

## The Consonants of English

We will begin this chapter by reviewing some of the gestures involved in producing the consonants of English. In the materials for this chapter on the CD, there are two movies. The first shows the pronunciation of consonants that, in the traditional terms, can be said to have different places of articulation. The stops [p, t, k] are illustrated in the nonsense utterances [hapa, hata, haka]. These stops are said to be bilabial, alveolar, and velar. But it is not just the different places on the roof of the mouth that distinguish these sounds. They are equally characterized by the movements of the lips and different parts of the tongue. Look at the movie on the CD and note the rapid movements of the lips for the first consonant, of the tip of the tongue for the second, and of the back of the tongue for the third.

CD 3.1

The second movie on the CD shows what are traditionally called different manners of articulation, illustrating the consonants [d, n, s] in the nonsense words [hade, hane, hase]. There are really three separate phrases in this utterance. Look at the movie and then go through it slowly. You can use the right arrow key, which is usually at the bottom right of the keyboard, to step through one frame at a time. In the first phrase, note how, at the left of the picture, the soft palate rises to form a velic closure in the first few frames, even before the tip of the tongue moves up to form a closure on the alveolar ridge for [hade]. Conversely, in the second phrase [hane], note that the soft palate moves up before the tongue moves, but this time only slightly. The soft palate does not make a complete closure and thus allows air to escape through the nose after the tongue tip has made a closure on the alveolar ridge for [n]. The third phrase in this movie, [hase], has tongue and soft palate gestures very similar to those in [hade]. The small differences in tongue shape are hard to see in this film, even when you step through it one frame at a time. But you may be able to see that in the [s] the center of the tongue is slightly hollowed while the tip of the tongue is not quite so close to the alveolar ridge. Also, during the [s] the teeth are closer together than during the [d]. Much of the sound of [s] is produced by a jet of air striking the edges of the teeth. The rapidly moving airstream is formed by the narrow gap between the tongue and the alveolar ridge.

CD 3.2



## STOP CONSONANTS

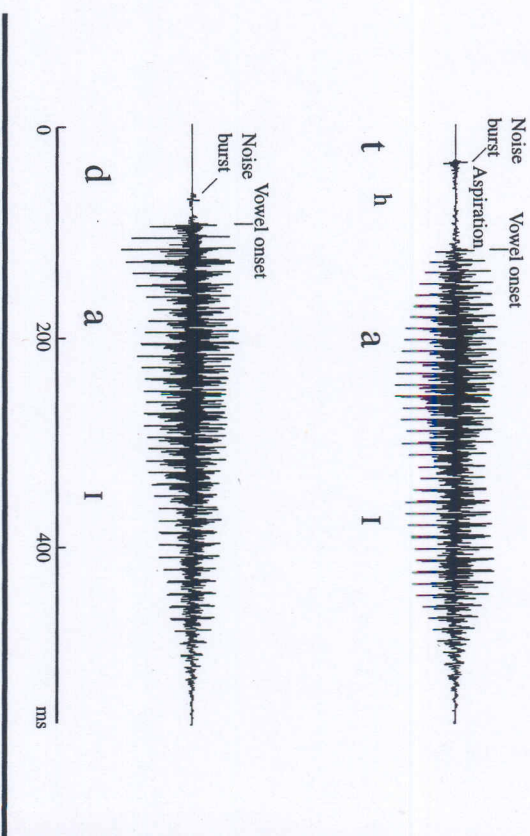
Consider the difference between the words in the first column in Table 3.1 and the corresponding words in the second column. This opposition may be said to be between the set of voiceless stop consonants and the set of voiced stop consonants. But the difference is really not just one of voicing during the consonant closure, as you can see by saying these words yourself. Most people have very little voicing going on while the lips are closed during either *pie* or *buy*. Both stop consonants are essentially voiceless. But in *pie*, after the release of the lip closure, there is a moment of **aspiration**, a period of voicelessness after the stop articulation and before the start of the voicing for the vowel. If you put your hand in front of your lips while saying *pie*, you can feel the burst of air that comes out during the period of voicelessness after the release of the stop.

In a narrow transcription, aspiration may be indicated by a small raised *h*, [ʰ]. Accordingly, these words may be transcribed as [pʰaɪ, tʰaɪ, kʰaɪ]. You may not be able to feel the burst of air in *tie* and *kye* because these stop closures are made well inside the mouth cavity. But listen carefully and notice that you can hear the period of voicelessness after the release of the stop closure in each of the words. It is this interval that indicates that the stop is aspirated. The major difference between the words in the first two columns is not that one has voiceless stops and the other voiced stops. It is that the first column has (voiceless) aspirated stops and the second column has (perhaps voiced) unaspirated stops. The amount of voicing in each of the stops [b, d, g] depends on the context in which it occurs. When it is in the middle of a word or phrase in which a voiced sound occurs on either side (as in column 3 in Table 3.1), voicing usually occurs throughout the stop closure. But most speakers of English have no voicing during the closure of so-called voiced stops in sentence initial position, or when they occur after a voiceless sound as in *that boy*.

One of the main objects of this book is to teach you to become a phonetician by learning to listen very carefully. You should be able to hear these differences, but you can also see them in computer-produced waveforms. Figure 3.1 is a record of the words *tie* and *die*. It is quite easy to see the different segments in the sound wave. In the first word, *tie*, there is a spike indicating the burst of noise which occurs when the stop closure is released, followed by a period of very small semi-random variations during the aspiration, and then a regular, repeating

TABLE 3.1 Words illustrating allophones of English stop consonants.

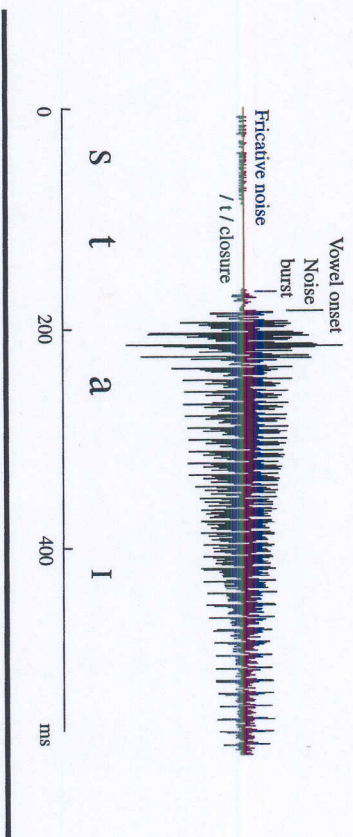
1	2	3	4	5	6
pie	buy	a buy	spy	nep	nab
tie	dye	a dye	sly	mat	mad
kye	guy	a sky	sky	knack	nag

FIGURE 3.1 The waveforms of the words *tie* and *die*.

wave as the vocal folds begin to vibrate for the vowel. In *die* there is no real spike at the beginning, just a somewhat irregular waveform before the start of the wave for the vowel. As you can see, the major difference between *tie* and *die* is the increase in time between the release of the stop and the start of the vowel. We will discuss this distinction further in Chapter 6.

