Linguistics for Language Teachers

Lessons for Classroom Practice

Sunny K. Park-Johnson and Sarah J. Shin

Routledge
Taylor & Francis Group
NEW YORK AND LONDON

First published 2020
by Routledge
52 Vanderbilt Avenue, New York, NY 10017

and by Routledge
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

Library of Congress Cataloging-in-Publication Data
A catalog record for this book has been requested

ISBN: 978-1-138-68182-8 (hbk)
ISBN: 978-1-138-68193-4 (pbk)
ISBN: 978-1-315-54546-2 (ebk)

Typeset in Times New Roman
by Apex CoVantage LLC

2 Phonetics

The Sounds of Language

2.1 Introduction

This chapter describes the physiological mechanisms of speech production and how we use articulators in the vocal tract to produce specific sounds. It will explain why a phonetic transcription is necessary to represent sounds in different languages and introduce the International Phonetic Alphabet. We will practice transcribing our own speech using these symbols and learn to group sounds into different classes based on their articulatory and acoustic properties. In addition, we will learn how a focus on prosodic properties (e.g., intonation, stress, and length), in addition to phonetic properties of individual sounds, can facilitate our teaching of pronunciation to second language learners.

In this chapter, we begin with individual sounds, the smallest unit of language. We normally do not think about how sounds are produced in our native languages. The reason that language teachers learn phonetics is to be sensitized to the properties of sounds in different languages—how they are produced (articulatory properties), and what they sound like (acoustic properties)—so that they can help their students to accurately distinguish and produce different sounds. In order to be effective in this endeavor, teachers will want to become familiar with the phonetic inventories of not only English but also the students' native languages so they can identify potential areas of difficulty for the students.

2.2 Why Do We Need a Phonetic Alphabet?

A good phonetic transcription system should have an unambiguous and consistent relationship between written symbols and the sounds that they represent. Each symbol should represent one sound only, and each sound should have only one symbol. The reason that we cannot use the English spelling system to describe sounds is that there is no one-to-one correspondence between English orthography and the sounds it represents. English spelling is highly irregular—the same letter can represent different sounds, or the same sound can be represented by different letters. For example, the letter *s* can represent a number of different sounds in English writing. It can represent the [s] sound

in words such as *sun, fast, phonetics,* the [z] sound in *is, use, thieves* or the [ʒ] sound in words like *pleasure, leisure* or no sound at all in *aisle, island, debris*. Conversely, the [i] sound in English can be written using different letters of the English alphabet, as in *see, sea, icy, ceiling, scenic, ravine, brief*. Additionally, the English alphabet is not able to accurately represent sounds that are not in English, such as the [ħ] sound in Arabic or the click sounds in Xhosa.

A phonetic alphabet solves these problems by representing each sound in human speech with a single symbol. Using a phonetic alphabet enables us to transcribe spoken language consistently and accurately. In this book, we will use the International Phonetic Alphabet (IPA), found in the inside back cover of this book (IPA Chart, 2015). The IPA is applicable to all spoken human languages and can be used to describe the sounds of any language. The symbols in the IPA are enclosed in slashes / / or brackets [ ] to indicate that the transcription is phonetic. The IPA does not represent the spelling system of any particular language.

2.3 Articulatory Phonetics: How Sounds Are Produced

Most speech sounds are made by pushing air out of the lungs. Try talking while breathing in and you will notice it's much more difficult than talking while breathing out. Air from the lungs goes up the trachea (also known as the windpipe) and into the larynx, where it passes through the space between the **vocal cords**, called the **glottis**. After the air passes through the glottis, it goes through the tube in the throat called the **pharynx**. Then the air goes out of the **oral cavity** through the mouth, or out of the **nasal cavity** through the nose. Sounds made when the vocal cords are vibrating are called **voiced**, while those made when the vocal cords are apart are called **voiceless**. To hear the difference between a voiced and a voiceless sound, put your fingertips lightly on your throat and say a long *s* sound (like a snake). You should feel no vibration as the vocal cords are separated to make this voiceless sound. Now with your fingertips still on your throat, say a long *z* sound. You will feel a vibration of the vocal cords for this voiced sound.

The difference in voiced and voiceless sounds is important for distinguishing sounds. Table 2.1 lists voiceless consonants in English in the left column and their voiced counterparts in the right column. The underlined sounds in the first row, [p] and [b], differ only in voicing. To check this for yourself, put your fingertips on your throat and say just the underlined consonant in each of these words. Say [p] and [b] alternately—[p, b, p, b, p, b]. Notice that both of these sounds are formed in the same way in the mouth. The only difference is that [p] is voiceless whereas [b] is voiced. The same goes for [f] and [v], for [θ] and [ð], for [s] and [z], and so on.

The air passages above the larynx are called the **vocal tract.** Figure 2.1 shows the principal parts of the vocal tract that can be used to make sounds.

Table 2.1 Voiceless Consonants and Their Voiced Counterparts

Voiceless	*Voiced*
pad [p]	bad [b]
face [f]	vase [v]
thigh [θ]	thy [ð]
see [s]	zee [z]
dilution [ʃ]	delusion [ʒ]
rich [ʧ]	ridge [ʤ]
tame [t]	dame [d]
coal [k]	goal [g]

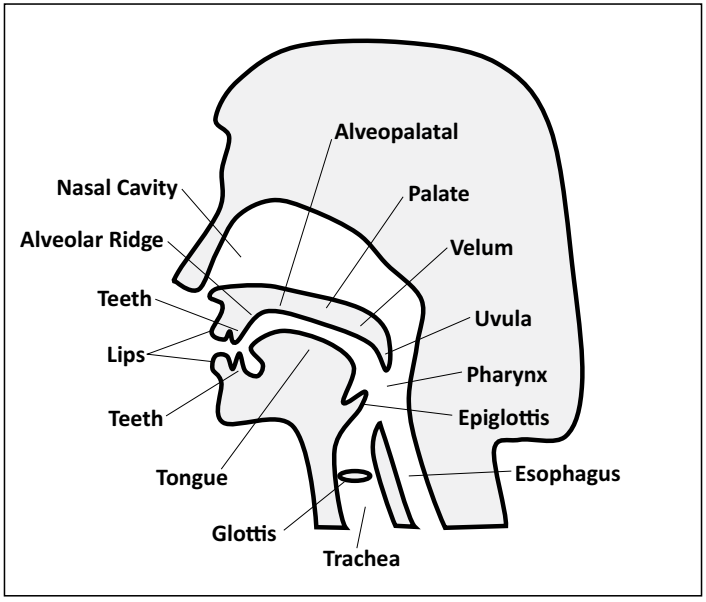

Figure 2.1 The Vocal Tract

These are called **articulators**. The lips and teeth are visible articulators in the front part of the mouth. Just behind the upper teeth is a small protuberance that can be felt with the tip of the tongue, called the **alveolar ridge**. Behind the alveolar ridge is the **hard palate**, a bony structure in the roof of the mouth. Going further back you will find the **soft palate,** or **velum**. If you have difficulty curling up the tongue far enough to touch the soft palate, you may need to use a fingertip to feel it. The soft palate is a muscular flap that can be raised to press against the back wall of the pharynx to prevent air from escaping through the nose. When this happens, the air can go out only

through the mouth, resulting in an **oral sound**. When the soft palate is lowered, however, the air goes out through the nose, resulting in a **nasal sound**. At the lower end of the soft palate is the **uvula**, a fleshy extension that hangs above the throat.

The most important articulator is the **tongue**, a fleshy muscular organ that can produce incredibly fine and complex movements. The tongue can be raised, lowered, pushed forward or pulled back. There are specific names for different parts of the tongue. The **tip** is the narrow area at the front of the tongue. Just behind the tip lies the **blade**. The main mass of the tongue is called the **tongue body**, which may be divided into the **front**, the **center**, and the **back**. The **root** of the tongue lies opposite the back wall of the pharynx.

The sounds of all languages fall into two major categories: consonants and vowels. Consonants are produced with some restriction in the vocal tract as the air from the lungs is pushed through the glottis out of the mouth or the nose. In the production of vowels, the passage of airstream is relatively unobstructed. In the following sections, we will first review characteristics of English consonants and vowels, then look at sounds found in other languages.

2.4 Consonants

Consonants are formed by the *obstruction* of the airstream through the vocal tract. Consonants can therefore be classified according to the place and manner of this obstruction, along with voicing at the larynx. Any given consonant is described by using the following three features: (1) Voicing, (2) Place of Articulation, and (3) Manner of Articulation. Table 2.2 shows the consonants of Mainstream American English, with places of articulation listed from left to right and manners of articulation listed from top to bottom.

Table 2.2 Consonants of Mainstream American English

	Bilabial	*Labiodental*	*Interdental*	*Alveolar*	*Post-Alveolar*	*Palatal*	*Velar*	*Glottal*
Stop	p b			t d			k g	ʔ
Fricative		f v	θ ð	s z	ʃ ʒ			h
Affricate					ʧ ʤ			
Nasal	m			n			ŋ	
Lateral Liquid				l				
Retroflex Liquid				r				
Glide	ʍ w					j	ʍ w	

Note: Where symbols appear in pairs, the one to the right represents a voiced consonant while the one to the left represents a voiceless consonant.

2.4.1 Places of Articulation

Place of articulation refers to the point where the airstream is obstructed to produce a different sound. Places of articulation are found at the lips, in the oral cavity, in the pharynx, and at the glottis.

(1) **Bilabials** are sounds produced by bringing the two lips together (*bi-* means two, *labial* means lips). In front of a mirror, say words such as *pay, bay, may* and see how the lips come together for the first sound of each of these words. Then say words such as *app, ab, am* and see how the lips come together for the last sound of each word. Then notice what happens as you say the initial sound in words such as *why, win*. This [w] sound involves rounding the lips without stopping the airstream. [w̥], the voiceless version of [w] occurs in some dialects of English, resulting in different pronunciation of words like *witch* and *which*.
(2) **Labiodentals** are produced by touching the lower lip to the upper front teeth (*labio-* means lips and *dental* means teeth). In front of a mirror, say the words *fine, vine* and see how the lower lip almost touches the upper front teeth for the first sound of each word.
(3) **Interdentals** are produced by inserting the tip of the tongue between the upper and lower teeth. English has two interdental sounds and they are both represented by the *th* in words such as *thigh, thy*.
(4) **Alveolars** are sounds made by raising the front part of the tongue to the alveolar ridge. Say the words *tee, dee, see, zee, knee* and notice how the tongue either touches or almost touches the alveolar ridge for the first part of each word. The [l] is called **lateral** liquid because the tongue is raised to the alveolar ridge with the sides of the tongue down, permitting air to escape laterally over the sides of the tongue. Hold the tongue in position in the last part of the word *tell* while taking in a breath through the mouth. You should feel the incoming air cooling the sides of the tongue. The English [r] is what's called a **retroflex** liquid, produced by curling the tip of the tongue back behind the alveolar ridge.
(5) **Post-Alveolars** (or, sometimes called **alveopalatals**) are produced by raising the front part of the tongue to a point on the hard palate just behind the alveolar ridge. The sound [ʃ] occurs in the beginning of words such as *shy, ship* and its voiced counterpart, [ʒ], occurs in the middle of words such as *measure, pleasure*. [ʧ] and [ʤ], also post-alveolar sounds, occur in the beginning of *cheese* and *geez* respectively.
(6) **Palatals** are produced with the tongue on or near the **palate**, the highest part of the roof of the mouth. The beginning sound in the words *yes, you* is a palatal sound.
(7) **Velars** are sounds made with the tongue on or near the velum, the soft area toward the rear of the roof of the mouth. Velars are the final sounds in words such as *tack, tag, tang*.

(8) **Glottals** are produced by using the **vocal cords** as primary articulators with no other modification of the airstream in the mouth. If the air is stopped at the glottis by tightly closed vocal cords, the sound is called a glottal stop, or [ʔ]. This sound occurs in the exclamation *uh-oh!*, or is used in place of the [t] sound in words such as *kitten, mutton*. The beginning sound in words such as *hi, hen, here* is also a glottal, but air passes through the open glottis.

2.4.2 Manners of Articulation

Manner of articulation refers to the way the airstream is modified by the vocal tract to produce sounds. The lips, tongue, velum, and glottis can be positioned in different ways to produce different sounds.

