4 The Sound Patterns of Language

Uans appona taim uas tri berres; mamma berre, pappa berre, e beibi berre. Live inne contri nire foresta. NAISE AUS. No mugheggia. Uanna dei pappa, mamma, e beibi go bice, orie e furghetta locche di dorra.

Bai ene bai commese Goldilocchese. Sci garra natingha tu du batte meiche troble. Sci puscia olle fudde daon di maute; no live cromma. Den sci gos appesterrese enne slipse in olle beddse.

Bob Belviso, quoted in Espy (1975)

In the preceding chapter, we investigated the physical production of speech sounds in terms of the articulatory mechanisms of the human vocal tract. That investigation was possible because of some rather amazing facts about the nature of language. When we considered the human vocal tract, we didn't have to specify whether we were talking about a fairly large person, over 6 feet tall, weighing over 200 pounds, or about a rather small person, about 5 feet tall, weighing less than 100 pounds. Yet those two physically different individuals would inevitably have physically different vocal tracts, in terms of size and shape. In a sense, every individual has a physically different vocal tract. Consequently, in purely physical terms, every individual will pronounce sounds differently. There are, then, potentially millions of physically different ways of saying the simple word *me*.

Phonology

In addition to those millions of different individual vocal tracts, each individual will not pronounce the word *me* in a physically identical manner on every occasion. Obvious differences occur when that individual is shouting, or has just woken from a deep sleep, or is suffering from a bad cold, or is trying to ask for a sixth martini, or any combination of these. Given this vast range of potential differences in the actual physical production of a speech sound, how do we manage consistently to recognize all those versions of *me* as the form [mi], and not [ni] or [si] or [mæ] or [mo] or something else entirely? The answer to that question is provided to a large extent by the study of phonology.

Phonology is essentially the description of the systems and patterns of speech sounds in a language. It is, in effect, based on a theory of what every adult speaker of a language unconsciously knows about the sound patterns of that language. Because of this theoretical status, phonology is concerned with the abstract or mental aspect of the sounds in language rather than with the actual physical articulation of speech sounds. If we can make sense of Bob Belviso's comic introduction to the Goldilocks story on page 42, we must be using our phonological knowledge of sounds in English words to overcome some very unusual spellings. We can use various different ways of spelling the words in the first and second lines below, but the underlying phonological representation in the third line is constant. (See the end of the chapter for a full translation of the story.)

Uans appona taim uas tri berres
Ones up on atam waz theree bars
/wʌns əpan ə taim wəz θri berz/

Phonology is about the underlying design, the blueprint of each sound type, which may vary in different physical contexts. When we think of the [t] sound in the words *tar, star, writer, butter* and *eighth* as being "the same," we actually mean that, in the phonology of English, they would be represented in the same way. In actual speech, these [t] sounds are all potentially very different from each other because they can be pronounced in such different ways in relation to the other sounds around them.

However, all these articulation differences in [t] sounds are less important to us than the distinction between the [t] sounds in general and the [k] sounds, or the [f] sounds, or the [b] sounds, because there are meaningful consequences related to the use of one rather than the others. These sounds must be distinct meaningful sounds, regardless of which individual vocal tract is being used to pronounce them, because they are what make the words *tar, car, far* and *bar* meaningfully distinct. Considered from this point of view, we can see that phonology is concerned with the abstract representation of sounds in our minds that enables us to recognize and interpret the meaning of words on the basis of the actual physical sounds we say and hear.

Phonemes

Each one of these meaning-distinguishing sounds in a language is described as a **phoneme**. When we learn to use alphabetic writing, we are actually using the concept of the phoneme as the single stable sound type that is represented by a single written symbol. It is in this sense that the phoneme /t/ is described as a sound type, of which all the different spoken versions of [t] are tokens. Note that slash marks are conventionally used to indicate a phoneme, /t/, an abstract segment, as opposed to the square brackets, as in [t], used for each phonetic or physically produced segment.

An essential property of a phoneme is that it functions contrastively. We know there are two phonemes /f/ and /v/ in English because they are the only basis of the contrast in meaning between the words *fat* and *vat*, or *fine* and *vine*. This contrastive property is the basic operational test for determining the phonemes in a language. If we change one sound in a word and there is a change of meaning, the sounds are distinct phonemes.