Now consider the words in the fourth column of Table 3.1. Are the sounds of the stop consonants more like those in the first column or those in the second? As in many cases, English spelling is misleading, and the sounds are in fact more like those in the second column. There is no opposition in English between words beginning with /sp/ and /sb/, or /st/ and /sd/, or /sk/ and /sg/. English spelling has words beginning with *sp*, *st*, *sc*, or *sk*, and none that begin with *sb*, *sd*, or *sg*, but the stops that occur after /s/ are really somewhere in between initial /p/ and /b/, /t/ and /d/, /k/ and /g/, and usually more like the so-called voiced stops /b, d, g/ in that they are completely unaspirated. Figure 3.2 shows the computer-produced waveform in *sty*. You can see the small variations in the waveform corresponding to the fricative /s/, followed by a straight line during the period in which there is no sound because there is a complete stop for the /t/. This is followed by a sound wave very similar to that of the /d/ in Figure 3.1. If you have access to a computer that can record sounds and let you see the waveforms of words, you can verify this for yourself. (The freeware program WaveSurfer, included on the CD, will let you do this.) Record words such as *spy*, *sty*, *sky*, *spill*, *still*, *skill*, each said as a separate word. Now find the beginning and end of each /s/, and cut this part out. When you play the edited recordings to



FIGURE 3.2 The waveform of the word *sty*.

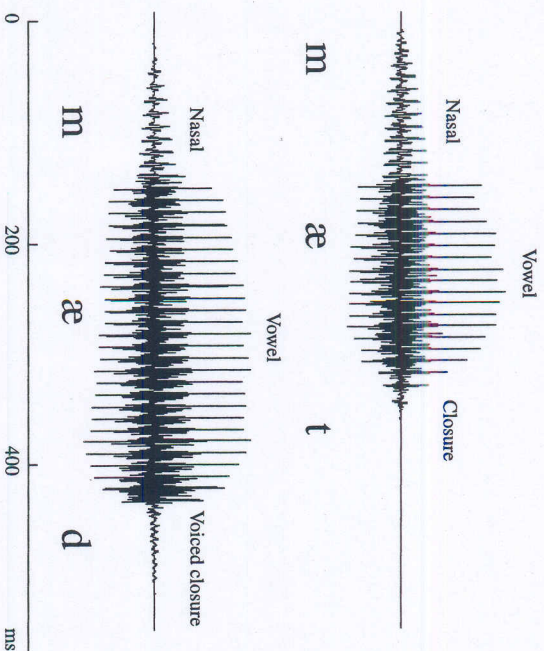
others and ask them to write down the words they hear, they will almost certainly write *buy*, *die*, *guy*, *bill*, *dill*, *gill*.

What about the differences between the words in the fifth and sixth columns? The consonants at the end of *nap*, *mat*, *knack* are certainly voiceless. But if you listen carefully to the sounds at the end of the words *nab*, *mad*, *nag*, you may find that the so-called voiced consonants /b, d, g/ have very little voicing and might also be called voiceless. Try saying these words separately. You can, of course, say each of them with the final consonant exploded and a short vowel-like sound afterward. But it would be more normal to say each of them without exploding the final consonants, or at least without anything like a vowel. You could even say *cab* and not open your lips for a considerable period of time if it were the last word of an utterance. In such circumstances it is quite clear that the final consonants are not fully voiced throughout the closure.

There is, however, a clear distinction between the words in the fifth and sixth columns. Say these words in pairs, *nap*, *nab*; *mat*, *mad*; *knack*, *nag*, and try to decide which has the longer vowel. In these pairs, and in all similar pairs, such as *cap*, *cab*; *cat*, *cad*; *back*, *bag* the vowel is much shorter before the voiceless consonants /p, t, k/ than it is before the voiced consonants /b, d, g/. The major difference between such pairs of words is in the vowel length, not in the voicing of the final consonants.

You can hear that both speakers on the CD also distinguish these words by vowel length. In these recordings each of the speakers said the words *nap*, *nab*; *mat*, *mad*; *knack*, *nag* in the same phonetic context, *I'll say \_\_\_ again*. By saying each word in a separate sentence it's easier to give each of them the same stress and intonation, and thus avoid the influence of these factors on the length of a word.

This length difference is very evident in Figure 3.3, which shows the waveforms of the words *mat* and *mad*. In this occasion the vowel in *mad* is almost twice as long as the vowel in *mat*. You can see small voicing vibrations during the /d/ in *mad*, but there is nothing noteworthy at the end of *mat* except the

FIGURE 3.3 The waveforms of the words *mat* and *mad*.

slightly irregular voicing at the time of the closure. We will return to this point later in this section.

Try comparing the length differences in short sentences such as *Take a cap now* and *Take a cab now*. If you say these sentences with a regular rhythm, you will find that the length of time between *Take* and *now* is about the same in both. This is because the whole word *cap* is only slightly shorter than the whole word *cab*. The vowel is much shorter in *cap* than in *cab*. But the consonant /p/ makes up for this by being slightly longer than the consonant /b/. It is a general rule of English (and of most other languages) that syllable final voiceless consonants are longer than the corresponding voiced consonants after this same vowel.

The phrases *Take a cap now* and *Take a cab now* also illustrate a further point about English stop consonants at the end of a word (or, in fact, at the end of a stressed syllable). Say each of these phrases without a pause before *now*. Do your lips open before the [n] of *now* begins, or do they open during the [n]? If they open before the [n], there will be a short burst of aspiration or a short vowel-like sound between the two words. Exploding the stops produces a somewhat unnatural pronunciation. Generally, final stops are unexploded when the next word begins with a nasal. The same is true if the next word begins with a stop. The final [t] in *cat* is nearly always unexploded in phrases like *the cat pushed...* In a narrow transcription we can symbolize the fact that a consonant is unexploded by adding a small raised mark [̚]. We could therefore transcribe the phrase as [t̚ə 'kæt̚ pʰʊʃt̚].



The same phenomenon occurs even within a word such as *apt* [æp't] or *act* [ækt]. Furthermore, across a word boundary the two consonants involved can even be identical, as in the phrase *white teeth*. To convince yourself that there are two examples of /t/ in this phrase, try contrasting it with *why teeth*. Not only is the vowel in *white* much shorter than the vowel in *why* (because the vowel in *white* is in a syllable with a voiceless consonant at the end), but also the stop closure in *white teeth* is much longer than the stop in the phrase with only one /t/. In *white teeth* there really are two examples of /t/ involved, the first of which is unexploded.

