CHAPTER 2

Phonetics

FILE 2.1

Representing Speech Sounds

2.1.1 Studying Pronunciation

"You're not from around here, are you?" Sometimes you can tell by the way a person pronounces words that he or she speaks a dialect that is different from yours. For example, some people do not pronounce *pin* differently from *pen*. In some parts of Ohio the word *push* is pronounced with a vowel sound like the one in *who*. If you hear someone say *poosh* you can guess where they are from. Such pronunciation differences have been noted for many thousands of years. For example, there is a story in the Bible (Judges 12:4–6) about a group who used a password that their enemies could not pronounce correctly. The password they used was *shibboleth*, since their enemies couldn't say the <sh> sound. This group would kill anyone with the telltale pronounciation *sibboleth*. These illustrations show that pronunciation is a part of what we know when we know a language.

There are numerous ways of studying pronunciation in spoken language. In recent years, phoneticians have begun to employ some very sophisticated instrumental techniques to study spoken language.

In articulatory phonetics, we want to know the way in which speech sounds are produced—what parts of the mouth are used and in what sorts of configurations. To investigate these aspects of sound production, phoneticians have used **X-ray photography** and cinematography, among other techniques. More recently, to avoid methods that expose talkers to dangerous amounts of radiation, phoneticians have used point-tracking devices such as the X-ray microbeam or the electromagnetic articulograph to track the locations of small receptors glued onto the lips, tongue, and jaw. Articulatory phonetics is also done with **palatography** (see Section 2.2.6) to observe contact between the tongue and the roof of the mouth, and instruments to measure airflow and air pressure during speech.

In acoustic phonetics, we are more interested in the characteristics of the sounds produced by these articulations. To study acoustic phonetics, phoneticians use pictures of the sounds, using tools such as the **sound spectrograph**. These pictures help acoustic phoneticians explore the physical properites of sounds. These days, you can download sound editing and analysis software from the Web. Try searching for a "waveform editor" or an "audio spectrograph," or simply for "phonetics analysis software," and see if you find any free software to enable you to look at and edit speech sounds on your computer.

The third branch of phonetics, auditory phonetics, focuses on how humans process speech sounds: how we perceive pronunciation. While the fundamentals of perception can be explored by using fairly simple experimental methods that look at human responses to particular stimuli, advanced study of this field depends on more modern equipment such as magnetic resonance imaging (MRI) and computerized tomography (CT).

All of these techniques give us great insight into the details of phonetics. But the simplest and most basic method of phonetic analysis—**impressionistic phonetic transcription**—is still a vital tool for phoneticians. An example of phonetic transcription is the line "you say tomato, I say tomahto" from Ira Gershwin's lyrics to the song "Let's Call the Whole Thing Off." The word *tomato* is pronounced differently by different people, and we can sym-

bolize two of the pronunciations as "tomato" and "tomahto" as Gershwin did. Or we could follow the pronunciation guide in *Webster's Third New International Dictionary* and write the two pronuncations as tə'mātō and tə'måto. Or we could refer to the *American Heritage Dictionary,* where the two pronunciations are written təmā'tō and təmä'tō. Confusing, isn't it? Yet we need to use phonetic transcription because the normal spelling of the word doesn't tell us enough about how it is pronounced by different people.

Spelling	Gershwin	Webster's	Amer. Heritage
tomato	tomato	tə'mātō	təmā'tō
tomato	tomahto	tə'måtō	təmä'tō

2.1.2 The "Right" Phonetic Alphabet

Did Gershwin write the two pronunciations of *tomato* correctly? Or does one of the dictionaries have the right way to symbolize the difference? It should be clear that there is no one "right" answer about how to write pronunciation in a phonetic transcription. The choices we make are largely arbitrary or influenced by typographical or historical considerations. However, it is absolutely crucial that both the reader and the author agree on the sound qualities that are assigned to the symbols in a phonetic alphabet. This is why almost all dictionaries give some guide to the pronunciation symbols where they list familiar words as examples of the sounds. For example, *father* is used to illustrate the sound intended by <å> in *Webster's* and by <ä> in the *American Heritage.* Whether the <a> has one mark or two is an arbitrary decision. This is fine, so long as we have a pronunciation guide.

If the goal of having a phonetic transcription system is to be able to unambiguously convey the important aspects of the pronunciation of a given set of sounds, using a written system of symbols, then such a system must have certain characteristics.