Natural Classes

The descriptive terms we used to talk about sounds in Chapter 3 can be considered "features" that distinguish each phoneme from the next. If the feature is present, we mark it with a plus sign (+) and if it is not present, we use a minus sign (−). Thus /p/ can be characterized as [−voice, +bilabial, +stop] and /k/ as [−voice, +velar, +stop]. Because these two sounds share some features, they are sometimes described as members of a **natural class** of phonemes. Phonemes that have certain features in common tend to behave phonologically in some similar ways. Table 4.1 presents an analysis of some of the distinguishing features of four English phonemes. Only /p/ and /k/ have sufficient features in common to be members of a natural class. They are both voiceless stops.

TABLE 4.1

/p/	/k/	/v/	/n/
–voice	–voice	+voice	+voice
+bilabial	+velar	+labiodental	+alveolar
+stop	+stop	+ fricative	+nasal

In contrast, /v/ has the features [+voice, +labiodental, +fricative] and so cannot be in the same natural class of sounds as /p/ and /k/. Although other factors will be involved, this feature analysis could lead us to suspect that there may be a good phonological reason why words beginning with /pl-/ and /kl-/ are common in English, but words beginning with /vl-/ or /nl-/are not. This type of feature analysis allows us to describe not only individual phonemes, but also the possible sequences of phonemes in a language.

Phones and Allophones

While the phoneme is the abstract unit or sound type ("in the mind"), there are many different versions of that sound type regularly produced in actual speech ("in the mouth"). We can describe those different versions as **phones**, which are phonetic units, in square brackets. When we have a set of phones, all of which are versions of one phoneme, we add the prefix "allo-" (= one of a closely related set) and call them **allophones** of that phoneme.

For example, the phoneme /t/ can be pronounced in a number of physically different ways as phones. The [t] sound in the word *tar* is normally pronounced with a stronger puff of air than is present in the [t] sound in the word *star*. If you put the back of your hand in front of your mouth as you say *tar*, then *star*, you should feel some physical evidence of **aspiration** (the puff of air) accompanying the [t] sound at the beginning of *tar* (but not in *star*). This aspirated phone is represented more precisely as [tʰ].

In the last chapter, we noted that the [t] sound between vowels in a word like *writer* often becomes a flap, which we can represent as [ɾ]. That's another phone.

We also saw that a word like *butter* can have a glottal stop as the middle consonant in the pronunciation, so the part written as "t" may be pronounced as [ʔ], which is yet another phone. In the pronunciation of a word like *eighth* (/eɪtθ/), the influence of the final dental [θ] sound causes a dental articulation of the [t] sound. This can be represented more precisely as [t̪]. That's yet another phone. There are even more variations of this sound which, like [tʰ], [ɾ], [ʔ] and [t̪], can be represented in a more precise way in a detailed, or narrow phonetic transcription. Because these variations are all part of one set of phones, they are referred to as allophones of the phoneme /t/, as shown in Table 4.2.

The crucial distinction between phonemes and allophones is that substituting one phoneme for another will result in a word with a different meaning (as well as a different pronunciation), but substituting allophones only results in a different (and perhaps unusual) pronunciation of the same word.

TABLE 4.2

Phoneme	Allophones	
/t/	[tʰ]	(*tar*)
	[ɾ]	(*writer*)
	[ʔ]	(*butter*)
	[t̪]	(*eighth*)

Complementary Distribution

When we have two different pronunciations (allophones) of a sound type (phoneme), each used in different places in words, they are said to be in **complementary distribution**. That is, the [tʰ] pronunciation of the phoneme /t/ with aspiration is used word-initially, as in *tar*, but never after another consonant in initial position, as in *star*. The places where /t/ occurs with aspiration, and without aspiration, never overlap and so the different pronunciations are in complementary distribution.

Minimal Pairs and Sets

Phonemic distinctions in a language can be tested via pairs and sets of words. When two words such as *fan* and *van* are identical in form except for a contrast in one phoneme, occurring in the same position, the two words are described as a **minimal pair**. When a group of words can be differentiated, each one from the others, by changing one phoneme (always in the same position in the word), they are described as a **minimal set**. Examples of contrasting pairs and sets are presented in Table 4.3.

