Show word stress
- D) None of the above

28. What does the symbol [ˈ] indicate in transcription?
- A) A secondary stress
- B) A primary stress
- C) A voiceless sound
- D) A long vowel

29. Which of the following is an example of a tag question?
- A) "You are coming."
- B) "You are coming, aren't you?"
- C) "Are you coming?"
- D) "I think you are coming."

30. How is a pause represented in ToBI?
- A) With a comma
- B) With a break index
- C) With a vertical line
- D) With a dash

31. What is the primary function of intonation in speech?
- A) To create pauses
- B) To convey emotion
- C) To indicate questions or statements
- D) All of the above

32. In prosodic transcription, which aspect is crucial to capture?
- A) Vowel quality
- B) Stress patterns
- C) Consonant length
- D) Syllable count

33. The concept of meter in phonetics refers to:
- A) The speed of speech
- B) The rhythm of speech
- C) The volume of speech
- D) The clarity of speech

34. What does the term "hyperspeech" refer to?
- A) Casual speech
- B) Formal speech with clear articulation
- C) Speech with many pauses
- D) Speech that is overly loud

35. How do phoneticians define a tonic syllable?
- A) The last syllable in a word
- B) The syllable that carries the most pitch change
- C) Any stressed syllable
- D) The first syllable in a phrase

36. Which type of transcription captures allophonetic details?

- A) Broad transcription
- B) Phonemic transcription
- C) Narrow transcription
- D) Orthographic transcription

37. What is the main focus of metrical phonology?
- A) Sound production
- B) Organizing segments into groups of relative prominence
- C) Analyzing syntax
- D) Studying historical linguistics

38. What does a high boundary tone (H%) indicate?
- A) A declarative statement
- B) A question
- C) A continuation rise
- D) A pause

39. Which phrase is an example of a continuation rise?
- A) "I love pizza."
- B) "He is nice, isn't he?"
- C) "She bought apples, bananas, and grapes."
- D) "That was fun."

40. In the context of prosodic transcription, what does "pitch contour" refer to?
- A) The loudness of speech
- B) The pattern of pitch changes
- C) The speed of speech
- D) The length of sentences

41. What do phoneticians use to denote stress in transcriptions?
- A) Numbers
- B) Letters
- C) Diacritics
- D) Marks like [ˈ] and [ˌ]

42. Which of the following best describes "hypospeech"?
- A) Speech in a formal setting
- B) Casual and relaxed speech
- C) Speech with many pauses
- D) Overly loud speech

43. How does the IPA recommend capturing prosody?

- A) Using one standard method
- B) Applying various rules and theories
- C) Ignoring prosody in transcription
- D) Relying solely on stress patterns

44. What does the break index in ToBI represent?
   - A) The length of words
   - B) The duration of pauses
   - C) The number of syllables
   - D) The volume of speech

45. In connected speech, what might affect the clarity of transcription?
   - A) Hesitations
   - B) Filled pauses
   - C) Rapid speech
   - D) All of the above

46. What is the purpose of marking pitch changes in speech?
   - A) To indicate breath groups
   - B) To show emphasis or focus
   - C) To clarify meaning
   - D) All of the above

47. Which of the following is a feature of linguistic prosody?
   - A) Emotional tone
   - B) Melody and timing
   - C) Syllable count
   - D) Word choice

48. What is indicated by a falling tone at the end of a sentence?
   - A) A question
   - B) A statement
   - C) Uncertainty
   - D) Excitement

49. The symbol [ɾ] represents which sound in English?
   - A) Voiced alveolar stop
   - B) Voiceless glottal stop
   - C) Flap or tap sound
   - D) Voiced velar stop

50. Which element is NOT part of an intonational phrase?
   - A) Tonic syllable

- B) Pre-head
- C) Word length
- D) Head

51. How is a filled pause represented in transcription?
   - A) Written as it is
   - B) Enclosed in parentheses
   - C) Ignored completely
   - D) Transcribed phonetically

52. What is a common characteristic of a tag question?
   - A) It always rises in intonation
   - B) It is always falling
   - C) It can be either rising or falling
   - D) It has no specific pattern

53. A speech that is characterized by many pauses and hesitations is known as:
   - A) Hyperspeech
   - B) Hypospeech
   - C) Hesitant speech
   - D) Fluent speech

54. In the context of prosody, what does "intonation" refer to?
   - A) The volume of speech
   - B) The melody of speech
   - C) The speed of speech
   - D) The clarity of speech