(1) **Stops** are made when there is a complete closure of the articulators so that air is stopped. There are two possible types of stops. If the air is stopped in the oral cavity but the velum is lowered so that air escapes through the nose, the sound produced is a **nasal stop**. In contrast, if the velum is not lowered, the air cannot escape through the nose, resulting in an **oral stop**. To see how nasal stops are different from oral stops, hold your nose while saying *my nanny*. It will sound something like *bye daddy*. This is because [m] and [b] are produced in the same way by stopping the airflow at the lips. The only difference is that for [m], the velum is lowered to allow the air to go out the nose. Holding your nose while saying [m] prevents the air from escaping through the nose, making it sound more like its oral counterpart, [b]. The same can be said for [n] and [d], which are produced in exactly the same way by having the front part of the tongue touch the alveolar ridge. Holding your nose while saying [n] prevents air from escaping through the nose and makes this nasal stop sound more like its oral counterpart, [d]. The same principle applies to [ŋ] and [g]. Although both the nasal stop and the oral stop can be classified as stops, the term **stop** by itself is typically used to indicate an oral stop, while the term **nasal** is used to indicate a nasal stop.
(2) **Fricatives** are produced when the airstream is partially obstructed to allow air to continuously flow through the mouth. When fricatives are produced, the air passes through a very narrow opening either at the glottis or in the vocal tract, resulting in a continuous turbulent noise, or frication. The beginning consonants in *face, vase*, *thigh, thy, sue, zoo,* and *ship*, as well as the consonants in the middle of the words *measure* and *aha*, are examples of fricative sounds.
(3) **Affricates** are a combination of two manners of articulation: a stop and a fricative. They are made by a stop closure followed immediately by a slight release of the articulators so that turbulent noise is produced. The beginning consonants in *chewed* and *Jude* are examples of the two affricates in English.

(4) **Nasals** are produced by lowering the velum and opening the nasal passage to the vocal tract. As explained earlier, nasals are sometimes called nasal stops because, just like the oral stops, there is complete obstruction of the airstream in the oral cavity. The only difference between nasal and oral stops is that for nasal stops, the velum is lowered to allow the air to escape through the nasal cavity. English has nasals in three different places of articulation: bilabial, alveolar, and velar.

(5) **Liquids** are consonants that involve substantial constriction of the vocal tract but the constriction is not sufficiently narrow to block the vocal tract or cause friction. There are mainly two types of liquids in English, namely, [l] and [r]. For the lateral liquid [l], the center of the vocal tract is completely obstructed, as in a stop, but air escapes through the sides of the tongue. Try saying the word *hill* and pause your tongue at the [l], then inhale sharply. The sides of your tongue should feel cool with the incoming air. Liquids are typically voiced in English. However, liquids become voiceless after voiceless stops, as in the English words *plead* [pl̥id] and *clay* [kl̥ɛj]. The voiceless lateral is written with an additional phonetic symbol, called a **diacritic**. In this case, the diacritic is a small circle beneath the [l] and denotes voicelessness on a sound that is otherwise voiced.

The other liquid in English is [r]. This sound is made either by curling the tongue tip back into the mouth or by bunching the tongue upward and back in the mouth. This r, which is known as retroflex r, is heard in words such as *rate* and *far* in American English. In IPA, the symbol [r] is reserved for a trilled r, as in Spanish *perro* "dog". The IPA transcribes the retroflex r as [ɹ], but in this book, we will use the symbol [r] for the retroflex r. In some languages, the r is produced by a single tap or a flap of the tongue against the alveolar ridge. In IPA, the flap is transcribed as [ɾ]. In Spanish both the trilled and tapped r occur as in *perro* "dog" and *pero* "but". Some speakers of British English pronounce the r in the word *very* with [ɾ]. Most American English speakers use [ɾ] instead of a [t] or [d] in words like *little* [lɪɾl̩], and *ladder* [lærər].

(6) **Glides** are sounds that are made with little or no obstruction of the airstream in the mouth. [w] is made by both rounding the lips and simultaneously raising the back of the tongue toward the velum. For that reason, [w] is called a labio-velar glide. [w̥] is made just like [w], except that it is voiceless. Some English speakers use this sound to distinguish the pronunciation of the word *which* [w̥ɪʧ] from that of *witch* [wɪʧ]. [j] is a palatal glide, which is made by raising the blade of the tongue toward the hard palate as if producing the vowel sound [i]. When occurring in a word, glides must always be either preceded or followed by a vowel. In articulating [w] or [j], the tongue rapidly glides toward or away from a neighboring vowel, hence the term **glide**. For example, in producing *yes* [yɛs], the tongue moves rapidly from the [j] to the [ɛ] vowel. Glides are transition sounds that are sometimes called **semivowels**.

2.5 Consonants in Other Languages

So far we have reviewed consonants in English. In this section, we will consider consonants that are in other languages. Learning to accurately produce these non-English sounds can be challenging but with practice one can achieve reasonable approximations of native pronunciations.

When we examine the IPA consonant chart (found at the top of the inside back cover of this book), we notice that there are three places of articulation that we have not considered so far: the retroflex, uvular, and pharyngeal sounds. Retroflex sounds do not occur in most varieties of English except for those spoken in India. They are produced by curling the tip of the tongue back behind the alveolar ridge. Thus the voiceless retroflex fricative [ʂ] is produced by sliding the tongue back and curling up the tip while saying [s]. You will notice that [ʂ] sounds something like [ʃ]. Similarly, the retroflex stops [ʈ] and [ɖ] are made by curling up the tongue tip while producing [t] and [d]. The curling of the tongue tip gives retroflex sounds an r-like quality.

Uvular sounds are made by raising the back of the tongue toward the uvula. The voiced uvular fricative [ʁ] is used in French words such as *roi* [ʁwa] "king" and *rapide* [ʁapid] "fast". The voiceless uvular fricative [χ] occurs in French after voiceless stops, as in *mètre* [mɛtχ] "meter". The voiceless uvular stop [q] is pronounced like [k], except that the tongue makes contact with the uvula. It occurs in the word [qɑzɑq] "Kazakh" in the Kazakh language. Uvular stops [q] and [ɢ] and uvular nasal [ɴ] occur in some of the indigenous languages of the Americas such as Inuktitut and Quechua.

For most English speakers, pharyngeal sounds are fairly difficult to make. The IPA lists two fricatives in the pharyngeal region, [ħ] and [ʕ], which occur in Semitic languages such as Arabic and Hebrew. Pharyngeal fricatives are made by pulling the tongue root toward the back wall of the pharynx. They occur in words like [ħadiqa] "garden" and [muħammad] "Mohamed" and [ʕarab] "Arab" in Arabic.

Other than uvular and pharyngeal sounds, the majority of non-English sounds that are found in other languages involve different manners of articulation at the same places of articulation as in English. For example, English has bilabial stops and nasals, but no bilabial fricatives. Acoustically, the voiceless and voiced bilabial fricatives, [ɸ] and [β], sound like labiodental fricatives, [f] and [v], but they are made by bringing the lips almost together so that there is a narrow opening between them. In the West African language, Ewe, the use of bilabial or labiodental fricatives results in meaning differences in word pairs such as [èβè] "Ewe" (the name of the language) and [èvè] "two". Similarly, Ewe contrasts [ɸ] with [f] in [éɸá] "he polished" and [éfá] "he was cold".

The voiceless palatal fricative [ç] occurs in German in words such as [ɪç] "I" and [nɪçt] "not" and is acoustically similar to the first part of the English word [hju] "hue". The voiceless velar fricative [x], as in the German words [axt] "eight" and [naxt] "night", is made by building up pressure behind the velar closure without completely stopping the airflow as one would when producing

[k]. Spanish has velar fricatives [x] and [ɣ] in words such as [xaβon] "soap" and [xuɣar] "to play".

Voices From the Classroom 2.1—How Knowledge of Students' Native Languages Can Help in Teaching Them English Pronunciation

I noticed that many of my Arabic-speaking ESL students had difficulty pronouncing [ʒ] as in the word *plea<u>s</u>ure* [plɛʒər]. But once I realized that Arabic has this sound in words like *<u>J</u>ordan* which is pronounced [ʒordan], I was able to point out to the students that the *s* in *pleasure* is the same sound as the beginning sound in *Jordan*.

Chris McKinnon, Adult ESL Instructor

2.6 Vowels

When we produce vowel sounds, none of the articulators come very close together, and the passage of the airstream is relatively unobstructed. Any given vowel is described by using the following four features: (1) the height of the tongue; (2) the front-back position of the tongue; (3) the degree of lip rounding; and (4) tense-lax gesture of the tongue.

To get a feel for differences in tongue height, place a lollipop on your tongue and say the following words with front vowels: *eat, ate, at*. You will notice that it is not very easy to say *eat* with the lollipop in your mouth, because the lollipop gets in the way of the tongue trying to bunch up toward the roof of the mouth. But as you produce the mid vowel in the word *ate*, the tongue is not as bunched up high in the mouth, leaving more room for the lollipop. It becomes easier still to produce the low vowel in the word *at* because the tongue is completely lowered.

You can try the same thing for the back vowels. With the lollipop on your tongue, say the following words: *ooze, o's, ah's*. You will notice that it is not difficult to say these words with the lollipop in your mouth because all of these vowels require the back of the tongue to move up and down in the mouth. Since the lollipop sits on the front part of the tongue, it does not interfere in the production of the back vowels. Even so, you should feel that it gets progressively easier to say the words as the tongue is progressively lowered.

Figure 2.2 shows the position of the tongue in the words *beet, bait, bet, bat, bot, bought, boat, boot*. In the first four vowels /i, e, ɛ, æ/, the highest point of the tongue is in the front of the mouth. Therefore, these vowels are called **front**

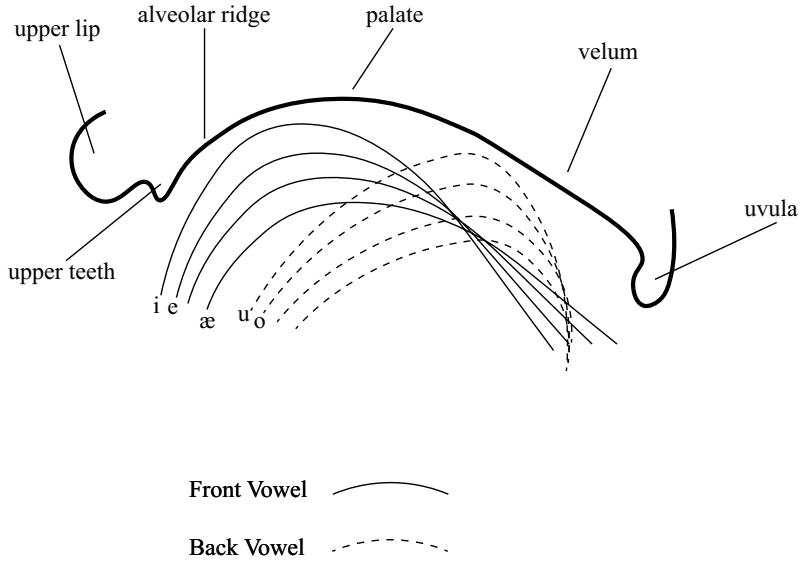

Figure 2.2 The Positions of the Tongue for the Vowels in "Beet, Bait, Bet, Bat, Bot, Bought, Boat, Boot"

vowels. The tongue is highest in the mouth for the vowel in *beet*, slightly lower for the vowel in *bait*, and lower still for the vowels in *bet* and *bat*. The vowel in *beet* is called a **high** front vowel, while the vowel in bat is called a **low** front vowel. The height of the tongue for the vowels in *bait* and *bet* is between these two extremes, and they are called **mid** front vowels. In the four **back** vowels /ɑ, ɔ, o, u/, the tongue is close to the back surface of the vocal tract. The tongue is highest in the vowel in *boot* and lowest in the vowel in *bot*. Thus the vowel in *boot* is called a high back vowel, while the vowel in *bot* is called a low back vowel. The vowels in *boat* and *bought* are classified as mid back vowels. In addition to the front and back vowels, English has a mid central vowel, which is written with the symbol [ə] to denote the vowel in the first syllable of *about* [əbawt], or with the symbol [ʌ] to denote the vowel in *but* [bʌt]. [ə] and [ʌ] are essentially the same vowel except that [ə] is used to denote a mid central vowel in unstressed syllables while [ʌ] is reserved for stressed (or only) syllables.