Other languages do not have this rule. For example, it is a mark of speakers with an Italian accent (at least as caricatured in films and on television) that they explode all their final stop consonants, producing an extra vowel at the end, as they normally would in their own language. Authors trying to indicate an Italian speaking English will write the sentence *It's a big day* as *It's a bigga day*. They are presumably trying to indicate the difference between the normal [It's a 'big'del] and the foreign accent [It's a 'biggə'del].

It is interesting that words such as *rap*, *rat*, *rack* are all distinguishable, even when the final consonants are unexploded. The difference in the sounds must therefore be in the way that the vowels end—after all, the rest is silence. The consonants before and after a vowel always affect it, so there is a slight but noticeable difference in its quality. Compare your pronunciation of words such as *pip*, *tit*, *kick*. Your tongue tip is up throughout the word *tit*, whereas in *pip* and *kick* it stays behind the lower front teeth. In *kick* it is the back of the tongue that is raised throughout the word, and in *pip* the lip gestures affect the entire vowel. The same is true for words with voiced consonants, such as *bib*, *did*, *gig*. The consonant gestures are superimposed on the vowel in such a way that their effect is audible throughout much of the syllable.

The sounds [p, t, k] are not the only voiceless stops that occur in English. Many people also pronounce a glottal stop in some words. A glottal stop is the sound (or, to be more exact, the lack of sound) that occurs when the vocal folds are held tightly together. As we have seen, the symbol for a glottal stop is [ʔ], resembling a question mark without the dot.

Glottal stops occur whenever one coughs. You should be able to get the sensation of the vocal folds being pressed together by making small coughing noises. Next, take a deep breath and hold it with your mouth open. Listen to the small plosive sound that occurs when you let the breath go. Now, while breathing out through your mouth, try to check and then release the breath by making and releasing a short glottal stop. Then do the same while making a voiced sound such as the vowel [a]. Practice producing glottal stops between vowels, saying [aʔa] or [iʔi], so that you get to know what they feel like.

One of the most common occurrences of a glottal stop is in the utterance meaning *no*, often spelled *nh-nh*. If someone asks you a question, you can reply *no* by saying [ʔaʔa] (usually with a nasalized vowel, which we will symbolize

later). Note that there is a contrast between the utterance meaning *no* and that meaning *yes* that is dependent on the presence of the glottal stop. If you had meant to say *yes*, you might well have said [ʔha]. We can tell that it is the glottal stop that is important in conveying the meaning by the fact that one could be understood equally well by using a syllabic consonant (shown by putting the mark [ ] under the consonant) instead of a vowel, and saying [ʔnhm] for *yes* and [ʔnm] for *no*. As long as there is a glottal stop between the two syllables, the utterance will mean *no*, irrespective of what vowel or nasal is used.

Glottal stops frequently occur as allophones of /t/. Probably most Americans and many British speakers have a glottal stop followed by a syllabic nasal in words such as *beaten*, *kitten*, *fatten* [bɪt̚n, kɪt̚n, fæt̚n]. London Cockney and many forms of Estuary English also have a glottal stop between vowels, as in *butter*, *kitty*, *father* [bʌʔə, kɪt̚r, fæʔə]. Many speakers in both Britain and America have a glottal stop just before final voiceless stops in words such as *rap*, *rat*, *rack*. Usually the articulatory gesture for the other stop is still audible, so these words could be transcribed [ræʔp, ræʔt, ræʔk]. When I recorded the word *nut* for Figure 3.3, I pronounced it as [mæʔt], with the glottal stop and the closure for [t] occurring almost simultaneously.

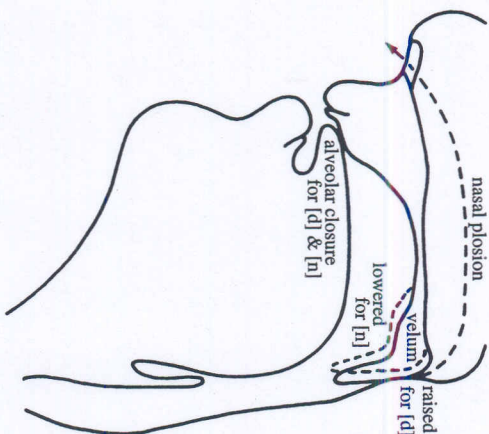
Practice producing words with and without a glottal stop. After you have some awareness of what a glottal stop feels like, try saying the words *rap*, *rat*, *rack* in several different ways. Begin by saying them with a glottal stop and a final release [ræʔp̚, ræʔt̚, ræʔk̚]. Next, say them without a glottal stop and with the final stops unexploded [ræp̚, ræt̚, ræk̚]. Then say them with a glottal stop and a final unexploded consonant [ræʔp̚, ræʔt̚, ræʔk̚]. Finally, say them with a glottal stop and no other final consonant [ræʔ, ræʔ, ræʔ].

When a voiced stop and a nasal occur in the same word, as in *hidden*, the stop is not released in the usual way. Both the [d] and the [n] are alveolar consonants. The tongue comes up and contacts the alveolar ridge for [d] and stays there for the nasal, which becomes syllabic [hɪdn̩]. Consequently, as shown in Figure 3.4, the air pressure built up behind the stop closure is released through the nose by the lowering of the soft palate (the velum) for the nasal consonant. This phenomenon, known as **nasal plosion**, is normally used in pronouncing words such as *sadden*, *swadden*, *leaden* [ˈsædn̩, ˈsɑdn̩, ˈledn̩]. It is considered a mark of a foreign accent to add a vowel [ˈsædn̩, ˈsɑdn̩, ˈledn̩]. Nasal plosion also occurs in the pronunciation of words with [t] followed by [n], as in *kitten* [kɪtn̩], for those people who do not have a glottal stop instead of the [t], but the majority of speakers of English pronounce this word with a glottal stop [kɪt̚n̩].

It is worth spending some time thinking exactly how you and others pronounce words such as *kitten* and *button*, in that it enables you to practice making detailed phonetic observations. There are a number of different possibilities. Most British and American English speakers make a glottal stop at the end of the vowel, before making an alveolar closure. Then, while still maintaining the glottal stop, they lower the velum and raise the tongue for the alveolar closure. But



FIGURE 3.4 Nasal plosion.



which comes first? If they lower the velum before making the alveolar closure, there is only [ɲn] and no [t]. If they make the alveolar closure first, we could say that there is [ɲn], but there would not be any nasal plosion, as there would be no pressure built up behind the [t] closure. Nasal plosion occurs only if there is no glottal stop, or if the glottal stop is released after the alveolar closure has been made and before the velum is lowered.

These are fairly difficult sequences to determine, but there are some simple things you can do to help you find out what articulations you use. First of all, find a drinking straw and something to drink. Put one end of the straw between your lips and hold the other end just (and only just) below the surface of the liquid. Now say [ɔpə], and note how bubbles form during [p]. This is because pressure is built up behind your closed lips. Now push the straw slightly farther into your mouth and say [ɑtə]. It will not sound quite right because the straw gets in the way of your tongue when it makes the alveolar closure. You may have to try different positions of the straw. Go on until you can see bubbles coming out, and convince yourself that pressure builds up behind the [t]. Now try saying *button*. Of course there will be bubbles during the [b], but are there any at the end of the word, or do you have a glottal stop and no [t] behind which pressure builds up?