First, each symbol should represent one sound (or **phone**) only, and there should be only one symbol for each sound. The letter <c> violates this principle in English spelling because it represents two sounds (the [k] sound in *cat,* and the [s] sound in *cymbal,* and both the [k] and [s] in *cynic,* for example). Hence using a <c> does not unambiguously tell the reader which sound is intended.

Second, if two sounds can distinguish one word from another, they should be represented by different symbols. The letters <th> in English violate this principle because the difference between the <th> sounds in *thy* and *thigh* is not captured by using <th> for both words. That is, there is an important difference in pronunciation that is not captured with these letters.

Third, if two sounds are very similar and their difference arises only from the context they are in, we should be able to represent that similarity. For example, the [k] sounds in *keep* and *cool* are different from each other in that the exact places they are articulated are dependent on the following vowel. The [k] in *keep* is produced farther forward in the mouth than the [k] in *cool.* But if we are not interested in this variation, because it is reasonably predictable, we want to make sure that these [k] sounds are not written with different symbols in our transcription system.

Based on the criteria above, the English spelling system is not a good phonetic alphabet because:

- sometimes the same sound is spelled using different letters, such as the [i] sound in *sea, see, scene, receive, thief, amoeba, machine,* and *Aesop;*
- sometimes the same letters can stand for different sounds, as in *sign, pleasure,* and *resign,* or *charter* and *character,* or *father, all, about, apple, any,* and *age;*

- sometimes a single sound is spelled by a combination of letters, as in *lock, that, book, boast, mountain, shop, apple,* or *special;*
- sometimes a single letter represents a combination of sounds, as in *exit* or *use;*
- sometimes letters stand for no sound at all, as in *know, doubt, though, island, rhubarb,* or *moose.*

A good phonetic transcription system is consistent and unambiguous because there is always a one-to-one correspondence between sounds and symbols. This is even true across languages, so that the symbols you will be learning can be used to transcribe the sounds of any language.

In this book we use the International Phonetic Alphabet (IPA for short). For our purposes this phonetic alphabet is the right one to use because it is applicable to all spoken human languages, rather than just English, and it has all of the properties of a "useful phonetic alphabet" discussed above. Here's what Anthony Burgess (the author of *A Clockwork Orange* and other novels) had to say about the IPA. "I propose that the reader become familiar with the International Phonetic Alphabet, or IPA. The reader may shudder in advance, but we have to do something about the accurate visualization of speech. I have the idealistic vision of phonetic symbols being added to our daily stock of alphabetic signs, so that I may tap the symbol /ə/ on my typewriter or word processor when I want to represent the sound that begins *apart* or ends *Asia*" (*A Mouthful of Air* [New York: Quill, 1992], pp. 26–27).

2.1.3 Types of Speech Sounds

In order to create a good phonetic transcription system, we need to know what types of sounds we are trying to transcribe. Phoneticians divide the speech stream into two main categories: **segments** and **suprasegmentals**. Segments are the discrete units of the speech stream and can be further subdivided into the categories consonants (File 2.2) and vowels (File 2.3). These sounds are transcribed easily using discrete symbols like [p] and [i]. Suprasegmentals, on the other hand, can be said to "ride on top of" segments in that they apply to entire strings of consonants and vowels—these are properties such as stress, tone, and intonation (File 2.5). These properties are somewhat more difficult to represent using an alphabetic-like transcription system, and there are many different ways they can be transcribed.

From an articulatory point of view, consonants and vowels are both made by positioning the vocal tract in a particular configuration. However, consonants are distinguished from vowels in that consonants are produced with a constriction somewhere in the vocal tract that impedes airflow, while vowels have at most only a slight narrowing and allow air to flow freely through the oral cavity. We can also distinguish consonants and vowels acoustically, based on the type of sounds they produce: consonants are much quieter than vowels and usually cannot function as the **nucleus** of a syllable. The syllable nucleus is the "heart" of the syllable, carrying suprasegmental information such as stress, loudness, and pitch, which vowels are much better suited to do than consonants.

Vowels in turn are often divided into two categories: **monophthongs** ([mɑnɔpθɑŋz]) and **diphthongs** ([dɪfθɑŋz] or [dɪpθɑŋz]). You can think of monophthongs as simple vowels, composed of a single configuration of the vocal tract, while diphthongs are complex vowels, composed of a sequence of two different configurations. We consider diphthongs to be "single" vowels, however, because the sequence of two configurations acts as the nucleus to a single syllable. To conceptualize this better, think of the two words *knives* and *naive*. The actual vowel sounds in these two words are essentially the same, but in *knives,* there is just one syllable nucleus (the diphthong [ɑɪ]), while in *naive,* there are two separate nuclei (the monophthong [ɑ] in the first syllable, followed by the monophthong [i] in the second syllable). The differences between monophthongs and diphthongs will be discussed in more detail in File 2.3.