Figure 2.3 shows the vowels in English. Notice that the tongue positions in Figure 2.2 roughly correspond to the placement of the symbols in the vowel chart in Figure 2.3. Although the vowel space is neatly divided into high, mid, and low for tongue height, and front, central, and back for placement of the tongue in the front or back of the mouth, it is important to note that there are really no distinct boundaries between one type of vowel and another. Vowels are much more fluid and continuous than consonants. See this for yourself by saying [i] as in *he* and move slowly to [æ] as in *had*. You will notice that you pass through sounds similar to [ɛ] in *head* and [e] in *hey*. It is perfectly possible

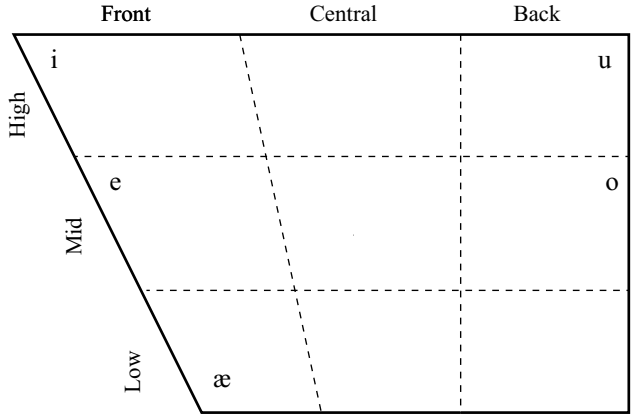

Figure 2.3 The Vowels in English

to make a vowel that is halfway between a mid vowel and a low vowel. In fact, it is possible to make a vowel at any spot on the vowel chart by making ever so slight changes in the shape and placement of the tongue and rounding of the lips. Consonants, on the other hand, are much more distinct from one another. Consider two different manners of articulation—a fricative and a liquid. A sound may be a fricative or a liquid, but it cannot be halfway between the two.

In addition to the height and front-back position of the tongue, vowel quality also depends on the shape of the lips. Look at your lips in a mirror while you say just the vowels in the words *boot, put, boat, bought*. You will notice that your lips are **rounded**. Now say the remaining vowels in the vowel chart in Figure 2.3 while looking in a mirror. You will notice that your lips are **unrounded**.

Finally, the vowels in English can be grouped into **tense** and **lax** vowels. Tense vowels generally require more extreme tongue gesture to reach the outer edges of the vowel space than lax vowels. There are four tense/lax vowel pairs that make a meaning difference in English words: (1) the [i, ɪ] pair as in *beet* and *bit*; (2) the [e, ɛ] pair as in *bait* and *bet;* (3) [u, ʊ] pair as in *boot* and *put*; and (4) the [o, ɔ] pair as in *boat* and *bought*. The difference between the vowels in *bait* and *bet* is that the vowel in *bet* is shorter, lower, and slightly more centralized than the vowel in *bait*. We call the vowel in *bet* a mid front lax vowel, and the vowel in *bait* a mid front tense vowel. Likewise, the lax vowels [ɪ] and [ʊ] are shorter, lower, and slightly more centralized than their tense counterparts [i] and [u].

In summary, the vowels in English can be described using the four aforementioned factors as follows:

[i], as in *beet*, is a high, front, unrounded, tense vowel.
[ɪ], as in *bit*, is a high, front, unrounded, lax vowel.

[e], as in *bait*, is a mid, front, unrounded, tense vowel.
[ɛ], as in *bet*, is a mid, front, unrounded, lax vowel.
[æ], as in *bat*, is a low, front, unrounded, lax vowel.
[u], as in *boot*, is a high, back, rounded, tense vowel.
[ʊ], as in *put*, is a high, back, rounded, lax vowel.
[o], as in *boat*, is a mid, back, rounded, tense vowel.
[ɔ], as in *bought*, is a mid, back, rounded, lax vowel.
[ɑ], as in *bot*, is a low, back, unrounded, lax vowel.
[ə], as in the first syllable of *about*, and [ʌ], as in *but*, is a mid, central, unrounded, lax vowel.

Voices From the Classroom 2.2—Using the Color Vowel Chart

Many of my students have had problems with various vowel sounds in English. Teaching the phonetic symbols was more time consuming than helpful for my students. Now, I use the Color Vowel Chart (https://americanenglish.state.gov/resources/color-vowel-chart) in class to help students remember the vowel sounds. The chart assigns a color and a common object to a vowel sound and helps students recognize the many vowel sounds in English as well as the multitude of vowel sounds for one letter. For instance, if students are challenged by the pronunciation of the word *refugee* (a common mispronunciation in my experience), students would remember *red dress, blue moon*, and *green tea*. These help young students to learn the correct pronunciation of a new word.

Morgan Nixon, ESL Teacher

2.6.1 Diphthongs

Diphthongs can be described as movements from one vowel to another. The first part of the diphthong is usually more prominent than the second, which is often transitory. The tense vowels in American English, [i], [e], [u], and [o], are phonetically diphthongs. This is why in some transcription systems, these vowels are transcribed as [ij], [ej], [uw], and [ow] respectively. Most other languages use "pure" vowels instead of diphthongs for these sounds. As such, English speakers should learn to hear the glides in these tense vowels and avoid them as they learn to speak other languages, unless they are specifically called for. For example, many American English speakers typically pronounce the Spanish word *peso* as /pejsow/. But if the goal is to sound more like a native Spanish speaker, one should use pure vowels and say /peso/. In addition

to [ij], [ej], [uw], and [ow], English has other diphthongs such as [aw], as in *brow*, [ɔj], as in *boy*, [aj], as in *buy*.

2.6.2 Phonics vs. Phonetics—What's the Difference?

At this point, it may be helpful to consider a question often asked by language arts teachers, namely, "what is the difference between phonics and phonetics?" Phonics is a teaching tool used to help children who already speak English to learn to read in English. As explained earlier, English orthography has a highly irregular spelling. While variability in spelling is found in both consonants and vowels, there is far greater variability when it comes to writing vowel sounds than consonant sounds. For example, the [o] sound can be written in at least six different ways (oval, nose, boat, bow, toe, though). Then there are words containing the letter o that sound like [ɑ] in American English (box, mop, hot, frog). For children who are just beginning to read, the label "short o" is helpful in associating the [ɑ] sound in words like box, mop, hot, and frog with the letter o written on the page.

Thus, by grouping vowels into "long" and "short" categories, phonics helps children recognize letter patterns in written text. As can be seen in Table 2.3, for each of the five vowels, a, e, i, o, and u, there is a long version and a short version (long e, short e, long a, short a, and so on). Notice that the words *beet* and *bet* which have the "long e" and "short e" sounds respectively, have e's in them. Similarly, *bait* and *bat* which have the "long a" and "short a" sounds respectively, have a's in them. The "long" and "short" characterizations roughly correspond to tense and lax vowels respectively. Thus, "long e", or the sound transcribed in IPA as [i], is a tense vowel while "short e", transcribed in IPA as [ɛ], is a lax vowel.

One may ask, "if the sound [i] corresponds to the long *e* sound, why is it represented with the letter *i*?" This is because in most languages that use the

Table 2.3 Phonics vs. Phonetics

What the vowel is called in phonics	*Words in English with the vowel*	*How the vowel is transcribed using phonetic symbols*
long e	beet	[i]
short e	bet	[ɛ]
long a	bait	[e]
short a	bat	[æ]
long i	bite	[aj]
short i	bit	[ɪ]
long o	boat	[o]
short o	bot	[ɑ]
long u	butte	[ju]
short u	but	[ʌ]

Roman alphabet, the long *e* sound is written with the letter *i*. For example, in Spanish, the letter *i* in words such as *picante* ("spicy") and *dinero* ("money") are pronounced with the long *e* or [i] sound. Similarly, the letter *i* in words such as *midi* ("noon") and *piscine* ("swimming pool") in French are pronounced with the long *e* or [i] sound.

A phonetic transcription, on the other hand, is a more accurate description of individual sounds, with each phonetic symbol representing only one sound. Additionally, when we look at the phonetic chart, it provides information about how and where to produce the sound. In order for teachers to correctly identify pronunciation difficulties of ESL students, they need to have an accurate and reliable way to refer to individual sounds. For students who are learning English for the first time as adolescents or adults, the "long" and "short" vowel characterizations will not make much sense since they do not already know how to speak English.

2.7 Vowels in Other Languages

Compared to many of the world's languages, English has a relatively large vowel inventory with eleven monophthongs in stressed syllables (check out Vowel Quality Inventories in the World Atlas of Language Structures Online at: http://wals.info/chapter/2). The five-vowel system of Spanish is much more common among the world's languages. Hundreds of languages have only five or six contrasting vowels (Japanese, Mandarin, and Swahili). It should be noted, however, a given symbol on these charts does not have the same value for all the languages. As we reviewed earlier in this chapter, tense vowels in English behave more like diphthongs. In contrast, the five vowels in Spanish are more pure vowels. Figure 2.4 shows the vowel inventories of English, Spanish, and French.

French has three front rounded vowels [y, ø, œ]. The high front rounded vowel [y] is produced by saying [i] with the lips rounded. To make this sound, begin by making the [i] sound, then slowly bring your lips to a rounded position while keeping the tongue position for [i]. The [y] sound is used in French words such as *tu* [ty] ("you" informal) and *vu* [vy] ("saw"), and contrasts with the high back vowel [u] in *tous* [tu] ("all") and *vous* [vu] ("you" formal). Similarly, the mid front rounded vowel [ø] is produced with a tongue position as in [e], but with rounded lips.

Earlier in this chapter, we saw that English has two kinds of stops: oral stops [p, b, t, d, k, g] and nasal stops [m, n, ŋ]. Vowels, on the other hand, are most often oral sounds. But vowels may be nasalized if the velum is lowered to allow part of the airstream to flow through the nose. Nasalization commonly occurs in vowels next to nasal consonants (more on this in the next chapter), but it can be a contrastive property in some languages like French, Hindi, and Navajo. In French, for instance, using an oral or a nasal vowel results in different words, as in *mais* [mɛ] ("but") and *main* [mɛ̃] ("hand"), and *beau* [bo] ("beautiful") and *bon* [bõ] ("good"). The diacritic [˜] placed

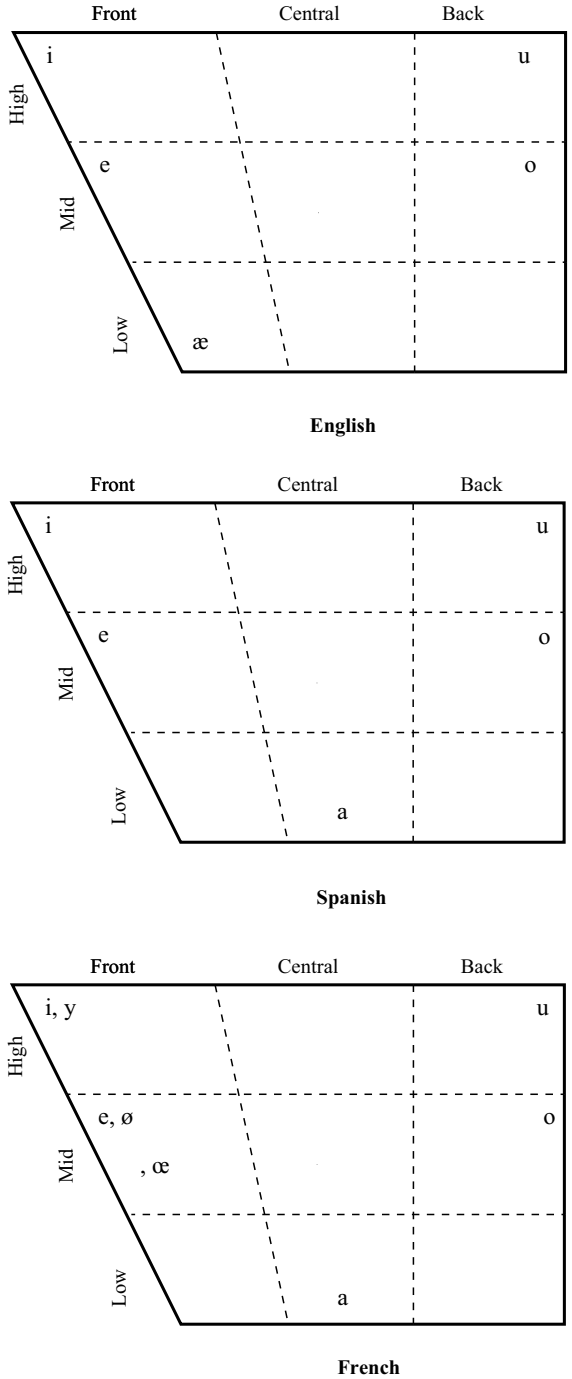

Figure 2.4 Vowels in English, Spanish, and French

above a vowel indicates that it is nasalized. A nasalized vowel is almost identical to its oral counterpart except for the fact that the velum is lowered to allow some of the air to escape through the nose. Although vowels can be nasalized in English, English does not contrast between oral vowels and their nasal counterparts.