When two sounds have the same place of articulation, they are said to be **homorganic**. Thus the consonants [d] and [n], which are both articulated on the alveolar ridge, are homorganic. For nasal plosion to occur within a word, there must be a stop followed by a homorganic nasal. Only in these circumstances can there be pressure first built up in the mouth during the stop and then

released through the nose by lowering the soft palate. Many forms of English do not have any words with a bilabial stop [p] or [b] followed by the homorganic nasal [m] at the end of the word. Nor in most forms of English are there any words in which the velar stops [k] or [g] are normally followed by the velar nasal [ŋ]. Consequently, both bilabial and velar nasal plosion are less common than alveolar nasal plosion in English. But when talking in a rapid conversational style, many people pronounce the word *open* as [ˈoʊpɪn], particularly if the next word begins with [m], as in *open my door, please*. Quite frequently when counting, people will pronounce *seven* as [ˈsɛvən], and I have also heard *something, captain, bacon* pronounced [ˈsʌmpn̩ kæptɪn, ˈbeɪkŋ]. You should try to pronounce all these words in these ways yourself.

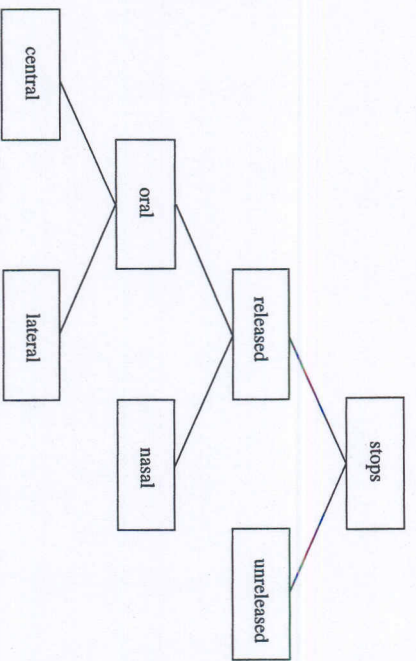
A phenomenon similar to nasal plosion may take place when an alveolar stop [t] or [d] occurs before a homorganic lateral [l] as in *little, ladle* [ˈlɪtl̩, ˈleɪdl̩]. The air pressure built up during the stop can be released by lowering the sides of the tongue; this effect is called **lateral plosion**. Say the word *middle* and note the action of the tongue. Many people (particularly British speakers) maintain the tongue contact on the alveolar ridge through both the stop and the lateral, releasing it only at the end of the word. Others (most Americans) pronounce a very short vowel in the second syllable. For those who have lateral plosion, no vowel sound occurs in the second syllables of *little, ladle*. The final consonants in all these words are syllabic. There may also be lateral plosion in words such as *Atlantic*, in which the [t] may be resyllabified so that it is at the beginning of the stressed (second) syllable. We should also note that most Americans, irrespective of whether or not they have lateral plosion, do not have a voiceless stop in *little*. There is a general rule in American English that whenever /t/ occurs after a stressed vowel and before an unstressed syllable other than [ŋ], it is changed into a voiced sound. For those Americans who have lateral plosion, this will be the stop [d].

This brings us to another important point about coronal stops and nasals. For many speakers, including most Americans, the consonant between the vowels in words such as *city, better, writer* is not really a stop but a quick tap in which the tongue tip is thrown against the alveolar ridge. This sound can be written with the symbol [ɾ] so that *city* can be transcribed as [ˈsɪɾi]. Many Americans also make this kind of tap when /d/ occurs after a stressed vowel and before an unstressed vowel. As a result they do not distinguish between pairs of words such as *latter* and *ladder*. But some maintain a distinction by having a shorter vowel in words such as *latter* that have a voiceless consonant in their underlying form. It is as if the statement that vowels are shorter before voiceless consonants had applied first, and then a later rule was applied changing [t] into [ɾ] when it occurred between a stressed and an unstressed syllable.

We can summarize the discussion of stop consonants by thinking of the possibilities there are in the form of a branching diagram, as shown in Figure 3.5. The first question to consider is whether the gesture for the stop is released



FIGURE 3.5 Stop consonant releases.



(exploded) or not. If it is released, then is it oral plosion or is the release due to the lowering of the velum, with air escaping through the nose, making it nasal plosion? If it is oral plosion, then is the closure in the mouth entirely removed, or is the articulation in the midline retained and one or both sides of the tongue lowered so that air escapes laterally? You should be able to produce words illustrating all these possibilities. For coronal stops, there is an additional point not shown in Figure 3.5, namely, is the [t] or [d] sound produced as a tap [ɾ]?

## FRICATIVES

The fricatives of English vary less than the stop consonants, yet the major allophonic variations that do occur are in many ways similar to those of the stops. Earlier we saw that when a vowel occurs before one of the voiceless stops /p, t, k/, it is shorter than it would be before one of the voiced stops /b, d, g/. The same kind of difference in vowel length occurs before voiceless and voiced fricatives. The vowel is shorter in the first word of each of the pairs *strife*, *strive* [straɪf, straɪv]; *teeth*, *teethe* [tiθ, tið]; *rice*, *rise* [raɪs, raɪz]; *mission*, *vision* [ˈmɪʃn, ˈvɪʒn].

Stops and fricatives are the only English consonants that can be either voiced or voiceless. Consequently, we can revise our statement that vowels are shorter before voiceless stops than before voiced stops. Instead, we can say that vowels are shorter before all voiceless consonants than before all voiced consonants. In this way we can capture a linguistically significant generalization that would have been missed if our statements about English had included two separate statements, one dealing with stops and the other dealing with fricatives.

We also saw that a voiceless stop at the end of a syllable (as in *hit*) is longer than the corresponding voiced stop (as in *hid*). Similarly, the voiceless fricatives are longer than their voiced counterparts in each of the pairs *safe*, *save* [seɪf, seɪv], *lace*, *laze* [leɪs, leɪz], and all the other pairs of words we have been discussing in this section. Again, because fricatives behave like stops, a linguistically significant generalization would have been missed if we had regarded each class of consonants completely separately.

Fricatives are also like stops in another way. Consider the degree of voicing that occurs in the fricative at the end of the word *ooze*, pronounced by itself. In most pronunciations, the voicing that occurs during the final [z] does not last throughout the articulation but changes in the last part to a voiceless sound like [s]. In general, voiced fricatives at the end of a word, as in *prove*, *smooth*, *choose*, *rouge* [pruːv, smuð, ʃuːz, ruːʒ], are voiced throughout their articulation only when they are followed by another voiced sound. In a phrase such as *prove it*, the [v] is fully voiced because it is followed by a vowel. But in *prove two times two is four* or *try to improve*, where the [v] is followed by a voiceless sound [t] or by a pause at the end of the phrase, it is not fully voiced.