TABLE 4.3

Minimal pairs		Minimal sets
fan – van	*bath – math*	*big – pig – rig – fig – dig – wig*
bat –beat	*math – myth*	*fat – fit – feet –fete – foot – fought*
sit –sing	*myth –Mick*	*cat –can – cap – cab – cash – cadge*

Phonotactics

This type of exercise with minimal sets also allows us to see that there are definite patterns in the types of sound combinations permitted in a language. The first minimal set in Table 4.3 does not include forms such as *lig* or *vig*. According to my dictionary, these are not English words, but they could be viewed as possible English words. That is, our phonological knowledge of the pattern of sounds in English words would allow us to treat these forms as acceptable if, at some future time, they came into use. They might, for example, begin as invented abbreviations (*I think Bubba is one very ignorant guy.* ~ *Yeah, he's a big vig!*). Until then, they represent "accidental" gaps in the vocabulary of English.

It is, however, no accident that forms such as [fsɪg] or [rnɪg] do not exist or are unlikely ever to exist. They have been formed without obeying some constraints on the sequence or position of English phonemes. Such constraints are called the **phonotactics** (i.e. permitted arrangements of sounds) in a language and are obviously part of every speaker's phonological knowledge. Because these constraints operate on a unit that is larger than the single segment or phoneme, we have to move on to a consideration of the basic structure of that larger phonological unit called the syllable.

Syllables

A **syllable** must contain a vowel or vowel-like sound, including diphthongs. The most common type of syllable also has a consonant (C) before the vowel (V) and is represented as CV. The basic elements of the syllable are the **onset** (one or more consonants) followed by the **rhyme**. The rhyme (sometimes written as "rime") consists of a vowel, which is treated as the **nucleus**, plus any following consonant(s), described as the **coda**.

Syllables like *me, to* or *no* have an onset and a nucleus, but no coda. They are known as **open syllables**. When a coda is present, as in the syllables *up, cup, at* or *hat*, they are called **closed syllables**. The basic structure of the kind of syllable found in English words like *green* (CCVC), *eggs* (VCC), *and* (VCC), *ham* (CVC), *I* (V), *do* (CV), *not* (CVC), *like* (CVC), *them* (CVC), *Sam* (CVC), *I* (V), *am* (VC) is shown in Figure 4.1.

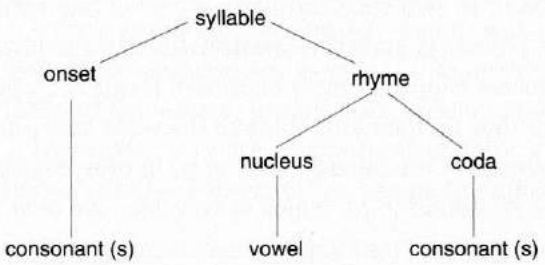

Figure 4.1

Consonant Clusters

Both the onset and the coda can consist of more than a single consonant, also known as a **consonant cluster**. The combination /st/ is a consonant cluster (CC) used as onset in the word *stop*, and as coda in the word *post*. There are many CC onset combinations permitted in English phonotactics, as in *black, bread, trick, twin, flat* and *throw*. Note that liquids (/l/, /r/) and a glide (/w/) are used in second position.

English can actually have larger onset clusters, as in the words *stress* and *splat*, consisting of three initial consonants (CCC). When we study the phonotactics of these larger onset consonant clusters, we can find a fairly regular pattern. The first consonant must always be /s/, followed by one of the natural class of voiceless stops (/p/, /t/, /k/), plus a liquid or a glide (/l/, /r/, /w/). We can check if this description is adequate for the combinations in *splash, spring, strong, scream* and *squeeze* (/skwiz/). Does the description also cover the second syllable in the pronunciation of *exclaim*? How about /ɛk-skleɪm/? Remember that it is the onset of the syllable that is being described, not the beginning of the word. See Task D on page 51 for more syllables and clusters.

Coarticulation Effects

It is quite unusual for languages to have large consonant clusters of the type just described. In English, large clusters may be reduced in casual conversational speech, particularly if they occur in the middle of a word. This is just one example of a process that is usually discussed in terms of **coarticulation effects**.