2.1.4 Phonetic Symbols for English

This section lists the IPA symbols for English segments that we will be using in this book. Phonetic symbols are written in square brackets, [ ], to distinguish them from letters or words written in ordinary spelling. It is important to remember that these symbols are not the same as letters of English. Rather, they represent the sounds of language. The following table gives the phonetic symbols for the sound inventory of Standard American English, and the example words make use of Standard American English pronunciations. (Other sounds and symbols will be introduced in File 2.4.)

Symbol	Sample Words	Name of Symbol
Consonants:		
[p]	pit, tip, spit, hiccough, appear	
[b]	ball, globe, amble, brick, bubble	
[t]	tag, pat, stick, pterodactyl, stuffed	
[d]	dip, card, drop, loved, batted	
[k]	kit, scoot, character, critique, exceed	
[g]	guard, bag, finger, designate, Pittsburgh	
[ʔ]	uh-oh, hatrack, Batman	glottal stop
[f]	foot, laugh, philosophy, coffee, carafe	
[v]	vest, dove, gravel, anvil, average	
[θ]	through, wrath, thistle, ether, teeth	theta
[ð]	the, their, mother, either, teethe	eth, [ɛð]
[s]	soap, psychology, packs, descent, peace, excruciating	
[z]	zip, roads, kisses, Xerox, design	
[ʃ]	shy, mission, nation, glacial, sure	esh, [ɛʃ]
[ʒ]	measure, vision, azure, casualty, decision	yogh, [joʊg] or ezh, [ɛʒ]
[h]	who, hat, rehash, hole, whole	
[tʃ]	choke, match, feature, constituent	
[dʒ]	judge, George, Jell-O, region, residual	
[m]	moose, lamb, smack, amnesty, ample	
[n]	nap, design, snow, know, mnemonic	
[ŋ]	lung, think, finger, singer, ankle	engma or eng
[l]	leaf, feel, Lloyd, mild, applaud	
[ɹ]	reef, fear, Harris, prune, carp	
[ɾ]	writer, butter, udder, clutter, cuter	flap
[w]	with, swim, mowing, queen, twilight	
[ʍ]	which, where, what, whale, why (for those dialects in which *witch* and *which* do not sound the same)	voiceless 'w'
[j]	you, beautiful, feud, use, yell	lower-case 'j'

Symbol	Sample Words	Name of Symbol
Syllabic Consonants:		
[m̩]	possum, chasm, Adam, bottomless	syllabic 'm'
[n̩]	button, chicken, lesson, kittenish	syllabic 'n'
[l̩]	little, single, simple, stabilize	syllabic 'l'
[ɹ̩]	ladder, singer, burp, percent	syllabic 'ɹ'
Vowels		
i. Monophthongs (Simple Vowels)		
[i]	beat, we, believe, people, money	
[ɪ]	bit, consist, injury, malignant, business	small capital 'i'
[ɛ]	bet, reception, says, guest	epsilon
[æ]	bat, laugh, anger, comrade, rally	ash
[u]	boot, who, sewer, duty, through	
[ʊ]	put, foot, butcher, could, boogie-woogie	upsilon
[ɔ]	bought, caught, wrong, stalk, core	open 'o'
[ɑ]	pot, father, sergeant, honor, hospital	script 'a'
[ʌ]	but, tough, another, oven	wedge or turned 'v'
[ə]	among, sofa, Asia	schwa
ii. Diphthongs (Complex Vowels)		
[ɑɪ]	bite, Stein, aisle, choir, island	
[ɑʊ]	bout, brown, doubt, flower, loud	
[ɔɪ]	boy, doily, rejoice, perestroika, annoy	
[oʊ]	boat, beau, grow, though, over	
[eɪ]	bait, reign, great, they, gauge	

In the list in the table above, we have given you examples of individual sounds in individual words. When we actually use language on a day-to-day basis, however, we speak in phrases and sentences, with all the words run together. This type of speech is known as **running speech** or **continuous speech**, and, although as linguists we sometimes need to break speech into its component parts of words and sounds, you should bear in mind that most everyday speech is not separated out into these pieces. In running speech, the pronunciations of words may be affected by the surrounding words (see Section 2.2.6 on phonetic coarticulation or File 3.2 on phonological assimilation), and one of the open research questions in the study of language processing is how the human mind processes running speech into its meaningful component parts (see Chapter 9).