Voices From the Classroom 2.3—Focusing on Pronunciation Errors That Interfere With Comprehension

I have come to believe that learners' objectives must be considered when developing a strategy to teaching language, with a high level of attention being paid to their personal and professional needs—whether there is a need for near-native speaking ability, or if comprehensibility is sufficient for their objectives. While my early approaches involved attempts to assist students in "overcoming" their L1 speech patterns (e.g., Spanish speakers saying *especial, eschool*), I now recognize that the time and effort to do so may be better directed towards elements of speech that promote increased comprehensibility (e.g., helping students correctly pronounce [ɪ] and [i] in *bit/beet, fit/feet*). Where comprehensibility is the critical target, for example, the ability to properly form a plural [s] or [z]; past [d] or [t]; or a flap versus the direct phonetic version of a spoken word might be more important than to spend valuable instructional time on some other, less obstructive, feature of a student's spoken language.

Steven Wagoner, Adult ESL Teacher

2.8 Suprasegmentals

So far we have reviewed the characteristics of individual sounds. We have seen that consonants can be described in terms of voicing, and place and manner of articulation, while vowels can be described in terms of tongue height and advancement, lip rounding, and tenseness of the tongue gesture. Consonants and vowels can be thought of as the **segments** that make up speech. Over and above these segments are other features known as **suprasegmentals**. These include variations in **length**, **stress**, and **pitch**. Improving pronunciation in a second language requires learning not only the features of individual consonants and vowels but also the suprasegmental features that extend across multiple segments in an utterance.

2.8.1 Length

The first suprasegmental feature we will consider is **length**. In many languages, differences in the duration of segments can be meaningful. For example, Japanese contrasts between short and long vowels in words like [ojisan] "uncle" and [oji:san] "grandfather", as well as [obasan] "aunt" and [oba:san] "grandmother". The colon [:] placed after a vowel indicates that it is lengthened. In Thai, vowel duration is linked to differences in meaning in words such as [tak] "to dip up" versus [ta:k] "to dry", and [khut] "to dig" versus [khu:t] "to scrape". Some languages contrast short and long consonants. In Arabic, consonant lengthening results in meaning difference in [qabala] "he accepted" versus [qabbala] "he kissed". In Finnish, [kuka] means "who" while [kukka] means "flower". The doubled consonant in these examples reflects lengthening.

2.8.2 Stress

Consonants and vowels together form syllables, which make up utterances. In any utterance, some syllables are perceived to be more prominent than others and are said to have greater **stress**. Stressed syllables are generally longer, louder, and higher in pitch than unstressed syllables. In English, stress can have a grammatical function. For example, variations in stress are used to distinguish between a noun and a verb in English. When you say the words *(an) ínsert* versus *(to) insért, (a) cóntrast* versus *(to) contrást*, or *(a) présent* versus *(to) presént*, you will notice that the stress is on the first syllable in the nouns whereas it is on the last syllable in the verbs. Stress can also be used to emphasize certain words in expressions like *She's wearing a pínk dress (not a blue dress)*, as opposed to *She's wearing a pink dréss (not a suit)*.

English uses several levels of stress in multisyllabic words such as *démocràt, demócracỳ, dèmocrátic, dèmocrátically*. An acute accent [ˊ] placed over the vowel marks the most prominent or **primary stress** whereas a grave accent [ˋ] marks the second most prominent or **secondary stress**. In IPA, [ˈ] is placed before the beginning of a syllable with primary stress, while [ˌ] is placed before the syllable with secondary stress. Thus, the same words would be transcribed as [ˈdɛməˌkræt], [dəˈmɑkrəˌsi], [ˌdɛməˈkrærɪk], [ˌdɛməˈkrærɪkli].

Placement of stress is predictable and relatively straightforward in some languages. In French, for example, stress is placed on the final syllable of a word. In Spanish, stress is placed on the penultimate (second to last) syllable in words that end in a vowel or *n* or *s* (e.g., *náda, ésta, éstas, cántan*) and on the final syllable in words that end in a consonant other than *n* or *s* (e.g., *hablár, libertád, comér*). However, placement of stress varies according to the word in English, which poses significant problems for those trying to learn it as a second language. Although there are some rules, they are rather complicated and unwieldy. Nonetheless there are some useful ways to teach English stress patterns to second language learners (see Voices from the Classroom 2.4—Teaching English Stress Patterns).

Voices From the Classroom 2.4—Teaching English Stress Patterns

- Raise awareness and build confidence

  Some learners love to learn about the "technical" side of language, while others like to "feel" or "see" the language more, hearing the music of word stress or seeing the shapes of the words. Try to use a variety of approaches: helping students to engage with English in different ways will help them in their goal to become more proficient users of the language.

- Mark the stress

  Use a clear easy-to-see way of marking stress on the board and on handouts for students.

- Integrate word stress into your lessons

  You don't need to teach separate lessons on word stress. Instead, you can integrate it into your normal lessons. The ideal time to focus students' attention on it is when introducing vocabulary. Quickly and simply elicit the stress pattern of the word from the students (as you would the meaning) and mark it on the board. Drill it, too!

- Troubleshooting

  Initially, many students (and teachers!) find it difficult to hear word stress. A useful strategy is to focus on one word putting the stress on its different syllables in turn. For example: com-pu**ter** com**pu**ter

- Say the word in the different ways for the students, really exaggerating the stressed syllable and compressing the unstressed ones. Ask the students which version of the word sounds "the best" or "the most natural." By hearing the word stressed incorrectly, students can more easily pick out the correct version.

Excerpted from a web article written by Emma Pathare, Teacher, Trainer, Dubai (Retrieved June 8, 2016 from: www.teachingenglish.org.uk/article/word-stress)

2.8.3 Pitch

Pitch is the auditory property of a sound that can be placed on a scale that ranges from low to high. It correlates to fundamental frequency and depends on the rate of vibration of the vocal folds. We can vary pitch by controlling the tension of the vocal folds and the amount of air that passes through the glottis.

While more tense vocal folds and greater air pressure lead to higher pitch, less tense vocal folds and lower air pressure result in lower pitch. There are mainly two types of controlled pitch movement that are found in the world's languages: (1) **tone** and (2) **intonation**.

Pitch differences that change the meaning of a word are called **tones**. Languages that exploit this feature are called **tonal languages**. They include many of the languages spoken in Southeast Asia (e.g., Burmese, Thai, Vietnamese), all varieties of Chinese, African languages such as Zulu, Luganda, and Igbo, as well as many of the Native languages of the Americas (e.g., Navajo, Cherokee, Zuni). For example, in Mandarin Chinese four different tones can combine with the syllable *ma* to produce different words: (1) *mā* (high level) "mother"; (2) *má* (high rising) "hemp"; (3) *mǎ* (low falling rising) "horse"; (4) *mà* (high falling) "scold". Similarly, Thai has five tones that result in differences in meaning: (1) /kʰāː/ (mid) "stick"; (2) /kʰàː/ (low) "galangal—an Asian plant of the ginger family"; (3) /kʰâː/ (falling) "value"; (4) /kʰáː/ (high) "to trade"; (5) /kʰǎː/ (rising) "leg".

Both Mandarin Chinese and Thai have **contour tones**, which can be described in terms of shifts in pitch rather than single points within a pitch range. The falling and rising tones in the preceding Mandarin Chinese and Thai examples illustrate contour tones. In contrast, **register tones** are level tones that describe specific points within a pitch range. The distinguishing feature in languages with register tones is the relative difference between the pitches such as high, mid, or low, not the shape of their movements. Luganda, Zulu, and Hausa are examples of register tone languages with two tones—high and low. Yoruba is an example of a tonal language with three tones—high, mid, and low. In a tonal language, it is not the absolute pitch of the syllables that is important but the relative pitch among different syllables. A tone is perceived as high if it is high relative to other pitches around it produced by the same speaker.

Intonation refers to pitch movement that is not related to differences in word meaning. Thus, in a **non-tonal language** such as English, whether one pronounces the word *sandwich* with a rising pitch or a falling pitch makes no difference to the meaning of the word. But intonation can be used to convey broader linguistic meaning in both tonal and non-tonal languages. For example, the falling pitch we hear at the end of a statement in English such as *He had a sandwich* signals that the utterance is complete. But a rising pitch at the end of the same utterance would signal that the speaker is asking a question (*He had a sandwich?*). If the word *sandwich* were part of a list, it would be pronounced with an intonation that first falls and then rises again slightly. Thus in the statement *He had a sandwich, a bag of chips, and a large soda* the words *sandwich* and *chips* would be marked by a rising intonation to indicate that the speaker is not finished speaking while the sentence-final word *soda* is pronounced with a falling intonation.

2.9 Syllable-Timed vs. Stress-Timed Languages

The world's languages can be classified into two main categories: **syllable-timed** and **stress-timed**. In syllable-timed languages, the more syllables

there are, the more time it takes to say something because each syllable is approximately the same length as other syllables. Examples of syllable-timed languages include Spanish, French, Korean, and Japanese. English is a stress-timed language, where syllables are not similar in length. A stress-timed language is a language where the stressed syllables are said at approximately regular intervals, and unstressed syllables shorten to fit this rhythm. Oral speech in English is rhythmic and moves from one stressed syllable to the next stressed syllable. Everything in between is shortened to fit the rhythm set by stressed syllables. To see how this works, use a metronome (you can access a free metronome online from sites such as www.metronomeonline.com/). Set the metronome to a comfortable walking tempo (around 80 to 90 beats per minute) and read the following sentences one by one, making sure that each stressed syllable falls *on* the beat and unstressed syllables fall *in between* the beats. Notice that, as more and more syllables are added, it still takes the same amount of time to produce each sentence. Each sentence is said in exactly three beats, which correspond to the number of stressed syllables in each sentence.

Kíds pláy gámes. (3 syllables)
The kíds pláy gámes. (4 syllables)
The kíds pláy the gámes. (5 syllables)
The kíds will pláy the gámes. (6 syllables)
The kíds will be pláying the gámes. (7 syllables)

If English were a syllable-timed language, it would take more time to produce the longer sentences. English learners whose native language is a syllable-timed language have a tendency to assign equal length to each syllable in English, whether stressed or not. These learners would do well to distinguish between stressed and unstressed syllables and learn to reduce the vowels in unstressed syllables. The reverse is true for English speakers who are learning a syllable-timed language like Spanish and French. English speakers may set the metronome to the syllable level and learn to pronounce each syllable with full vowels.

2.10 The Importance of the [ə] in Improving Pronunciation in English

When words in English are produced in isolation, at least one syllable is stressed. In case of words that are composed of one syllable, that syllable will be stressed. Thus, each of the words in the sentence *She said that to you and me*, when produced in isolation, would sound something like this:

[ʃi], [sɛd], [ðæt], [tu], [ju], [ænd], [mi]

However, when the words are said in quick succession in conversation, they undergo considerable change. Some words will be completely unstressed, the

vowel may be reduced to [ə] or may disappear altogether, and one or more of the consonants may be dropped or altered. Thus, *She said that to you and me* in connected speech might sound like: [ʃɪ ˈsɛd ðət tə ˈju ən ˈmi]. Notice that the vowels in unstressed words are reduced and the final consonant in the word "and" is dropped.

The mid central vowel in English, [ə] (also known as *schwa*), is critical in English because vowels in unstressed syllables are often reduced to it. Let's see how placement of stress changes vowel quality. Consider the words *emphátic* [ɛmˈfæɬik] and *émphasis* [ˈɛmfəsɪs]. Both have three syllables. However, notice that the stress falls on the second syllable in *emphatic* whereas it falls on the first syllable in *emphasis*. The low front vowel [æ] is reduced to a mid central vowel [ə] in an unstressed vowel. Teachers sometimes hear English learners pronounce the second vowel in *emphasis* as [æ]. To help these students, teachers might point out that the stress falls on the first syllable in *emphasis* and that the vowel in the unstressed second syllable is reduced to a [ə].

To help students achieve more natural-sounding and less stilted pronunciation in English, teachers can review the different ways that some of the monosyllabic function words in English are pronounced in stressed or unstressed positions (see Table 2.4).

Since English is a stress-timed language, English speakers tend to pronounce full vowels only when they are in stressed syllables. The remaining syllables are reduced, usually to [ə]. However, English speakers learning other languages will want to produce each vowel as a full vowel when speaking other languages.