55. The IPA symbol for a voiceless stop is:
   - A) [b]
   - B) [d]
   - C) [p]
   - D) [g]

56. Which of the following is NOT a characteristic of a stressed syllable?
   - A) Increased loudness
   - B) Shorter duration
   - C) Higher pitch
   - D) Distinct articulation

57. How do you identify a tonic syllable in connected speech?
   - A) It is always the first syllable

- B) It carries the most pitch change
- C) It is always stressed
- D) It is the last syllable

58. In prosodic transcription, what does a low boundary tone (L%) indicate?
   - A) A question
   - B) A continuation
   - C) A final fall
   - D) A statement

59. Which of the following best describes the role of stress in phrases?
   - A) It is irrelevant
   - B) It can change the meaning and focus
   - C) It is always the same
   - D) It only applies to nouns

60. The term "sonority" pertains to:
   - A) The height of a vowel
   - B) The loudness and quality of a sound
   - C) The length of a consonant
   - D) The speed of speech

61. Which transcription method includes fine-grained details of speech?
   - A) Broad transcription
   - B) Narrow transcription
   - C) Phonemic transcription
   - D) Orthographic transcription

62. What does the term "connected speech" refer to?
   - A) Speech that is grammatically correct
   - B) Speech produced in isolation
   - C) Natural speech in conversation
   - D) Formal speech

63. Which of the following terms describes a syllable that stands out due to pitch change?
   - A) Pre-head
   - B) Tonic syllable
   - C) Head
   - D) Tail

64. What is an example of a non-linguistic vocalization?

- A) "Um"
- B) Laughter
- C) Coughing
- D) All of the above

65. How is emphasis indicated in prosody?
- A) Through loudness
- B) Through pitch changes
- C) Through stress patterns
- D) All of the above

66. What does the term "intonational phrase" refer to?
- A) A single word
- B) A group of sounds with a specific pitch pattern
- C) A grammatical unit
- D) A type of vowel sound

67. Which diacritic indicates primary stress?
- A) [ˌ]
- B) [ˈ]
- C) [˩]
- D) [ː]

68. The pitch contour indicates:
- A) The grammatical structure
- B) The emotional tone
- C) The pattern of pitch changes in speech
- D) The number of syllables

69. What does the term "break index" in ToBI indicate?
- A) The number of words
- B) The strength of pauses
- C) The length of phrases
- D) The amount of stress

70. In prosodic transcription, how are breath groups indicated?
- A) With commas
- B) With vertical lines
- C) With dashes
- D) With spaces

71. The feature of "dark l" typically occurs:
- A) At the beginning of a word

- B) At the end of a syllable
- C) Before vowels
- D) After voiceless stops

72. Which of the following is true about syllabic consonants?
   - A) They cannot stand alone
   - B) They often occur at the end of words
   - C) They are always voiced
   - D) They have no vowel quality

73. What is the main focus of this document?
   - A) Historical linguistics
   - B) Phonetics and prosody
   - C) Syntax and semantics
   - D) Language acquisition

74. How do phoneticians analyze intonation in speech?
   - A) By ignoring pitch changes
   - B) By applying various transcription methods
   - C) By focusing solely on stress
   - D) By analyzing grammar

75. The IPA symbol [d] represents which sound?
   - A) Voiced alveolar stop
   - B) Voiceless alveolar stop
   - C) Voiced velar stop
   - D) Voiceless velar stop

76. Which element is NOT part of an intonational phrase structure?
   - A) Nucleus
   - B) Pre-head
   - C) Stress
   - D) Tail

77. What is the function of a phrase accent in ToBI?
   - A) To indicate pitch changes
   - B) To mark sentence boundaries
   - C) To show word stress
   - D) To denote syllable length

78. Which transcription method is typically used for clinical notes?
   - A) Narrow transcription
   - B) Broad transcription

- C) Orthographic transcription
- D) Phonemic transcription

79. In the phrase "The boy fed the dog," which word is emphasized if it is the tonic syllable?
   - A) The
   - B) Boy
   - C) Fed
   - D) Dog

80. How do phoneticians define "prosody"?
   - A) The study of syntax
   - B) The melody and rhythm of speech
   - C) The sounds of individual phonemes
   - D) The meaning of words

81. What does "downdrift" refer to in prosodic analysis?
   - A) A rise in pitch
   - B) A gradual decline in pitch
   - C) A sudden drop in volume
   - D) A lengthening of syllables

82. Which of the following is an example of a rising tone in a tag question?
   - A) "You're coming, aren't you?"
   - B) "It's a nice day, isn't it?"
   - C) "He sold you a fake Rolex, didn't he?"
   -