

# 5

## 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 furchetta 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 six feet tall, weighing over 200 pounds, or about a rather small person, about five 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*.

In addition, each individual will not pronounce the word *me* in a physically identical manner on every occasion. Obvious differences occur when that individual is shouting, is suffering from a bad cold or is asking for a sixth martini. 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 [ma] or [mo] or something else entirely? The answer to that question is provided to a large extent by the study of phonology.

---

### 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 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 manage to make sense of Bob Belviso's comic introduction to the story of Goldilocks and the Three Bears quoted earlier, we must be using our phonological knowledge of likely combinations of sounds in English words to overcome some very unusual spellings of those words. (See the end of the chapter for a translation.)

Phonology is about the underlying design, the blueprint of each sound type, which serves as the constant basis of all the variations in different physical articulations of that sound type in different contexts. When we think of the [t] sound in the words *tar*, *star*, *writer* 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 very different.

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 set of sounds in a language that allows us to distinguish meaning in 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 considered the basis of alphabetic writing in chapter 3, we were actually working with the concept of the phoneme as the single sound type which came to be 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 that exist in a language. If we substitute one sound for another in a word and there is a change of meaning, then the two sounds represent different phonemes. The basic phonemes of English are listed in the consonant and vowel charts in chapter 4.

The technical terms used in creating those charts 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's 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 sounds. The prediction would be that sounds which have features in common would behave phonologically in some similar ways. A sound which does not share those features would be expected to behave differently.

For example, /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 /p1-/ and /k1-/ are common in English, but words beginning with /v1-/ are not. Could it be that there are some definite sets of features required in a sound in order for it to occur word-initially before /l/? If so, then we will be on our way to producing a phonological account of permissible sound sequences in the 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**. Phones are phonetic units and appear in square brackets. When we have a group of several phones, all of which are versions of one phoneme, we add the prefix ‘allo-’ (= one of a closely related set) and refer to them as **allophones** of that phoneme.

For example, 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 be able to 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 version is represented more precisely as [t<sup>h</sup>]. That’s one phone. 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 [D]. That’s another phone. In the pronunciation of a word like *eighth* (/etθ/), 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 [t<sup>h</sup>], [D], 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 typically referred to as allophones of the phoneme /t/.

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.

Let's look at another quick example, using a vowel sound. In English, there is a subtle difference in the pronunciation of /i/ in the words *seed* and *seen*. In the second word, the effect of the nasal consonant [n] makes the [i] sound nasalized. We can represent this **nasalization** with a small mark (˜), called 'tilde', over the symbol [i] in a narrow phonetic transcription. So, there are at least two phones, [i] and [ĩ], used to realize the single phoneme. They are both allophones of /i/ in English.

It is possible, of course, for two languages to have the same pair of phonetic segments, but to treat them differently. In English, the effect of nasalization on a vowel is treated as allophonic variation because the nasalized version is not meaningfully contrastive. Whether we say [sin] or [sɪ̃n], people will only recognize one word *seen*. In French, however, the pronunciation [mɛ] is used for one word *mets* ('dish') and the pronunciation [mɛ̃] is used for another word *main* ('hand'). Also, [so] for the word *seau* ('pail') contrasts with [sɔ̃] for the word *son* ('sound'). Clearly, in these cases, the distinction is phonemic.

## Minimal pairs and sets

Phonemic distinctions in a language can be tested via pairs and sets of words. When two words such as *pat* and *bat* 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**. More accurately, they would be classified as a minimal pair in the phonology of English. (Arabic, for example, does not have this contrast between /p/ and /b/.) Other examples of English minimal pairs are *fan–van*, *bet–bat*, *site–side*. Such pairs have traditionally been used in the teaching and testing of English as a second or foreign language to help students develop the ability to understand the contrast in meaning based on the minimal sound contrast.