Table 2.4 How Functions Words Are Pronounced Differently in Stressed vs. Unstressed Positions

Function Word	*Stressed*	*Unstressed*	*Example in Conversation*
and	ænd	ənd, ən, nd, n̩	ˈTom and ˈJerry [ˈtam n̩ ˈʤɛri]
as	æz	əz	as ˈfast as light [əz ˈfæst əz lait]
at	æt	ət	ˈsee you at ˈnoon [ˈsi ju ət ˈnun]
can	kæn	kən	I can ˈdig. [aɪ kən ˈdɪg]
has	hæz	həz, əz	ˈHe has done it. [ˈhi əz dən ɪt]
he	hi	hɪ, ɪ	Will he ˈlook? [wɪl ɪ ˈlʊk]
to	tu	tə	You ˈhave to ˈtell her. [ju ˈhæv tə ˈtɛl ər]

Recommended Websites

Interactive IPA Chart with Audio Recordings
https://www.internationalphoneticassociation.org/IPAcharts/inter_chart_2018/IPA_2018.html
The World Atlas of Language Structures Online
http://wals.info/
Type IPA Phonetic Symbols
https://ipa.typeit.org/

Further Reading

Ashby, P. (2011). *Understanding phonetics*. Oxon, UK: Hodder Education.
Collins, B., & Mees, I. M. (2013). *Practical phonetics and phonology: A resource book for students* (3rd ed.). London and New York: Routledge.
Ladefoged, P., & Johnson, K. (2015). *A course in phonetics* (7th ed.). Stamford, CT: Cengage Learning.

Exercises

1. Using the words from the Word Bank, fill in the names of the vocal organs in Figure 2.5.

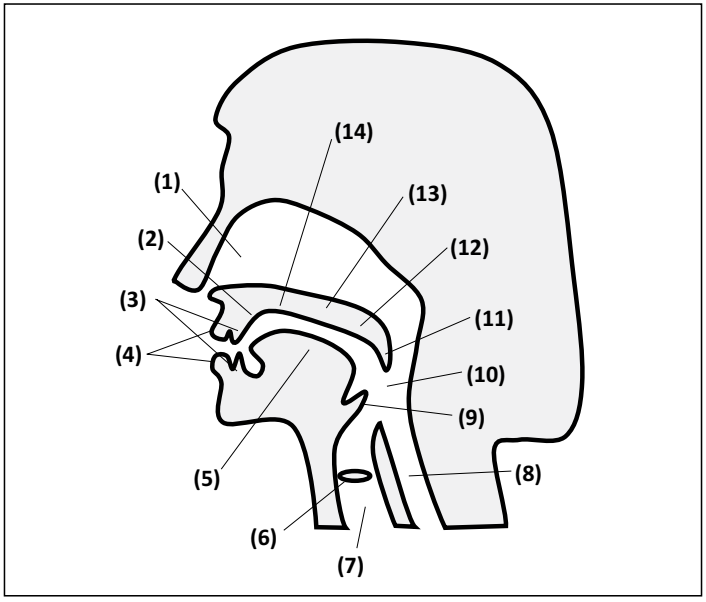

Figure 2.5 The Vocal Tract

Word Bank
nasal cavity, uvula, teeth, lips, velum (soft palate), palate, tongue, glottis, alveolar ridge, esophagus, epiglottis, alveopalatal, pharynx, trachea

2. Following the examples below, provide three-part description of each consonant, starting with voicing, place of articulation, and manner of articulation in that order.

   a) /p/—voiceless bilabial stop
   b) /g/—voiced velar stop
   c) /f/—voiceless labiodental fricative
   d) /s/—
   e) /n/—
   f) /l/—
   g) /θ/—
   h) /ʤ/—

3. Transcribe the following words:

   a) image [ ]
   b) choice [ ]
   c) grief [ ]
   d) advise [ ]
   e) advice [ ]
   f) remain [ ]
   g) pursue [ ]

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

   a) mothball [maðbɔl] should be [maθbɔl]
   b) wives [wajvs] should be [ ]
   c) recommend [rɛcəmɛnd] should be [ ]
   d) treasure [trɛzər] should be [ ]
   e) shipping [ʃɪppɪŋ] should be [ ]
   f) conniving [kənnajvɪŋ] should be [ ]
   g) resounding [rəsawndɪŋ] should be [ ]

5. Find errors in the transcription of the vowel sounds in the following words. In each word there is one error, indicating an impossible pronunciation of

that word for a native speaker of English of any variety. Make a correct transcription in the space provided after the word.

a) minerals [minərəlz] should be [mɪnərəlz]
b) fishing [fiʃɪŋ] should be [ ]
c) regular [rɛgulər] should be [ ]
d) biennial [bajænɪəl] should be [ ]
e) couldn't [kudnt] should be [ ]
f) programming [programɪŋ] should be [ ]
g) football [futbɔl] should be [ ]

Reference

IPA Chart. (2015). Retrieved from http://www.internationalphoneticassociation.org/content/ipa-chart, available under a Creative Commons Attribution-Sharealike 3.0 Unported License. Copyright © 2015 International Phonetic Association.

3 Phonology

The Patterning of Sounds

3.1 Introduction

Phonology is the study of how sounds pattern in a language. While phonetics is concerned with the physical production and perception of speech sounds, phonology describes the ways in which sounds function within a language to encode meaning. Phonological knowledge helps us to produce sounds that form meaningful utterances. For example, it helps us to choose the appropriate sound of the plural *-s* morpheme to turn singular nouns into plural nouns in English (e.g., [s] for *cup<u>s</u>*, [z] for *spoon<u>s</u>*, and [ɪz] for *dish<u>es</u>*), or the appropriate sound of the past tense *-ed* morpheme to form past tense verbs (e.g., [t] for *pass<u>ed</u>*, [d] for *reason<u>ed</u>*, and [ɪd] for *last<u>ed</u>*). For native speakers of English, knowing which sound to use in the formation of the plural noun or the past tense verb is largely subconscious (it is not so apparent for nonnative speakers of English, and this chapter will show how teachers can help students gain this knowledge). Phonology is about describing this subconscious knowledge that speakers have of their language.

In this chapter, we will learn that speakers and listeners of any language pay attention to contrastive sounds that signal a difference in meaning (i.e., phonemes) and ignore phonetically conditioned variation between non-contrastive sounds (i.e., allophones). We will see that what constitutes a phoneme is language-specific. For example, the difference between the vowels [i] and [ɪ] is crucial to English, as we can see from minimal pairs like *keen* [kin] and *kin* [kɪn]. But in Korean, this difference in pronunciation is not distinctive. A Korean speaker may pronounce the word for "long" as [kin] or [kɪn], and it will make no difference to the meaning.

When learning a second language, people naturally resort to their knowledge of their native language phonology. Thus, a Korean speaker may not always make a distinction between the vowels [i] and [ɪ] when pronouncing English words. This explains why many people speak a second language with an accent. We will review some strategies teachers can use to help students learn second language phonology and improve their pronunciation.

3.2 What Does It Mean When We Say We Know a Word?

Knowing a word means knowing both its sounds and its meaning. Most of the words in a language differ both in sounds and meaning. Consider the following words in English:

pat [pæt]	teal [til]	thigh [θaɪ]
bat [bæt]	deal [dil]	thy [ðaɪ]

Each word differs from the other words in both its sounds and meaning. Sometimes two words differ by just one sound. For example, the sounds of *pat* and *bat* are identical except for the initial consonant sounds. The *p* and *b* sounds can therefore distinguish or contrast words, and are called phonemes. **Phonemes** of a given language are heard by its native speakers as distinct sounds that result in a difference in meaning.

We see from the contrast between *teal* and *deal* and between *thigh* and *thy* that [t], [d], [θ], [ð] are also phonemes in English for the same reason—substituting a [d] for [t] or a [θ] for [ð] produces a different word. The difference between *pat* and *bat*, between *teal* and *deal*, and between *thigh* and *thy* are minimal (i.e., they are identical except for one sound that occurs in the same place in the word). For that reason, they are called **minimal pairs**. We can come up with more words that differ with *pat* and *bat* in the beginning consonant such as *mat, sat, fat*, and *vat*. These words form a **minimal set** with *pat* and *bat*. Likewise, we can form a minimal set by substituting the vowel sound in *pat*, as in *pet*, *pit*, *pot*, and *put*. We can form yet a different minimal set by substituting the final consonant in *pat*, as in *pad*, *pal*, *path*, and *pack*.

3.3 Phonemes and Allophones

In our discussion of voiceless stops in Chapter 2, we did not make a distinction between the *p* sound in the word *pit* from the *p* sound in the word *spit*. But try pronouncing these two words while holding a strip of paper in front of your lips and you will see (and hear) a clear difference between the two *p* sounds. When you say *pit*, a puff of air will push the paper. The paper will not move when you say *spit*. We call the *p* in *pit* **aspirated** and transcribe it using a small raised [ʰ] after the [p]. The *p* in *spit* is **unaspirated** and is transcribed with a regular [p]. Thus, *pit* is transcribed as [pʰɪt] and *spit* is transcribed as [spɪt].

After the release of the aspirated *p* in *pit* [pʰɪt], there is a brief delay before the start of voicing of the following vowel, [ɪ]. This lag in the onset of voicing is accompanied by release of air, giving rise to the puff of air (aspiration). In contrast, voicing of the vowel starts very soon after release of the

unaspirated *p* in *spit* [spɪt]. There is a clear audible difference between the aspirated *p* and the unaspirated *p*. However, speakers of English typically perceive the [pʰ] in *pit* and the [p] in *spit* to be the same *p* sound. Why is this? This is because the difference between the two sounds is predictable and their occurrence can be determined by specific phonetic environments. Furthermore, changing one sound for the other will not change the meaning of the word.

To understand this point, let us examine some data. Table 3.1 lists examples of aspirated and unaspirated consonants in monosyllabic words (words made up of a single syllable). Notice that the sounds that have both aspirated and unaspirated varieties are all voiceless stops. The *t* in *tuck* and the *k* in *cone* are also aspirated voiceless stops, while the *t* in *stuck* and the *k* in *scone* are unaspirated voiceless stops.

Looking at Table 3.1, you will notice that all the aspirated stops occur at the beginning of the word while the unaspirated stops do not. But what happens when the word has more than one syllable? How do voiceless stops behave in words that have multiple syllables?

Let us consider some multisyllabic words in English such as *contain* [kən.tʰéɪn], *repeal* [rɪ.pʰíl], *articulate* [ar.tʰí.kjə.lɛɪt], *respect* [rɪ.spɛ́kt], *compass* [kʰám.pəs], and *marker* [már.kər] (note that the "." marks syllable boundaries while an acute accent [´] marks the primary stress in the word). When we examine these words, we notice that all the aspirated stops occur at the beginning of a syllable with a stressed vowel. The unaspirated stop, on the other hand, occurs in non-syllable-initial positions and in unstressed syllables. This observation leads us to the following phonological rule:

> The phonemes /p/, /t/, and /k/ become [pʰ], [tʰ], and [kʰ] respectively when they occur at the beginning of a stressed (or, only) syllable. In non-syllable-initial positions or in non-stressed syllables, /p/, /t/, and /k/ are uttered as [p], [t], and [k] respectively.

Notice that where the aspirated *p* occurs, the unaspirated *p* does not, and where the unaspirated *p* occurs, the aspirated *p* does not. This mutual exclusivity makes it possible for us to determine when to use the aspirated *p* and when to use the unaspirated *p*. We call the aspirated *p* and the unaspirated *p*

Table 3.1 Monosyllabic English Words With Aspirated and Unaspirated Consonants

Aspirated	*Unaspirated*
paid [pʰɛɪd]	spade [spɛɪd]
tuck [tʰʌk]	stuck [stʌk]
cone [kʰown]	scone [skown]

allophones of the /p/ phoneme. **Allophones** are predictable variants of one phoneme. In other words, the choice of which allophone to use is not random but can be predicted from examining the phonetic environment. Notice in the aforementioned rule that slashes / / are used to enclose phonemes and square brackets [ ] are used for allophones. The phoneme, /p/, is the representation of the sound in the mind of the English speaker while the allophones, [pʰ] and [p], are how that mental representation is produced in actual speech depending on the phonetic environment.

While native speakers of English know when to produce aspirated voiceless stops and when to produce unaspirated voiceless stops, they are not explicitly taught these rules. Rather, they acquire these rules subconsciously as children. But people who are learning English as a second language do not have this knowledge and need to learn the rule explicitly if they want to pronounce voiceless stops appropriately. This is where the ESL teacher would come in. By teaching the aspiration rule in English voiceless stops, teachers can help their students improve their pronunciation.