Briefly stated, then, fricatives are like stops in three ways. First, stops and fricatives influence vowel length in similar ways—vowels before voiceless stops or fricatives are shorter than before voiced stops or fricatives. Second, final voiceless stops and fricatives are longer than final voiced stops and fricatives. Third, the final stops and fricatives classified as voiced are not actually voiced throughout the articulation unless the adjacent sounds are also voiced. In addition, both these types of articulation involve an obstruction of the airstream. Because they have an articulatory feature in common and because they act together in phonological statements, we refer to fricatives and stops together as a natural class of sounds called **obstruents**.

However, fricatives do differ from stops in that they sometimes involve actions of the lips that are not immediately obvious. Try saying *fin*, *thin*, *sin*, *shin* [fɪn, θɪn, sɪn, ʃɪn]. There is clearly a lip action in the first word as it involves the labiodental sound [f]. But do your lips move in any of the other three words? Most people find that their lips move slightly in any word containing /s/ (*sin*, *kiss*) and quite considerably in any word containing /ʃ/ (*shin*, *wish*), but that there is no lip action in words containing /θ/ (*thin*, *teeth*). There is also lip movement in the voiced sounds corresponding to /s/ and /ʃ/, namely /z/ as in *zeal*, *zeer* and /ʒ/ in *leisure*, *treasure*, but none in /ð/, as in *that*, *teethe*.

The primary articulatory gesture in these fricatives is the close approximation of two articulators so that friction can be heard. The lip rounding is a lesser articulation in that the two articulators (the lower lip and the upper lip) approach one another but not sufficiently to cause friction. A lesser degree of closure by two articulators not involved in the primary articulation is called a secondary articulation. This particular one, in which the action of the lips is added to another articulation, is called labialization. The English fricatives /ʃ, ʒ/ are strongly labialized, and the fricatives /s, z/ are slightly labialized.



## AFFRICATES

This is a convenient place to review the status of affricates in English. An affricate is simply a sequence of a stop followed by a homorganic fricative. Some such sequences, for example the dental affricate [tθ] as in *eight* or the alveolar affricate [ts] as in *cats*, have been given no special status in English phonology. They have been regarded just as consonant clusters comparable with those at the end of *lapse* and *sacks* (which are not affricates, as the stops and the fricatives are not homorganic). But, as we noted in the discussion of symbols for transcribing English, it is appropriate to regard the sequences [tʃ] and [dʒ] as different from other sequences of consonants. They are the only affricates in English that can occur at both the beginning and the end of words. In fact, even the other affricates that can occur at the end of words will usually do so only as the result of the formation of a plural or some other suffix, as in *eighth*. From the point of view of a phonologist considering the sound pattern of English, the palato-alveolar affricates are plainly single units, but [ts] as in *cats* is simply a sequence of two consonants.

## NASALS

The nasal consonants of English vary even less than the fricatives. Nasals, together with [r, l], can be syllabic when they occur at the end of words. As we have seen, the mark [ ] under a consonant indicates that it is syllabic. (Vowels, of course, are always syllabic and therefore need no special mark.) In a narrow transcription, we may transcribe the words *sadden*, *table* as [sædn̩, 'teɪb̩]. In most pronunciations, *prism*, *prison* can be transcribed ['prɪzn̩, 'prɪzn̩], as these words do not usually have a vowel between the last two consonants. Syllabic consonants can also occur in phrases such as *Jack and Kate* ['dʒæk ŋ 'keɪt].

The nasal [ŋ] differs from the other nasals in a number of ways. No English word can begin with [ŋ]. This sound can occur only within or at the end of a word, and even in these circumstances it does not behave like the other nasals. It can be preceded only by the vowels /ɪ, e, æ, ʌ/ and /ɑ/ (American English) or /ɒ/ (British English), and it cannot be syllabic (except in slightly unusual pronunciations, such as *bacon* ['beɪkŋ], and phrases such as *Jack and Kate* mentioned above).

One way to consider the different status of [ŋ] is that it is a sequence of the phonemes /n/ and /g/. Looking at it this way, *sing* would be phonemically /sɪŋ/ and *sink* would be /sɪnk/. There would be a rule saying that /n/ has the allophone [ŋ] whenever it occurs before /g/ and /k/, turning [sɪŋ] into [sɪŋ] and [sɪnk] into [sɪŋk]. Another rule would allow for the dropping of /g/ (but not of /k/) whenever it occurs after [ŋ] at the end of either a word (as in *sing*) or a stem followed by a suffix such as *-er* or *-ing*. In this way the /g/ would be

dropped in *singer* which contains a suffix *-er* but is retained in *finger* in which the *-er* is not a suffix. This second rule would not apply in the case of some speakers from the New York area who make *singer* rhyme with *finger*.

## APPROXIMANTS

The voiced approximants are /w, r, j, l/ as in *whack*, *rack yak*, *lack*. The first three of these sounds are central approximants, and the last is a lateral approximant. The articulation of each of them varies slightly depending on the articulation of the following vowel. You can feel that the tongue is in a different position in the first sounds of *we* and *water*. The same is true for *rap* and *raw*, *lee* and *law*, and *ye* and *yaw*. Try to feel where your tongue is in each of these words.

These consonants also share the possibility of occurring in consonant clusters with stop consonants. The approximants /r, w, l/ combine with stops in words such as *pray*, *bray*, *tray*, *dray*, *Cray*, *gray*, *twirl*, *dwelt*, *quell*, *Gwen*, *play*, *blade*, *clay*, *glaze*. The approximants are largely voiceless when they follow one of the voiceless stops /p, t, k/ as in *play*, *twice*, *clay*. This voicelessness is a manifestation of the aspiration that occurs after voiceless stops, which we discussed at the beginning of this chapter. At that time I introduced a small raised *h* symbol [ʰ], which could be used to show that the first part of the vowel was voiceless. When there is no immediately following vowel, we can use the diacritic [◌̥] to indicate a voiceless sound. We can transcribe the words *play*, *twice*, *clay*, in which there are approximants after initial voiceless plosives, as [pl̥eɪ, t̥waɪs, kleɪ]. The approximant /j/ as in *you* [ju] can occur in similar consonant clusters, as in *pew*, *cue* [pj̥u, kj̥u], and, for speakers of British English, *tune* [tj̥n̩]. We will discuss the sequence [ju] again when we consider vowels in more detail.