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), then we have a **minimal set**. For example, one minimal set based on the vowel phonemes of English could include *feat*, *fit*, *fat*, *fate*, *fought*, *foot*, and another minimal set based on consonant phonemes could have *big*, *pig*, *rig*, *fig*, *dig*, *wig*.

## Phonotactics

This type of exercise involving minimal sets also allows us to see that there are definite patterns in the types of sound combinations permitted in a language. In English, the minimal set we have just listed does not include forms such as *lig* or *wig*. According to the dictionary, these are not English words, but they could be viewed as possible English words. That is, our phonological knowledge of

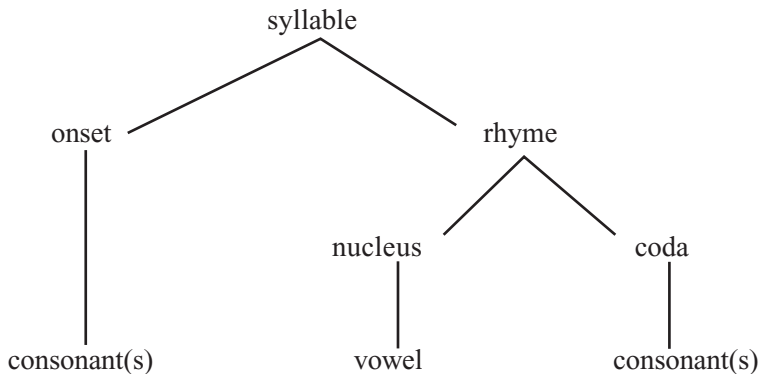
the pattern of sounds in English words would allow us to treat these forms as acceptable if, at some future time, they came into use. They might, for example, begin as invented abbreviations (*I think Bubba is one very ignorant guy. ~ Yeah, he's a big vig). Until then, they represent 'accidental' gaps 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 and clusters

A **syllable** must contain a vowel (or vowel-like) sound. The most common type of syllable in language also has a consonant (C) before the vowel (V) and is typically represented as CV. Technically, the basic elements of the syllable are the **onset** (one or more consonants) and 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 the accompanying diagram.



Both the onset and the coda can consist of more than one 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 being used in second position.

English can actually have larger onset clusters, as in the words *stress* and *splat*, consisting of three initial consonants (CCC). The phonotactics of these larger onset consonant clusters is not too difficult to describe. The first consonant must always be /s/, followed by one of the voiceless stops (/p/, /t/, /k/) and then one of these liquids or glides (/l/, /r/, /w/). You can check if this description is adequate for the combinations in *splash*, *spring*, *strong*, *scream* and *square*. Does the description also cover the second syllable in the pronunciation of *exclaim*? How about /ɛk-sklejm/? Remember that it is the onset of the syllable that is being described, not the beginning of the word.

It is quite unusual for languages to have consonant clusters of this type. Indeed, the syllable structure of many languages (e.g. Japanese) is predominantly CV. It is also noticeable in English that large consonant clusters are frequently 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 co-articulation effects.

## Co-articulation effects

In much of the preceding discussion, we have been describing speech sounds in syllables and words as if they are always pronounced carefully and deliberately, almost in slow motion. Speech isn't 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 **co-articulation**. There are two well-known co-articulation effects, described as assimilation and elision.

## 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**. If we think of the physical production of speech, we realize that this regular process happens simply because it's quicker, easier and more efficient for our articulators as they do their job. In isolation, we would typically pronounce [ɪ] and [æ] without any nasal quality at all. However, when we say words like *pin* and *pan* in everyday speech, the anticipation of forming the final nasal consonant will make it easier to go into the nasalized articulation in advance and consequently the vowel sounds in these words will be, in more precise transcription, [ɪ̃] and [æ̃]. This is a very regular feature of English speakers' pronunciation. It is so regular, in