It is important to remember that phonological rules are language-specific. While the aspiration rule applies to English voiceless stops, the same rule does not apply to voiceless stops in other languages. In English, the aspirated *p* and the unaspirated *p* are allophones of the same phoneme, /p/. Therefore, if an ESL student were to pronounce *paid* as [pɛɪd] rather than [pʰɛɪd], it would sound a little "strange" but it would not result in a word with a different meaning. However, in languages such as Hindi or Korean, aspirated and unaspirated stops are separate phonemes (i.e., the use of one or the other results in difference in meaning). In these languages, making the distinction between aspirated and unaspirated stops is crucial because if you use the wrong sound, you may wind up saying something entirely different by mistake. Korean makes a three-way distinction in voiceless stops:

[t̰am] "sweat" (unaspirated, tense)
[tam] "fence" (slightly aspirated, lax)
[tʰam] "greed" (heavily aspirated)

This minimal triplet shows that [t̰], [t], and [tʰ] are separate phonemes in Korean that result in a difference in meaning depending on which one is used.

English speakers who are learning other languages should remember that while syllable-initial /p/, /t/, and /k/ are pronounced with strong aspiration in American English as [pʰ], [tʰ], and [kʰ], they are not produced with aspiration in many other languages. Those who wish to sound more "native-like" in languages like Spanish, for example, should pronounce the unaspirated versions of these consonants and use them only. Thus, Spanish words such as *casa* "house", *pero* "but", and *tostada* "toast" with voiceless stops in word-initial position should be pronounced as [kasa], [pɛro], [tostaða], and not as [kʰasa], [pʰɛro], and [tʰostʰada].

Voices From the Classroom 3.1 — Knowledge of Other Languages Comes in Handy When Teaching English Pronunciation

My knowledge of students' native languages helps me understand some of my students' pronunciation issues. In Arabic, [p] and [b] are allophones so Arabic speakers often say *labtob* for *laptop*. I tell my students both [p] and [b] are made the same way with the lips and mouth but that [b] vibrates in the throat. I have them put their fingers on the throat and feel the vibration as they practice saying the sounds [p] and [b] and then have them practice the sounds in words. I like to use mirrors and discuss mouth shapes when practicing pronunciation, especially for vowels. Lax vowels are rare in other languages so I have students watch as they make the [i] and [ɪ] with their mouth opening wider for [i]. Predicting what issues beginners will have helps me understand what they are trying to say. Likewise, warning students that certain sounds are hard for speakers of certain languages alerts them that they should try to focus on these sounds.

Hilary Reintges, Adult ESL Teacher

3.4 Natural Classes

In generating phonological rules, it is helpful to refer to the concept of natural classes. A **natural class** is a group of sounds in a language that share one or more articulatory or auditory property, to the exclusion of other sounds in that language. For example, consonants constitute a natural class to the exclusion of vowels. Similarly, voiced sounds form a natural class to the exclusion of voiceless sounds. The same principle applies to labial sounds vs. non-labial sounds, nasals vs. non-nasals, front vowels vs. non-front vowels, etc.

When trying to identify natural classes for consonants, look for what a group of consonants have in common in terms of voicing, place of articulation, and manner of articulation. When identifying natural classes for vowels, look for what the vowels have in common in terms of tongue height, front/back, tenseness, and lip rounding.

Natural classes are important because it is easier to learn a rule that applies to a whole class of sounds than to individual sounds. For example, in the previously discussed English aspiration rule, we saw that a natural class—namely, voiceless stops, /p/, /t/, and /k/—are affected by this phonological rule. Voiced stops and other classes of sounds are not affected by this rule. Natural classes help us to understand how sounds pattern in a language by allowing us to efficiently and economically describe how a whole group of sounds are affected by a variety of phonological processes.

Let us illustrate the significance of natural classes by way of two examples: Canadian Raising and Vowel Length in English.

3.4.1 Canadian Raising

Canadian Raising is a phonological rule in many dialects of North American English that changes the pronunciation of diphthongs, [aɪ] and [aʊ], to [ʌɪ] and [ʌʊ] respectively. It gets its name from a low vowel, [a], being raised to a mid vowel, [ʌ]. Speakers of these dialects would say *out and about* as [ʌʊt ən əbʌʊt] rather than [aʊt ən əbaʊt], and *bright light* as [brʌɪt lʌɪt] rather than [braɪt laɪt]. However, the pronunciation changes do not happen across the board to all words containing these diphthongs. Some words with these diphthongs retain the [aɪ] and [aʊ] pronunciation as in other dialects of English.

Our job as students of phonology is to determine when the change in the pronunciation occurs and when it does not. In other words, is the change in pronunciation governed by a phonological rule? If so, what is that rule? To answer this question, we need to examine words that are pronounced with [aɪ] and words that are pronounced with [ʌɪ] in these dialects of North American English.

Consider the following data. Notice that the left column has words that are pronounced with the diphthong [aɪ], and the right column, with the diphthong [ʌɪ].

[aɪ]	*[ʌɪ]*
rise [raɪz]	bite [bʌɪt]
time [taɪm]	bike [bʌɪk]
file [faɪl]	life [lʌɪf]
ninth [naɪnθ]	type [tʌɪp]

Based on these data, we can list the phonetic environments of the two diphthongs as follows. The "phonetic environment" typically means the sound that is immediately to the left and the sound that is immediately to the right of the sound in question. Notice that we use a dash (—) to stand for the sounds in question, [aɪ] and [ʌɪ]:

[aɪ]	*[ʌɪ]*
r—**z**	b—**t**
t—**m**	b—**k**
f—**l**	l—**f**
n—**n**	t—**p**

Look at the sounds that occur before [aɪ] in the left column and see if they form a natural class. In other words, look for what they have in common in

terms of voicing, place of articulation, and manner of articulation. Then, look at the sounds to the right of [aɪ] and see what they have in common in terms of voicing, place of articulation, and manner of articulation. Now, look at the sounds to the left of [ʌɪ] and see if you can identify any natural classes. Finally, repeat the same process for the sounds to the right of [ʌɪ].

When we do this, we notice that the sounds that are immediately to the right of [aɪ] are all voiced consonants whereas the sounds that are immediately to the right of [ʌɪ] are all voiceless consonants. Voiced consonants and voiceless consonants are mutually exclusive categories (i.e., they form natural classes) and we can therefore say that the distribution of [aɪ] and [ʌɪ] is determined by the phonetic environment. In other words, we can predict when the diphthong /aɪ/ will be pronounced as [aɪ] and when it will be pronounced as [ʌɪ]. So, we can write a phonological rule that captures this fact as follows:

(1) [aɪ] occurs before a voiced consonant whereas [ʌɪ] occurs before a voiceless consonant.

We can also write the phonological rule in (1) in a different way, as (2):

(2) /aɪ/ becomes [ʌɪ] before a voiceless consonant.

Notice that (1) and (2) are the same rule written in different ways. While (1) shows that [aɪ] and [ʌɪ] are allophones of the same phoneme, (2) shows that /aɪ/ is the phonemic form of the diphthong, which is pronounced as [ʌɪ] before a voiceless consonant and [aɪ] before a voiced consonant.

Let us now consider words that contain [aʊ] and [ʌʊ] in dialects with Canadian Raising:

[aʊ]	*[ʌʊ]*
howl [haʊl]	about [əbʌʊt]
browse [braʊz]	house [hʌʊs]
cloud [klaʊd]	south [sʌʊθ]
pound [paʊnd]	clout [klʌʊt]

We can list the phonetic environments of these words as follows:

[aʊ]	*[ʌʊ]*
h—**l**	b—**t**
r—**z**	h—**s**
l—**d**	s—**θ**
p—**n**	l—**t**

Looking at the phonetic environments, we notice that the same rule applies to [aʊ] and [ʌʊ]—namely, that [aʊ] occurs before a voiced consonant whereas [ʌʊ] occurs before a voiceless consonant.

This analysis shows that the natural classes of voiced consonants vs. voiceless consonants elegantly describe the distribution of the two types of diphthongs in some dialects of North American English. The mutually exclusive nature of voiced consonants vs. voiceless consonants helps us predict when we will get the raised diphthongs, [ʌɪ] and [ʌʊ], and when we will not.

3.4.2 Vowel Length in English

Now, let us consider another example of allophonic variation—that of vowel length. The colon [ː] placed after [ɪ] in the words in the right column means that the vowel is long.

[ɪ]	*[ɪː]*
hiss [hɪs]	his [hɪːz]
rich [rɪtʃ]	ridge [rɪːdʒ]
trick [trɪk]	trig [trɪːg]
hit [hɪt]	hid [hɪːd]
slip [slɪp]	slim [slɪːm]

We can list the phonetic environments of the two sounds in question as follows:

[ɪ]	*[ɪː]*
h—**s**	h—**z**
r—**tʃ**	r—**dʒ**
r—**k**	r—**g**
h—**t**	h—**d**
l—**p**	l—**m**

When we consider the phonetic environments of [ɪ] and [ɪː], we notice that the sounds that are immediately to the right of [ɪ] are all voiceless consonants whereas the sounds immediately to the right of [ɪː] are all voiced consonants. Therefore, we can say that the distribution of [ɪ] and [ɪː] is phonologically conditioned. In other words, we can predict when the phoneme /ɪ/, will be pronounced as [ɪ] and when it will be pronounced as [ɪː]. So, we can write a phonological rule that captures this fact as follows:

(3) [ɪ] occurs before a voiceless consonant while [ɪː] occurs before a voiced consonant.

We can also write the phonological rule in (3) in a different way, as (4):

(4) /ɪ/ becomes [ɪː] before a voiced consonant.

Notice that (4) shows that /ɪ/ is the phonemic form of the vowel, which is uttered as [ɪː] before a voiced consonant and [ɪ] before a voiceless consonant. Therefore, [ɪ] and [ɪː] are allophones of the phoneme /ɪ/. As explained earlier, the phoneme /ɪ/ is the representation of the sound in the mind of the English speaker while the allophones [ɪ] and [ɪː] are how that mental representation is produced in actual speech depending on the phonetic environment.

Vowel length is allophonic in most dialects of English. That is, a monosyllabic word that ends in a voiceless consonant will be shorter than the same word that ends in a voiced consonant. You can see this in a minimal pair such as *pat* and *pad*. Try saying these two words and you will notice that the vowel [æ] in [pæd] is slightly longer than the vowel [æ] in [pæt]. The same rule applies to *neat* and *need*, to *but* and *bud*, and even to diphthongs like *write* and *ride*, and *mouth* (noun) and *mouth* (verb). Words that end in a voiced consonant have vowels that are slightly longer than those that end in a voiceless consonant.

The vowel lengthening rule applies to the speech of all speakers of English regardless of style or rate of speaking. While the length differences are ever so small and not very easy to notice, they are nonetheless an important part of a native accent. Not applying the rule of vowel lengthening would make someone sound like a nonnative speaker of English. Also, keep in mind that vowel length here is used differently than the notion of long vowels and short vowels in phonics (see Chapter 2); vowel length here is literally the time each vowel lasts, while phonics uses long and short vowels to refer to two different vowels in English, like the long i (as in *bi̲te*) and the short i (as in *bi̲t*).

More importantly, not getting the right vowel length can interfere with communication. For example, the final stops, [t] and [d], in *pat* and *pad* are often unreleased. That is, the final stop has no audible release. Unreleased stops are marked by a raised [˺] after the stop, as in [pæt˺] and [pæːd˺]. We see instances of unreleased stops in the first consonant of clusters in words such as *doctor* [dak˺tər] and *kept* [kɛp˺t].

When the final stop in words like *pat* and *pad* is unreleased, the length of the preceding vowel plays a crucial role in communicating to the listener which word was meant by the speaker. The slightly longer vowel in [pæːd˺] will signal to the listener that the word *pad* was meant by the speaker and not the word *pat* [pæt˺]. Without an audible release in the final stop, the only clue that the listener has is the length of the preceding vowel.

Not getting the right vowel length is a common pronunciation problem for many ESL students. Few students know that there is a difference between the vowels in *pat* and *pad*. Here again, the ESL teacher can help students produce the right vowel length by explicitly teaching the vowel lengthening rule in English. Teachers can do exercises involving minimal pairs (e.g., cap/cab, heat/heed, muck/mug, ate/aid, bright/bride). By exaggerating vowel length

differences in minimal pairs such as these, teachers can sensitize students to accurately perceive and produce words in English.