In most forms of British English, there is a considerable difference in the articulation of /l/ before a vowel or between vowels, as in *leaf* or *feeling*, as compared with /l/ before a consonant or at the end of a word, as in *field* or *feel*. In most forms of American English, there is less distinction between these two kinds of /l/. Note the articulation of /l/ in your own pronunciation. Try to feel where the tongue is during the /l/ in *leaf*. You will probably find that the tip is touching the alveolar ridge, and one or both sides are near the upper side teeth, but not quite touching. Now compare this articulation with the /l/ in *feel*. Most (but not all) speakers make this sound with the tongue tip touching the alveolar ridge. But in both British and American English the center of the tongue is pulled down and the back is arched upward as in a back vowel. If there is contact on the alveolar ridge, it is the primary articulation. The arching upward of the back of the tongue forms a secondary articulation, which we will call velarization. In most forms of American English, all examples of /l/ are comparatively velarized, except, perhaps, those that are syllable initial and between high front vowels, as



in *freely*: In British English /l/ is usually not velarized when it is before a vowel, as in *lamb* or *swelling*, but it is velarized when word final or before a consonant, as in *ball* or *filled*. Also compare the velarized /l/ in *Don't kill dogs* with the one in *Don't kill it*. Most people don't have a velarized /l/ in *kill it*, despite the fact that it is seemingly at the end of a word. This is because the *it* in *kill it* acts like a suffix (technically a clitic), just like the suffix *-ing* in *killing*. (Note: The differences between the two types of /l/ are more noticeable in British English. American English examples of the phenomena cited above are not included on the CD.)

One symbol for velarization is the mark [~] through the middle of the symbol. Accordingly, a narrow transcription of *feel* would be [fi:~l]. In my own speech, the whole body of the tongue is drawn up and back in the mouth so that the tip of the tongue no longer makes contact with the alveolar ridge. Strictly speaking, therefore, this sound is not an alveolar consonant but more like some kind of back vowel in the speech of some English speakers.

Finally, we must consider the status of /h/. Earlier I suggested that English /h/ is the voiceless counterpart of the surrounding sounds. At the beginning of a sentence /h/ is like a voiceless vowel, but /h/ can also occur between vowels in words or phrases like *behind the head*. As you move from one vowel through /h/ to another, the articulatory movement is continuous, and the /h/ is signaled by a weakening of the voicing, which may not even result in a completely voiceless sound.

In many accents of English, /h/ can occur only before stressed vowels or before the approximant /j/ as in *hue* [hju:]. Some speakers of English also sound /h/ before /w/, so that they contrast *which* [hwɪtʃ] and *wich* [wɪtʃ]. The symbol [w] (an inverted *w*) is sometimes used for this voiceless approximant. The contrast between /w/ and /w/ is disappearing in most forms of English. In those dialects in which it occurs, [w] is more likely to be found only in the less common words such as *whether* rather than in frequently used words such as *what*.

## OVERLAPPING GESTURES

All the sounds we have been considering involve movements of the articulators. They are often described in terms of the articulatory positions that characterize these movements. But, rather than thinking in terms of static positions, we should really consider each sound as a movement. This makes it easier to understand the overlapping of consonant and vowel gestures in words such as *bib*, *did*, *gig*, mentioned earlier in this chapter. As we noted, in the first word, *bib*, the tongue tip is behind the lower front teeth throughout the word. In the second word, *did*, the tip of the tongue goes up for the first /d/ and remains close to the alveolar ridge during the vowel so that it is ready for the second /d/. In the third word, *gig*, the back of the tongue is raised for the first /g/ and remains near the soft palate during the vowel. In all these cases the gestures for the vowels and consonants overlap.

The same kind of thing happens with respect to gestures of the lips. Lip rounding is an essential part of /w/. Because there is a tendency for gestures to overlap with those for adjacent sounds, stops are slightly rounded when they occur in clusters in which /w/ is the second element, as in *twice*, *dwindle*, *quick* [kwɪks], *dwindle*, *kwɪk*]. This kind of gestural overlapping, in which a second gesture starts during the first gesture, is sometimes called anticipatory coarticulation. The gesture for the approximant is anticipated during the gesture for the stop. In many people's speech /r/ also has some degree of lip rounding. Try saying words such as *read* and *head*. Do you get some movement of the lips in the first word but not in the second? Use a mirror to see whether you get anticipatory lip rounding for the stops [t, d] so that they are slightly rounded in words such as *tree* and *dream*, as opposed to *tee* and *deem*.

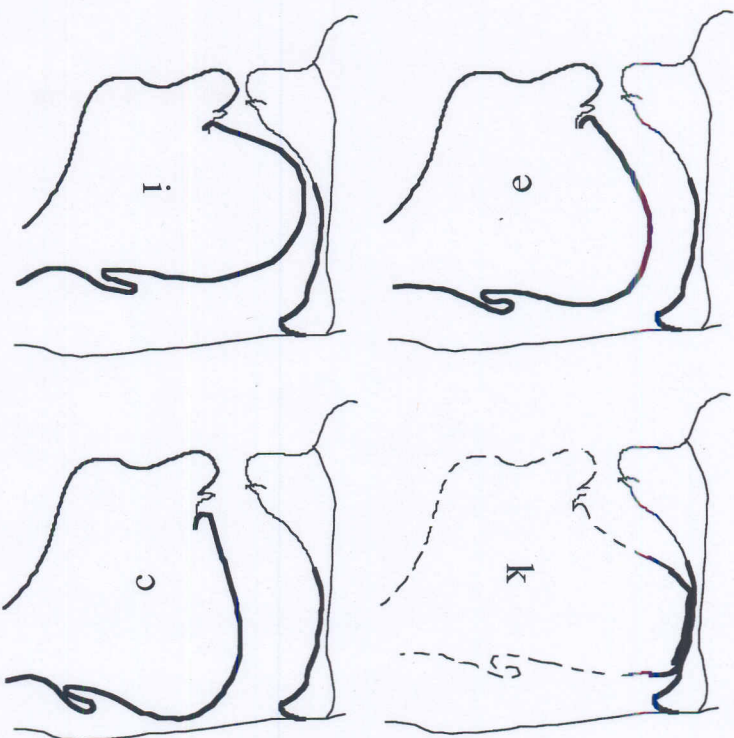
We can often think of the gestures for different articulations as movements towards certain targets. A target is something that one aims at but does not necessarily hit, perhaps because one is drawn off by also having to aim at another target. Ideally, the description of an utterance might consist of the specification of a string of target gestures that must be made one after another. It's important in any science to be able to give numerical specifications. We should be able to state the articulatory targets in terms of numerical values indicating the endpoints of the gestures of the tongue and lips.