In much of the preceding discussion, we have been describing speech sounds in syllables and words as if they are always pronounced carefully in slow motion. Speech is not normally like that. Mostly our talk is fast and spontaneous, and it requires our articulators to move from one sound to the next without stopping. The process of making one sound almost at the same time as the next sound is called coarticulation.

Assimilation

When two sound segments occur in sequence and some aspect of one segment is taken or "copied" by the other, the process is known as **assimilation**. In the physical production of speech, this regular process happens simply because it is quicker, easier and more efficient for our articulators as they do their job. Think of the word *have* /hæv/ by itself, then think of how it is pronounced in the phrase *I have to go* in everyday speech. In this phrase, as we start to say the /t/ sound in *to*, which is voiceless, we tend to produce a voiceless version of the preceding sound, resulting in what sounds more like /f/ than /v/. So, we typically say [hæftə] in this phrase and you may even see it written informally as "hafta," showing how the assimilation from a voiced to a voiceless sound is perceived.

Nasalization

Vowels are also subject to assimilation. In isolation, we would typically pronounce [ɪ] and [æ] with no nasal quality at all. However, when we say the words *pin* and *pan* in everyday talk, the anticipation of the final nasal consonant makes it easier to go into the nasalized articulation in advance. This process is known as **nasalization** and can be represented with a small diacritic (~), called "tilde," over the vowel symbol. The vowel sounds in those words will be, in more precise transcription, [ĩ] and [æ̃]. This process is such a regular feature of English that a phonological rule can be stated in the following way: "Any vowel becomes nasal whenever it immediately precedes a nasal."

This type of assimilation process occurs in a variety of different contexts. By itself, the word *can* may be pronounced as [kæn], but, when we say *I can go*, the influence of the following velar [g] in *go* will typically make the preceding nasal sound come out as [ŋ] (velar) rather than [n] (alveolar). The most commonly observed conversational version of the phrase is [aɪkəŋgoʊ]. Notice that the vowel in *can* has also changed to schwa [ə] from the isolated-word version [æ]. We may also pronounce *and* as [ænd] by itself, but in the normal use of the phrase *you and me*, we usually say [ən], as in [juənmi].

Elision

In the last example, illustrating the normal pronunciation of *you and me*, the [d] sound of the word *and* was not included in the transcription. That is because it is not usually pronounced in this phrase. In the environment of a preceding nasal [n] and a following nasal [m], we simply don't devote speech energy to including the stop sound [d].

There is also typically no [d] sound included in the everyday pronunciation of a word like *friendship* [frɛnʃɪp]. This process of not pronouncing a sound segment that might be present in the deliberately careful pronunciation of a word in isolation is described as **elision**. In consonant clusters, especially in coda position, /t/ is a common casualty in this process, as in the typical pronunciation [æspɛks] for *aspects*, or in [himəsbi] for the phrase *he must be*. We can, of course, slowly and deliberately pronounce each part of the phrase *we asked him*, but the process of elision (of /k/) in casual conversation is likely to produce [wiæstəm].

Vowels also disappear through elision, with the result that sometimes a whole syllable may not be pronounced, as in [ɛvri] for *every*, [ɪntrɪst] for *interest*, [kæbnət] for *cabinet*, [kæmrə] for *camera*, [prɪznər] for *prisoner* and [spoʊz] for *suppose*.

These processes are summarized in Table 4.4. We use a pair of symbols (/ ____) to indicate "in the context of" or "under the influence of" the following element.

TABLE 4.4

Assimilation: making a sound segment more similar to the next one
voiced (→ voiceless) /____ + voiceless: hæv + tu → hæftə
Nasalization: adding a nasal quality to a sound segment before a nasal sound
non-nasal (→ nasal) /____ + nasal: pæ + n → pæ̃n
Elision: leaving out a sound segment
consonant cluster (→ reduced) / ____ + consonant: məst + bi → məsbi
three syllables (→ two syllables) /____ + syllable: prɪzənər → prɪznər

Normal Speech

These processes of assimilation, nasalization and elision occur in everyone's normal speech and should not be regarded as some type of sloppiness or laziness in speaking. In fact, consistently avoiding the regular patterns of assimilation, nasalization and elision used in a language would result in extremely artificial-sounding talk. The point of investigating these phonological processes is not to arrive at a set of rules about how a language should be pronounced, but to try to come to an understanding of the regularities and patterns that underlie the actual use of sounds in language.