## 4 <br> 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 $m e$.

## 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 $m e$ 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
/w^ns əpan ə taim wəz Өri 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 $[\mathrm{k}]$ 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 $/ \mathrm{p} /$ and $/ \mathrm{k} /$. 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 [ $\mathrm{t}^{\mathrm{h}}$ ].

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 [r]. 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 [?], which is yet another phone. In the pronunciation of a word like eighth (/eIt $\theta /$ ), the influence of the final dental [ $\theta$ ] sound causes a dental articulation of the [ t ] sound. This can be represented more precisely as [ t$]$. That's yet another phone. There are even more variations of this sound which, like [ $\mathrm{t}^{\mathrm{h}}$ ], [ r$],[\mathrm{P}]$ and [tr], 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.

Table 4.2

| Phoneme | Allophones |  |
| :--- | :--- | :---: |
| $[$ th] | $(\underline{\text { tar })}$ |  |
|  | $[\mathrm{c}]$ | $($ writer $)$ |

## (eighth)

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.

## 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

| fan - van | $\underline{\text { bath - math }}$ | $\begin{aligned} & \underline{\text { big }} i g-\text { pig - rig }-\mathrm{fig}-\underline{\operatorname{d} i g}- \\ & \text { wig } \end{aligned}$ |
| :---: | :---: | :---: |
| bat -beat | math - myth | fat $-f$ ít - 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 [fsig] or [rnig] 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 codla.

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), $a m$ (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 $/ \mathrm{s} /$, 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 / $\varepsilon$ k-skleIm/? Remember that it is the onset of the syllable that is being described, not the beginning of the word. See Task D on page $\underline{51}$ 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 [I] 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 ( $\sim$ ), called "tilde," over the vowel symbol. The vowel sounds in those words will be, in more precise transcription, [ I$]$ and [ $\tilde{\mathbb{x}}]$. 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 [arkəŋgov]. 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 [frenfip]. 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 [æspદks] 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 $/ \mathrm{k} /$ ) 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 [عvri] for every, [Intrist] for interest, [kæbnət] for cabinet, [kæmrə] for camera, [prIznər] for prisoner and [spoひz] 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

$$
\text { voiced ( } \rightarrow \text { voiceless) /___ + voiceless: hæv + tu } \rightarrow \text { hæftə }
$$

Nasalization: adding a nasal quality to a sound segment before a nasal sound

$$
\text { non-nasal ( } \rightarrow \text { nasal) /___ + nasal: pæ + n } \rightarrow \text { p } \tilde{n}
$$

Elision: leaving out a sound segment

$$
\text { consonant cluster ( } \rightarrow \text { reduced) / ___ + consonant: məst + bi } \rightarrow
$$ məsbi

three syllables ( $\rightarrow$ two syllables) /___+ syllable: prizənər $\rightarrow$ prIznər

## 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 artificialsounding 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.

## Study Questions

1 In French, the words /bo/ for beau ("handsome") and /bõ/ for bon ("good") seem to have different vowels. Are these two vowels allophones or phonemes in French?

2 Which English phoneme has the features: -voice, +velar, +stop?
3 What is an aspirated sound and which of the following words would normally be pronounced with one?
kill, pool, skill, spool, stop, top

4 Does this phrase (big black bag) contain a minimal pair, a minimal set, or neither?

5 Which of the following words would be treated as minimal pairs?
ban, fat, pit, bell, tape, heat, meal, more, pat, tap, pen, chain, vote, bet, far, bun, goat, heel, sane, tale, vet

6 What is meant by the phonotactics of a language?
7 In the pronunciation of track, which sound(s) would be the nucleus?
8 What is the difference between an open and a closed syllable?
9 Is the nasal consonant in the everyday pronunciation of I can go alveolar or velar?

10 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 (e) victory

## Tasks

A What are diacritics and which ones were used in this chapter to identify sounds?

B Individual sounds are described as segments. What are suprasegmentals?
C
(i) In the phonology of the Hawaiian language there are only open syllables. Using this information, can you work out how English "Merry Christmas" became "Mele Kalikimaka" for people in Hawai'i? Also, based on this slender evidence, which two English consonants are probably not phonemes in Hawaiian?
(ii) Including the glottal stop $/ \mathrm{R} /$, described in Chapter 3, Hawaiian has eight consonant phonemes. Looking at the list of Hawaiian names below, can you identify the other seven Hawaiian consonants?
(iii) Can you pair each Hawaiian name with its matching English name from the second list (e.g. Henele = Henry)?

| Henele, Kala, Kalona, Kania, | Bev, David, Fabian, Fred, |
| :--- | :--- |
| Kawika, Keoki, Kimo, Likeke, | George, Henry, Jim, Richard, |
| Lopaka, Papiano, Peleke, Pewi | Robert, Sarah, Sharon, Tanya |

D 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?