3.5 Phonological Processes

Phonological rules may be obligatory or optional. **Obligatory rules** apply in the speech of all speakers of a language or dialect having the rule, regardless of style or rate of speaking. Canadian Raising and Vowel Length in English, which we discussed previously, are obligatory rules that apply in the speech of all speakers of those dialects. On the other hand, the unreleased final stop is an optional rule that sometimes applies in casual speech. **Optional rules** may or may not apply in any given utterance and are responsible for variation in speech.

Another optional rule in English is **deletion** in unstressed syllables. Deletion rules eliminate a sound, frequently in unstressed syllables and in casual speech. Deletion is common in fast speech because it saves time and effort for the speaker. So, the sentence *I want him to see this* may be said either as [aɪ wənt hɪm tə si ðɪs] in careful speech, or as [aɪ wən ɪm tə si ðɪs] in more casual speech. The voiceless stop, /t/, after a nasal at the end of the word *want* is often deleted, especially when the final stop is unreleased. Sounds like /h/ that are not very noticeable are often deleted because speakers can save time and effort by eliminating them without sacrificing meaning.

In addition, an unstressed vowel [ə], is often deleted when the next syllable is stressed in fast speech, as shown in Table 3.2.

As deleting sounds can make speech production easier in these examples, **insertion** of sounds can accomplish the same goal in others. Insertion is especially helpful when producing two neighboring sounds that have very different phonetic features. For example, the words *warmth, tenth*, and *length* are typically pronounced [wɔrmθ], [tɛnθ], and [lɛŋθ]. In casual speech, however, speakers may insert a [p] between the [m] and the [θ] and pronounce the word [wɔrmpθ]. Similarly, a [t] may be inserted between [n] and [θ] to produce [tɛntθ], and a [k] may be inserted between [ŋ] and [θ] to produce [lɛŋkθ].

There is an articulatory explanation to this process. The insertion of a voiceless stop consonant makes it easier for the articulators to transition from a voiced nasal to a voiceless fricative in these words. Table 3.3 describes [m] and [θ] in terms of voicing, place, and manner of articulation. Notice that these

Table 3.2 Deletion of Vowel [ə] in Unstressed Syllables

	Slow speech	*Fast speech*
corruption	[kʰərʌ́pʃən]	[kr̥ʌ́pʃən]
ferocious	[fəróʊʃəs]	[fr̥̥óʊʃəs]
ballistic	[bəlístɪk]	[blístɪk]

Note: The [̥] placed below [r] in [kr̥ʌ́pʃən] and [fr̥̥óʊʃəs] denotes devoicing.

two sounds have no common feature—they are very different sounds. Articulatorily speaking, it is not very easy to go from [m] to [θ] in quick succession.

Now look at Table 3.4, which lists the features for [m], [p], and [θ]. Notice that as the articulators go from [m] to [p] to [θ], there is overlapping place and voicing that the inserted sound, [p], has in common with the preceding and the following sound. In other words, the transition is more gradual than if the articulators were to go directly from [m] to [θ].

Aside from those explained so far, there are a few phonological processes that are very common in the world's languages.

One of the most common phonological processes found in many of the world's languages is assimilation. Quite simply, **assimilation** causes a sound to become more like a neighboring sound in terms of one or more of its phonetic characteristics. Assimilation occurs mainly for ease of articulation; that is, assimilation makes it easier to move the articulators to produce different sounds consecutively in fluent speech.

In Table 3.2, we saw that [r], a typically voiced sound, became voiceless after voiceless consonants, [k] and [f], in fast speech. This process is called **devoicing**. Liquids and glides after voiceless consonants in English words such as *play* [pl̥ɛɪ], *prod* [pr̥ad], and *cure* [kj̊ur] are often devoiced. Devoicing is a kind of assimilation because the lack of voicing in the [p] and [k] sounds spreads to [l], [r], and [j], making these normally voiced sounds voiceless. Speaking in terms of articulatory phonetics, devoicing happens because the vocal folds do not start vibrating immediately after the release of the voiceless consonant closure.

Table 3.3 [m] and [θ] Described in Terms of Voicing, Place, and Manner

	Voicing	*Place*	*Manner*
[m]	Voiced	Bilabial	Nasal
[θ]	Voiceless	Interdental	Fricative

Table 3.4 [m], [p], and [θ] Described in Terms of Voicing, Place, and Manner

	Voicing	*Place*	*Manner*
[m]	Voiced	**Bilabial**	Nasal
[p]	**Voiceless**	**Bilabial**	Stop
[θ]	**Voiceless**	Interdental	Fricative

Voices From the Classroom 3.2—Teaching the Different Pronunciations of the *-ed* Ending

When teaching past tense, I use phonology to teach the different pronunciations of the *-ed* ending. For this lesson, I explain the difference between voiced and voiceless sounds, asking students to touch their throats as we go through different sounds together so they can feel their vocal cords vibrating or not. Then, in groups, they receive a stack of cards, each with three words with *-ed* on them (for example, one card might have *asked, filled*, and *wanted*). Together, students decide on the pronunciation for each word, label the words with /t/, /d/, or /ɪd/, and practice saying them. It's very rewarding, as a teacher, to walk around and see students touching their vocal cords, trying out the different pronunciations—and getting them right based on their new understanding of voiced and voiceless sounds.

Erica Ashton, ESL Teacher

Other than **voicing assimilation**, assimilation for place and manner of articulation is also widespread in the world's languages. An example of **place assimilation** is the pronunciation of the prefix *un-* in English. Words like *unbeatable, unjust*, and *uncut* are often pronounced [əmbiɾəbl̩], [əndʒʌst], and [əŋkʌt]. The nasal /n/ is often pronounced as a bilabial nasal, [m], before a bilabial sound, as in [ə**m**biɾəbl̩], and as a velar nasal, [ŋ], before a velar sound, as in [ə**ŋ**kʌt]. The nasal stays as an alveolar nasal, [n], before a post-alveolar affricate, [dʒ], in [ə**n**dʒʌst]. Thus, when an alveolar nasal immediately precedes a bilabial consonant, assimilation causes it to become a bilabial nasal. When it immediately precedes a velar consonant, the same process causes the alveolar nasal to become a velar nasal. In other words, the nasal is assimilated to the place of articulation of the following consonant.

An example of **manner assimilation** can be seen in nasalization of vowels before a nasal consonant, as shown in Table 3.5. Nasalization of a vowel before a nasal consonant is caused by speakers anticipating the lowering of the velum in preparation for a nasal sound. The result is that the preceding sound takes on the nasality of the following nasal consonant. Nasalized vowels occur only before nasal consonants in English syllables and are therefore predictable variants (allophones) of vowel phonemes.

However, as shown by minimal pairs in Table 3.6, nasal and oral vowels contrast in French, that is, they are separate phonemes in that language. The distribution of oral vs. nasal vowels is not determined by the phonetic environment since the oral vowels and their nasal counterparts occur in the same environment.

Table 3.5 Vowel Nasalization in English

Oral vowels	*Nasalized vowels*
lead [lid]	lean [lĩn]
sack [sæk]	sang [sæ̃ŋ]
face [fɛɪs]	fame [fɛ̃ɪm]

Table 3.6 How Oral and Nasal Vowels Contrast in French

Oral vowels	*Nasal vowels*
gras [grɑ] "fatty"	grand [grɑ̃] "tall"
beau [bo] "beautiful"	bon [bõ] "good"
paix [pɛ] "peace"	pain [pɛ̃] "bread"

The opposite of assimilation is **dissimilation**, in which a sound becomes less like a neighboring sound in terms of one of more of its phonetic characteristics. Compared to assimilation, dissimilation is a much rarer phonological process. However, when it does occur, it frequently serves the purpose of achieving clearer communication by breaking up sounds that are too similar. For example, in English, the word *fifths* [fɪfθs] ends in a consonant cluster made up of three consecutive fricatives. Some speakers pronounce this word as [fɪfts], breaking up the sequence of three fricatives with a stop.

Strengthening is a phonological process that makes sounds stronger. Aspiration of voiceless stops in English, discussed earlier in this chapter, is an example of strengthening. Aspirated stops are stronger sounds than unaspirated stops because the lag in the onset of voicing in aspirated stops is accompanied by a puff of air, which gives the sound a stronger quality.

The opposite of strengthening is **weakening**, a process by which sounds become weaker. **Flapping** is an example of weakening. Like a stop, a flap involves completely obstructing the oral cavity. But the duration of obstruction is much shorter than that of a stop. American English has an alveolar flap, [ɾ], which involves the tip of the tongue quickly striking the roof of the mouth and returning to its rest position. The flap occurs as the middle sound in the words, *butter* [bʌɾər], *writer* [raɪɾər], *ladder* [lærər], and *tidal* [taɪɾəl]. Notice that in these words, a *t* or a *d* sound becomes [ɾ] when it occurs between two vowels—specifically, after a stressed vowel and before an unstressed vowel. The flap is a weaker sound than [t] or [d] because it is shorter in duration and obstructs less air. Notice that when [t] changes to [ɾ] in these words, it involves voicing assimilation: [t], a voiceless sound, takes on the voicing of its surrounding vowels by becoming [ɾ], a voiced sound.

Finally, **metathesis** is a phonological process that reorders sounds. In many cases, sounds metathesize to make words easier to pronounce or easier to understand. In some dialects of English, for example, the word *ask* is pronounced

as *aks*. Some English-speaking adults pronounce *prescription, introduce*, and *cavalry* as *perscription, interduce*, and *calvary*. In each of these instances, metathesis facilitates the pronunciation of consonant cluster sequences.

3.6 Syllables

So far, we have discussed phonemes and allophones, and have seen that allophonic variation is often conditioned by neighboring sounds. We turn now to a different unit of phonological representation, namely, the syllable. Just as phonemes and allophones are part of speakers' knowledge, syllables are also part of speakers' knowledge about how sounds pattern in their language.

In any given language, a word is composed of one or more syllables, and a **syllable** is in turn composed of one or more phonemes. Every syllable has a **nucleus**, or a vowel. The nucleus may be preceded by one or more consonants called the **onset** and followed by one or more consonants called the **coda**. The nucleus and coda together are called the **rhyme**.

Figures 3.1 through 3.3 show the structures of monosyllabic English words, *ah, ark*, and *shark*. Notice that a syllable can be made up of just a vowel nucleus (as in *ah*), or nucleus and coda (as in *ark*), or onset, nucleus, and coda (as in *shark*).

Whenever native speakers of a language count syllables in a word, they demonstrate their knowledge of the syllable as a phonological unit. For example, English speakers intuitively know that the word *composition* [kam.pə.zɪ.ʃən] has four syllables because it has four vowels that serve as syllable nuclei (“.” marks syllable divisions). Children learn from an early age that the nucleus

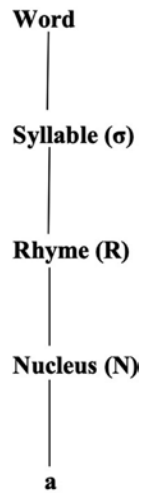

Figure 3.1 English Syllable Structure for *“Ah”*

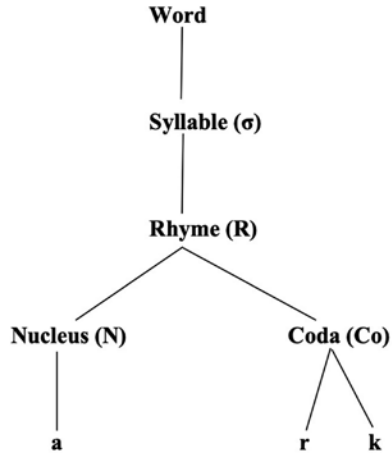

Figure 3.2 English Syllable Structure for *"Ark"*

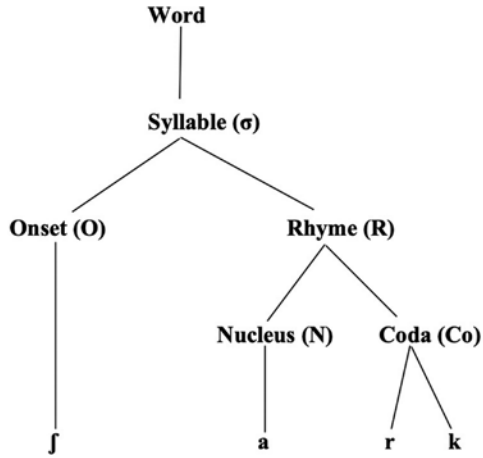

Figure 3.3 English Syllable Structure for *"Shark"*

and the coda of the final syllable in rhyming words are identical, as shown in the following nursery rhyme:

Humpty Dumpty sat on a w<u>all</u> /ɔl/
Humpty Dumpty had a great f<u>all</u> /ɔl/
All the king's horses and all the king's m<u>en</u> /ɛn/
Couldn't put Humpty together ag<u>ain</u> /ɛn/

Just as what constitutes a phoneme is language-specific, what constitutes a syllable is also language-specific. Each language has its own definition of what

may be considered a syllable in that language. Figure 3.4 shows the syllable structure of a monosyllabic word in English, *sprint*. The vowel, /ɪ/, constitutes the nucleus, while the remaining five sounds /s/, /p/, /r/, /n/, and /t/ are part of either the onset or the coda.