The data in Figure 3.6 indicate how we can specify graphically, and hence measurably, the articulatory targets for the four English sounds /ə/, k, i, ɔ/, the four sounds used in phrases like *a key* /ski/, *a cow* /sko/, and (in a British accent) *occur* /sko/. Similar diagrams could be drawn for each of the other English sounds that have separate targets. The first point to note is that the target specifies only a part of the tongue shape for /k/. Some parts of the tongue are left out because their movements are not specified for the gesture for /k/. The thickness of the line shows how much it matters for that part of the tongue to achieve the /k/ target shape. What is most important for /k/ is the position of the back of the tongue. For the vowels, what matters most is the position of the tongue in the region where the vocal tract is most constricted, but the other parts of the tongue still need be approximately specified, as indicated by the thinner line. When saying a phrase such as *occur* /sko/, the parts of the tongue that are unspecified for /k/ will be in the target position for /ə/ throughout the phrase. When saying *a key* /ski/ or *a cow* /sko/, some parts of the tongue will move from /ə/ to /i/ or /ɔ/ without being affected by the /k/, but others will achieve as much of the target position for /k/ as possible.

Coarticulation between sounds will always result in the positions of some parts of the vocal tract being influenced largely, whereas others will not be so much affected by neighboring targets. The extent to which anticipatory coarticulation occurs depends on the extent to which the position of that part of the vocal tract is specified in the two targets. The degree of coarticulation also depends on the interval between them. For example, a considerable amount of lip rounding



FIGURE 3.6



occurs during [k] when the next sound is rounded as in *cool* [ku]. Slightly less lip rounding occurs if the [k] and the [u] are separated by another sound, as in *clue* [klu], and even less occurs if there is also a word boundary between the two sounds, as in the phrase *sack Lou* [sæklu]. Nevertheless, some rounding may occur, and sometimes anticipatory coarticulations can be observed over even longer sequences. In the phrase *tackle Lou* [tæklɪu], the lip rounding for the [u] may start in the [k], which is separated from it by two segments and a word boundary.

There is no simple relationship between the description of a language in terms of phonemes and the description of utterances in terms of gestural targets. A phoneme is an abstract unit that may be realized in several different ways. Sometimes the differences between the different allophones of a phoneme can be explained in terms of targets and overlapping gestures. The differences between the [k] in *key* and the [k] in *cow* are simply due to their overlapping with different vowels. Similarly, we do not have to specify separate targets for the alveolar [n] in *ten* and the dental [ŋ] in *tenth*. Both are the result of aiming

at the same target, but in *tenth* the realization of the phoneme /n/ is influenced by the dental target required for the following sound. However, the differences between some allophones are actually the result of aiming at different targets. For many American English speakers the initial [r] in *reed* is made with a tongue gesture that is very different from that for the final [r] in *deer*. In most forms of British English the [l] in *leaf* and the [l] in *feel* differ in ways that cannot be ascribed to coarticulation. Perhaps the most extreme example of the difference between phonemes and gestures is in the realization of the /t/ phoneme in *ten* [tʰɛn] and in *bitton* [bɪtʰn], in which the one phoneme is realized by two completely different gestures, [tʰ] and [ʔ]. Sometimes the differences between allophones are the result of overlapping gestures, producing what have been called intrinsic allophones, and sometimes they involve different gestures, which may be called extrinsic allophones. Because phonemes are composed of these two types of allophones, they cannot be equated with gestures.

To summarize, gestural targets are units that can be used in descriptions of how a speaker produces utterances. Phonemes are more abstract units that can be used in descriptions of languages to show how one word contrasts with another. Virtually all the gestures for neighboring sounds overlap. Differences in the timing of one gesture with respect to another account for a wide range of the phenomena that we observe in speech. The next section provides a number of additional examples.

### RULES FOR ENGLISH CONSONANT ALLOPHONES

A good way of summarizing (and slightly extending) all that I have said about English consonants so far is to list a set of formal statements or rules describing the allophones. These rules are simply descriptions of language behavior. They are not the kind of rules that prescribe what people ought to do. Like most phoneticians, I would not presume to set myself up as an arbiter of fashion and declare what constitutes “good” speech. But phonetics is part of an exact scientific discipline, and that means we should be able to formalize descriptions of speech in terms of a set of precise statements.

Some of the statements apply to all the consonants of English. One of the statements dealing with consonant length is the following:

- (1) Consonants are longer when at the end of a phrase.

You can see the application of this statement by comparing the consonants in words such as *bit*, *did*, *don*, *nod*. Use WaveSurfer (on the CD) to make a recording of these words, and then play the recording backward. Are the first two words the same backward and forward? Do the third and fourth words sound like each other when played in reverse?



Most of the formal statements apply to only selected groups of consonants, as in (2) through (18) below:

- (2) Voiceless stops (i.e., /p, t, k/) are aspirated when they are syllable initial, as in words such as *pip*, *test*, *kick* [p<sup>h</sup>ɪp, t<sup>h</sup>ɛst, k<sup>h</sup>ɪk].
- (3) Obstruents—stops and fricatives—classified as voiced (that is, /b, d, g, v, ð, z, ʒ/) are voiced through only a small part of the articulation when they occur at the end of an utterance or before a voiceless sound. Listen to the /v/ when you say *try to improve*, and the /d/ when you say *add two*.
- (4) So-called voiced stops and affricates /b, d, g, dʒ/ are voiceless when syllable initial, except when immediately preceded by a voiced sound (as in *a day* as compared with *this day*).
- (5) Voiceless stops /p, t, k/ are unaspirated after /s/ in words such as *spew*, *stew*, *skew*.
- (6) Voiceless obstruents /p, t, k, tʃ, f, θ, s, ʃ/ are longer than the corresponding voiced obstruents /b, d, g, dʒ, v, ð, z, ʒ/ when at the end of a syllable.

Words exemplifying this rule are *cap* as opposed to *cab* and *back* as opposed to *bag*. Try contrasting these words in sentences, and you may be able to hear the differences more clearly.

- (7) The approximants /w, r, j, l/ are at least partially voiceless when they occur after initial /p, t, k/, as in *plaw*, *twin*, *cue* [pl<sup>h</sup>ɛr, twɪn, kj<sup>h</sup>u].

This is due to the overlapping of the gesture required for aspiration with the voicing gesture required for the approximants. (Note that the formal statement says at least partially voiceless, but the transcription marks the approximants as being completely voiceless. Conflicts between statements and transcriptions of this kind will be discussed further below.)

- (8) The gestures for consecutive stops overlap, so that stops are unexploded when they occur before another stop in words such as *apt* [əp<sup>h</sup>t] and *ribbed* [rɪb<sup>h</sup>d].
- (9) In many accents of English, syllable final /p, t, k/ are accompanied by an overlapping glottal stop gesture, as in pronunciations of *tip*, *pit*, *kick* as [tɪp̚, pɪt̚, kɪk̚]. (This is another case where transcription cannot fully describe what is going on.)

This rule does not apply to all varieties of English. Some people do not have any glottal stops in these circumstances, and others have glottal stops completely replacing some or all of the voiceless stops. In any case, even for those who simply add a glottal stop, the statement is not completely accurate. Many

people will have a glottal stop at the end of *car* in phrases such as *that's a car* or *the car sat on the mat*, but they will not have this allophone of /t/ in *the cat eats fish*.