E The English words lesson and little are typically pronounced with syllabic consonants.
(i) What exactly is a syllabic consonant and how would it appear in a phonetic transcription?
(ii) Which of these words would most likely be pronounced with a syllabic consonant?
bottle, bottom, button, castle, copper, cotton, paddle, schism, wooden
F We can use the phonology of English as a guide to some regularities in spelling, based on a distinction between shorter vowels and longer vowels (including diphthongs).
(i) Complete Table 4.5 by adding the minimal pairs to appropriate columns to illustrate differences in spelling that match differences in pronunciation here. Can you also add the appropriate phonological forms of the three diphthongs at the top of each column?
(ii) Can you state a general principle connecting spelling to phonology (in these examples)?
back, bake, cock, coke, dame, damn, diner, dinner, dole, doll, hoping, hopping, later, latter, Mick, Mike, mile, mill

Table 4.5 English phonology and spelling


G In the Spanish words mismo ("same") and isla ("island"), the " s " is pronounced as [z], but in the words este ("this") and pescado ("fish"), the " s " is pronounced [s].
(i) Based on these and the following examples, what is the rule for choosing [z] or [s]?

$$
(\underline{s}=[\mathrm{z}]) \quad(\underline{s}=[\mathrm{s}])
$$

béisbol ("baseball")
desde ("from")
rasgado ("torn")
socialismo ("socialism")

## España ("Spain")

casa ("house")
sistema ("system")
socialista ("socialist")
(ii) Based on this (rather slim) evidence, would you say that the difference is phonemic or allophonic?
$\mathbf{H}$ A general distinction can be made among languages depending on their basic rhythm, whether they have syllable-timing or stress-timing. How are these two types of rhythm distinguished and which type characterizes the pronunciation of English, French and Spanish?

I The following examples are from Cree, a Native American or Aboriginal language spoken in many areas across Canada. It has been noted that voiced and voiceless stops have different distributions in Cree than in English.
(i) Can you describe the distribution of $[\mathrm{p}]$ and $[\mathrm{b}]$ by analyzing the places they occur in the following words (from Cowan and RakuŠan, 1999)?
(ii) Based on this limited evidence, would you say that [p] and [b] are likely to be phonemes or allophones in Cree?
(1) peyak
("one")
(5) asabap
("thread")
(2) nistosap
("twelve")
(6) kiba
("soon")
(3) tanispi
("when")
(7) mibit
("tooth")
(4) ospuagan
("pipe")
(8) nabeu
("man")

## 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?
balance, compatible, complete, decent, glorious, gratitude, legal, literate, mature, perfect, possible, rational, responsible, sane, tolerant, variable
(For background reading, see chapter 3 (pages 75-78) of Payne, 2006.)
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/ or/Iz/ to words like bus, bush, church, judge and maze.
(i) Can you identify the sets of sounds that regularly precede each of these alternative pronunciations of the plural ending?
(ii) What are the features that each of these sets has in common?
(For background reading, see chapter 2 (pages 55-56) of Jeffries, 2006.)

## 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

## Basic Treatments

Davenport, M. and S. Hannahs (2013) Introducing Phonetics and Phonology (3rd edition) Routledge

McCully, C. (2009) The Sound Structure of English: An Introduction Cambridge University Press

McMahon, A. (2002) An Introduction to English Phonology (16-17) Edinburgh University Press

## More Detailed Treatments

Carr, P. (2013) English Phonetics and Phonology (2nd edition) Wiley-Blackwell

Odden, D. (2014) Introducing Phonology (2nd edition) Cambridge University Press

Roach, P. (2009) English Phonetics and Phonology (4th edition) Cambridge University Press

## Syllables

Duanmu, S. (2008) Syllable Structure Oxford University Press

Ladefoged, P. and K. Johnson (2015) A Course in Phonetics (7th edition) (chapter 10) Wadsworth, Cengage

## Phonotactics

Bauer, L. (2015) "English phonotactics" English Language and Linguistics 19: 437-475

Herbst, T. (2010) English Linguistics (chapter 5) de Gruyter

## Coarticulation

Hardcastle, W. and N. Hewlett (2006) Coarticulation: Theory, Data and Techniques Cambridge University Press

## Assimilation and Elision

Brown, G. (1990) Listening to Spoken English (2nd edition) Longman

## Other References

Cowan, W. and J. Rakušan (1999) Source Book for Linguistics (3rd edition) John Benjamins

Jeffries, L. (2006) Discovering Language Palgrave Macmillan
Payne, T. (2006) Exploring Language Structure Cambridge University Press

## 5

## Word Formation



Autocorrect I could do without. It thinks I am stupid and clumsy, and while it's true that I don't know how to disable it and I can't text with thumbs like a teenager (though I am prehensile), why would I let a machine tell me what I want to say? I text someone "Good night" in German, and instead of "Gute Nacht" I send "Cute Nachos." I type "adverbial," and it comes out "adrenal,"