While *sprint* is a monosyllabic word in English, when the Japanese borrow the word, *sprint*, it is pronounced as *su.pu.rin.to* with four syllables (see Figure 3.5). This is because Japanese syllable structure does not allow consonant clusters in the onset or in the coda, as does English. Therefore, when borrowing words from other languages, Japanese speakers break up consonant clusters into separate consonants by inserting a vowel after each consonant. Notice that the only coda that is allowed in Japanese is /n/.

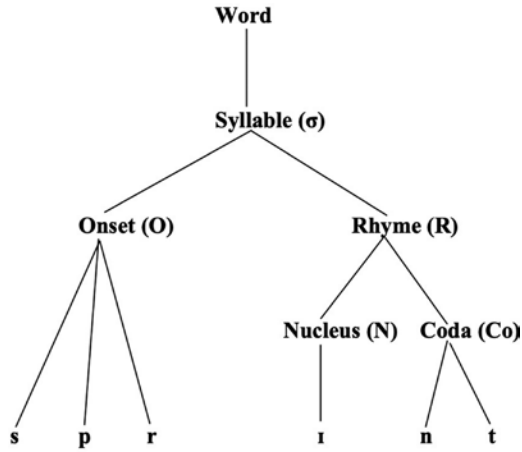

Figure 3.4 English Syllable Structure for *"Sprint"*

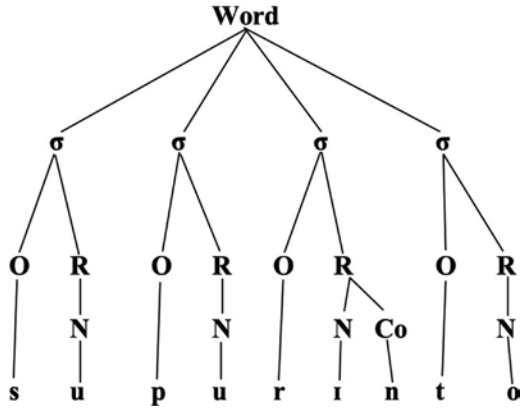

Figure 3.5 Japanese Syllable Structure for *"Sprint"*

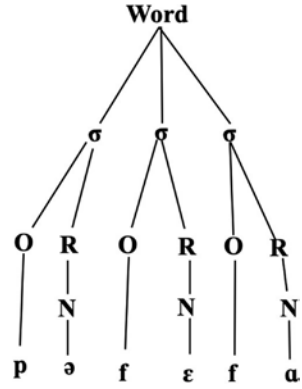

Figure 3.6 English Syllable Structure for "*Pfeffer*" [pəfɛfɑ]

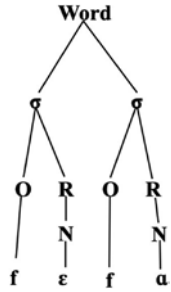

Figure 3.7 English Syllable Structures for "*Pfeffer*" [fɛfɑ]

Just as a Japanese-speaking learner of English would have difficulty pronouncing *sprint* as one syllable, an English speaker who is learning German would have difficulty pronouncing a word like *pfeffer* /pfɛfɑ/ "pepper" since the /pf/ sequence is not a permissible onset for a syllable in English. Thus, many English speakers would pronounce this German word as either [pəfɛfɑ] by inserting an extra vowel, or as [fɛfɑ] by deleting the initial consonant (see Figures 3.6 and 3.7).

3.7 How to Help Students Improve Their Pronunciation in a Second Language

When trying to help students improve their pronunciation in a second language, it is important to remember that second language pronunciation problems are often rooted in differences between what constitutes a phoneme or a syllable in the student's native language versus the target language. Consequently, it is helpful for teachers to become familiar with the phonology of the student's native language.

The Internet provides a wealth of information on the sound inventories of the world's languages. Teachers can do a quick search using key words such as *Spanish phonemic inventory, Arabic phonology*, or *Finnish vowels* and find articles written on these topics. Wikipedia provides an easy reference on the consonant and vowel inventories of many languages.

Let's see how a study of the phonemic inventory of students' native language can help teachers address students' pronunciation difficulties. A quick look at the Castilian Spanish consonant phoneme inventory (see Table 3.7) reveals that this language does not have /v/. Although the letter *v* is used in Spanish writing, the *v* sound does not exist in most dialects of Spanish. Therefore, you might hear Spanish speakers substitute the *v* sound with a *b* sound in English and say *bat* for *vat*, and *Boyd* for *void*. They may also substitute /v/ with /β/, a voiced labial fricative and say [sɛβən] for *seven*. [β] is an allophone of the phoneme /b/ that occurs between vowels.

Another observation from the Spanish consonant inventory is that it does not have a voiceless palatal fricative, /ʃ/, which suggests that Spanish speakers may have difficulty pronouncing words with /ʃ/ in English. In fact, many Spanish speakers will substitute /ʃ/ with /tʃ/, and pronounce *shock* as *chock* and *wish* as *witch*.

What can teachers do to help Spanish-speaking students accurately produce sounds like /v/ and /ʃ/ that do not exist in their native language? First, teachers can explain how the *v* sound is produced by showing how the upper teeth touch the lower lip and how the air passes through the narrow opening to produce the turbulent noise heard in the production of /v/. Another way to approach this is to draw students' attention to the fact that pronouncing /v/ involves simply adding voicing to /f/, which they already know how to produce. Teachers might have students place the tip of their fingers on their throat while saying /f/ and /v/ alternately. They can explain that both these sounds are produced in the same way in the mouth and that the only difference between /f/ and /v/ is the absence or presence of voicing.

As for /ʃ/, even though it is not a phoneme in many dialects of Spanish, Spanish speakers can nonetheless produce this sound. Teachers can have students practice saying "shhh" as if telling someone to be quiet. Then, students

Table 3.7 Castilian Spanish Consonant Phoneme Inventory

	Labial	*Dental*	*Alveolar*	*Palatal*	*Velar*
Stop	p b	t d		tʃ	k g
Fricative	f		s	ʝ	x
Nasal	m		n	ɲ	
Lateral			l		
Flap			ɾ		
Trill			r		

Source: Adapted from: Martínez-Celdrán et al., 2003

can add different vowel nuclei and coda to form words like *shhhock*, *shhhip, shhhine*, or add different onsets and vowels to form words like *wishhh, mashhh, lushhh*.

When addressing pronunciation problems, teachers will want to exaggerate the differences in articulation between the sounds they are trying to contrast, using slower tempo and repetition to help students hear and feel the differences. As discussed in Chapter 2, teachers may also use a variety of teaching tools such as lollipops and mirrors to help students see the position and shape of the articulators (lips, teeth, tongue).

English has a particularly large inventory of vowel phonemes compared with most other languages in the world. This causes problems for students whose native languages have far fewer vowels. For example, the tense/lax distinction found in English vowels does not exist in many other languages. Thus, many ESL students will pronounce *deep* and *dip* in the same way and *coat* and *caught* in the same way. One strategy that works well involves doing exercises with minimal pairs such as the following:

/i/	/ɪ/
seat	sit
heed	hid
green	grin
steal	still

After practicing with minimal pairs in isolation, teachers can have students create sentences using the words in minimal pairs (e.g., "When invited to a party, don't sit in the most important seat" or "The big green monster had a wide grin on his face") and have students practice reading them accurately with a partner.

Further Reading

Avery, P., & Ehrlich, S. (1992). *Teaching American English pronunciation*. Oxford: Oxford University Press.

Graham, C. (2001). *Jazz Chants old and new: Student book* (2nd ed.). Oxford: Oxford University Press.

Exercises

1. Transcribe the following words phonetically. Then circle all words that form a minimal pair with *sap* /sæp/.

lap	cop	tab	tap	hope	tan	tip	top	cone
tar	zap	sip	team	map	tow	rap	nine	tad

2. Flapping: When do /d/ and /t/ become [ɾ] in American English? Examine the following data and determine the phonological rule. Note that in the following transcriptions, an acute accent [ˊ] marks the primary stress in the word.

/t/ or /d/	*/ɾ/*
sandy [sǽndi]	pity [píri]
attention [əténʃən]	magnetic [mægnéɾɪk]
tainted [téɪntɪd]	beauty [bjúɾi]
atone [ətóʊn]	ladle [léɪɾl]
compete [kəmpít]	daughter [dɔ́ɾər]
disdain [dɪsdéɪn]	udder [ʌ́ɾər]

3. Consider the distribution of [s] and [ʃ] in Korean in the following words. [ɯ] is a high back unrounded vowel.

saram "person"	ʃidʒaŋ "market"
suhak "mathematics"	ʃihʌm "test"
sʌnsɛŋ "teacher"	ʃinpal "shoe"
sɯmul "twenty"	ʃilsɯp "training"
kasu "singer"	kaʃi "thorn"

Based on these data, are [s] and [ʃ] allophones of the same phoneme or allophones of different phonemes? If they are allophones of the same phoneme, state their distribution.

4. Consider the distribution of [r] and [l] in Korean in the following words and determine whether these two sounds are allophones of one phoneme or of different phonemes.

baram "wind"	palt^{h}op "toenail"
kirin "giraffe"	sʌnpul "advance payment"
noritʌ "playground"	kalmang "craving"
suri "repair"	p^{h}ilsu "necessary"
harapʌdʒi "grandfather"	halmʌni "grandmother"

If [r] and [l] are allophones of the same phoneme, which is the more basic form? How do you know?

5. Consider the distribution of [t], [ts], and [tʃ] in the following Japanese words and determine whether these three sounds are allophones of one phoneme or of different phonemes. Note that [ts] is an alveolar affricate and should be taken as a single symbol. Similarly, [tʃ] is an alveopalatal

affricate and should be taken as a single symbol. If the three sounds in question are allophones of the same phoneme, state the phonological rule.

tabemono “food”	tsukue “desk”
tʃigau “wrong”	todana “cupboard
tomodatʃi “friend”	tenki “weather”
totemo “very”	tsukuru “to make”
tʃikatetsu “subway”	tʃitʃi “father”
tokubetsu “special”	tsumori “plan”

6. Consider the distribution of [b], [d], [g], and [β], [ð], [ɣ] in the following Spanish data. [β] is a voiced bilabial fricative, [ð] is a voiced dental fricative, and [ɣ] is a voiced velar fricative. [b] and [β] are allophones of the same phoneme, as are [d] and [ð], as well as [g] and [ɣ]. State a rule that describes the conditioning environments for each pair.

[laðrar]	“to bark”	[kaða]	“each”	[doβle]	“double”
[boka]	“mouth”	[beβo]	“I drink”	[nombre]	“name”
[toðo]	“all”	[teŋgo]	“I have”	[tjenda]	“shop”
[graβaðor]	“recording”	[amiɣo]	“friend”	[siɣlo]	“century”

7. Consider the distribution of palatal affricates, [tʃ], [tʃʰ], [dʒ], and [dʒʱ], in the following data from Gujarati, a language spoken in the Indian state of Gujarat. [tʃʰ] is a voiceless aspirated palatal affricate and [dʒʱ] is a voiced aspirated palatal affricate. Determine if the four are allophones of the same phoneme or of separate phonemes. If they are allophones of the same phoneme, state a rule that describes the distribution. If they are allophones of separate phonemes, give a minimal set.

[tʃəl]	“walk”	[tʃidʒʱ]	“thing”	[dʒiɳɖgi]	“life”
[səmədʒ]	“understand”	[dʒəl]	“water”	[dʒʱəndə]	“flag”
[bidʒu]	“second”	[ɲĩtʃko]	“swing”	[dʒʱəl]	“glimmer”
[ekədʒ]	“one”	[tʃʰəl]	“deceit”	[tʃʰətri]	“umbrella”

Reference

Martínez-Celdrán, E., Fernández-Planas, A. M., & Carrera-Sabaté, J. (2003). Castilian Spanish. *Journal of the International Phonetic Association*, *33*(2), 255–259.