fact, 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] will almost certainly make the preceding nasal sound come out as [ŋ] (the velar nasal) rather than [n] (the alveolar nasal). The most commonly observed conversational version of the phrase is [ajkəŋgo]. Notice that the vowel in *can* has also changed to schwa [ə] from the isolated-word version [æ]. In many words spoken carefully, the vowel receives stress, but in the course of ordinary everyday talk, that vowel may no longer receive any stress and naturally reduce to schwa. We may, for example, 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’s because it isn’t 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]. This isn’t laziness, it’s efficiency. 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 [æspeks] for *aspects*, or in [hɪməsbɪ] 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 in casual conversation is likely to produce [wiæstəm]. Vowels also disappear, as in [ɛvri] for *every*, [ɪntrɪst] for *interest*, [kæbnət] for *cabinet*, and [spowz] for *suppose*.

## Normal speech

These two processes of assimilation 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 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 which underlie the actual use of sounds in language.

### ■ Study questions

- 1 What is the difference between an allophone and a phoneme?
- 2 What is an aspirated sound?
- 3 Which of the following words would be treated as minimal pairs?

*ban, fat, pit, bell, tape, heat, meal, more, pat, pen, chain, vote, bet, far, bun, goat, heel, sane, tale*

- 4 What is meant by the phonotactics of a language?
- 5 What is the difference between an open and a closed syllable?
- 6 Which segments in the pronunciation of the following words are most likely to be affected by elision?  
(a) *government* (b) *postman* (c) *pumpkin* (d) *sandwich*

### ■ Research tasks

- A What are diacritics and which ones were used in this chapter to identify sounds?
- B Is the difference between using *a* (*a banana*) and *an* (*an apple*) based on a spelling rule of written English or a phonological rule of spoken English and what kind of examples would provide clear evidence in support of the rule?
- C According to Radford *et al.* (2006), the word *central* has a consonant cluster (-*ntr-*) in the middle and two syllables. What do you think is the best way to divide the word into two syllables (*ce + ntral*, *centr + al*, *cen + tral*, *cent + ral*) and why?
- D Individual sounds are described as segments. What are suprasegmentals?

### ■ Discussion topics/projects

- I We can form negative versions of words such as *audible* and *edible* in English by adding *in-* to produce *inaudible* and *inedible*. How would you describe the special phonological processes involved in the pronunciation of the negative versions of the following words?

*complete, decent, gratitude, legal, literate, mature, mortal, perfect, possible, rational, responsible, sane*

(For background reading, see chapter 1 of Harris, 1994.)

- II The use of plural *-s* in English has three different, but very regular, phonological alternatives. We add /s/ to words like *bat*, *book*, *cough* and *ship*. We add /z/ to words like *cab*, *cave*, *lad*, *rag* and *thing*. We add /əz/ to words like *bus*, *bush*, *church*, *judge* and *maze*.  
(a) Can you identify the sets of sounds that regularly precede each of these alternative pronunciations of the plural ending?



- (b) What features do each of these sets have in common?  
(For background reading, see chapter 4 of Finegan, 2004.)

### ■ **Bob Belviso translated**

One attempt to interpret those very unusual spellings might be as follows:

*Once upon a time was three bears; mama bear, papa bear, and baby bear. Live in the country near the forest. NICE HOUSE. No mortgage. One day papa, mama, and baby go beach, only they forget to lock the door.*

*By and by comes Goldilocks. She got nothing to do but make trouble. She push all the food down the mouth; no leave a crumb. Then she goes upstairs and sleeps in all the beds.*

### ■ **Further reading**

For another introduction to phonology, see chapter 7 in Fromkin *et al.* (2003), or chapter 3 in O'Grady *et al.* (2005). Introductory texts are Carr (1999), Gussenhoven & Jacobs (1998) and a number of the texts listed in the Further Reading section of chapter 4. More specialized texts are Gussman (2002), McMahon (2002) and Roach (2001b). For a dictionary of technical terms, see Crystal (2003a) or Trask (1996a). For more on assimilation and elision, see Brown (1990), and on assimilation rules, see chapter 3 of Hudson (2000).