- (10) In many accents of English, /t/ is replaced by a glottal stop when it occurs before an alveolar nasal in the same word, as in *beaten* [ˈbi:t̚n].
- (11) Nasals are syllabic at the end of a word when immediately after an obstruent, as in *laden*, *chasm* [ˈlɛd̩n, ˈkæz̩m].

Note that we cannot say that nasals become syllabic whenever they occur at the end of a word and after a consonant. The nasals in *kiln*, *film* are not syllabic in most accents of English. We can, however, state a rule describing the syllabicity of /l/ by saying simply:

- (12) The lateral /l/ is syllabic at the end of a word when immediately after a consonant.

This statement summarizes the fact that /l/ is syllabic not only after stops and fricatives (as in *paddle*, *whistle* [ˈpædl̩, ˈwɪsl̩]), but also after nasals (as in *ken- nel*, *channel* [ˈkɛnl̩, ˈtʃænl̩]). The only problem with this rule is what happens after /r/. It is correct for words such as *barrel* [ˈbær̩l̩] but does not work in most forms of American English in words such as *snarl* [ˈsnɑrl̩], when /r/ has to be considered as part of the vowel.

When it is not part of the vowel, /r/ is like /l/ in most forms of American English in that it, too, can be syllabic when it occurs at the end of a word and after a consonant, as in *sabre*, *razor*, *hammer*, *tailor* [ˈsɛɪbr̩, ˈreɪzr̩, ˈhæmɹ̩, ˈteɪlɹ̩]. If we introduce a new term, **liquid**, which is used simply as a cover term for the consonants /l, r/, we may rephrase the statement in (12) and say:

- (12a) The liquids /l, r/ are syllabic at the end of a word when immediately after a consonant.

The next statement also applies more to American English than to British English. It accounts for the /t/ in “fatty, data” [ˈfæɹi, ˈdæɹə]. But note that these are not the only contexts in which these changes occur. This is not simply a change that affects /t/ after a stressed vowel and before an unstressed one, in that /t/ between two unstressed vowels (as in *divinity*) is also affected. However, not all cases of /t/ between vowels change in this way. The /t/ in *attack* (i.e., before a stressed syllable) is voiceless, and /t/ after another consonant (for example, in *haspy* and *caprive*) is also voiceless. Note also that most American English speakers have a very similar articulatory gesture in words containing /d/ and /n/ in similar circumstances, such as *daddy* and *many*. The first of these two words could well be transcribed [ˈdæɹɪ]. The second has the same sound, except that it is nasalized, so it could be transcribed [ˈmæɹ̩] in a narrow transcription.



Nasalization is shown by the diacritic [ ~ ] over a symbol. The following statement accounts for all these facts:

- (13) Alveolar stops become voiced taps when they occur between two vowels the second of which is unstressed.

Many speakers of American English require a similar rule to describe a sequence of an alveolar nasal followed by a stop. In words such as *painter* and *splinter* the /t/ is lost and a nasal tap occurs. This has resulted in *winter* and *winner* and *paning* and *panning* being pronounced in the same way. For these speakers we can restate (13), making it:

- (13a) Alveolar stops and alveolar nasal plus stop sequences become voiced taps when they occur between two vowels the second of which is unstressed.

There is a great deal of variation among speakers with respect to this statement. Some make taps in familiar words such as *antite*, but not in less common words such as *Dante*. Some make them only in fast speech. Try to formulate a statement in a way that describes your own speech.

- (14) Alveolar consonants become dentals before dental consonants, as in *eight*, *tenth*, *wealth* [eɪθ, tɛnθ, weɪθ]. Note that this statement applies to all alveolar consonants, not just stops, and often applies across word boundaries, as in *at this* [æt ðɪs]. This is a statement that in English the gestures for these two consonants overlap so much that the place of articulation for the first consonant is changed.

In a more rapid style of speech, some of these dental consonants tend to be omitted altogether. Say these words first slowly and then more rapidly, and see what you do yourself. It is difficult to make precise statements about when consonants get deleted, because this depends so much on the style of speech being used. Alveolar stops often appear to get dropped in phrases such as *fact finding*. Most people say *most people* as [ˈmʊs ˈpipl̩] with no audible [t], and they produce phrases such as *send papers* with no audible [d]. We could state this as follows:

- (15) Alveolar stops are reduced or omitted when between two consonants.

Formal statement (15) raises an interesting point of phonetic theory. Note that I said “alveolar stops often appear to get dropped,” and there may be “no audible [d]”. However, the tongue tip gesture for the alveolar stop in *most people* may not be omitted but just not audible because it is completely overlapped by the labial stop following. More commonly, it is partially omitted, that is to say, the tongue tip moves up for the alveolar stop but does not get to make a complete closure. When we think in terms of phonetic symbols, we can write [ˈmʊs ˈpipl̩] or [ˈmʊst ˈpipl̩]. This makes it a question of whether the [t] is there or not. But

that is not really the issue. Part of the tongue tip gesture may have been made, a fact that we have no way of symbolizing.

Check how you say phrases such as *best game* and *grand master*. Say these and similar phrases with and without the alveolar stop. You may find it difficult to formulate a statement that takes into account all the contexts where alveolar stops may not appear in your speech.

We must state not only where consonants get dropped, but also where they get added. Words such as *something* and *youngster* often get pronounced as [ˈsʌmpθŋ] and [ˈjʌŋkstɚ]. In a similar way, many people do not distinguish between *prince* and *prints*, or *tense* and *tents*. All these words may be pronounced with a short voiceless stop between the nasal and the voiceless fricative. But the stop is not really an added gesture. It is simply the result of changing the timing of the nasal gesture with respect to the articulatory gesture. By delaying the lowering of the velum for the nasal, a moment of complete closure—a stop—occurs. The apparent insertion of a stop into the middle of a word in this way is known as **epenthesis**. If we want to make a formal statement of this phenomenon we could say:

- (16) A homorganic voiceless stop may occur after a nasal before a voiceless fricative followed by an unstressed vowel in the same word.

Note that it is necessary to mention that the following vowel must be unstressed. Speakers who have an epenthetic stop in the noun *concern* do not usually have one in verbal derivatives such as *concerned*, or in words such as *concern*. Nothing need be said about the vowel before the nasal. Epenthesis may—like the [t] to [r] change in statement (13)—occur between unstressed vowels. I have heard an inserted [t] in both *agency* and *grievances*.

Statement (16) raises a theoretical point similar to that discussed in connection with (15), where we were concerned with whether or not a segment had been deleted. Now we are concerned with whether a segment has been added. In each case it is better to treat these as false questions and to think about the gestures involved rather than worry about the symbols that might or might not represent separate segments. It may be convenient to transcribe *something* as [ˈsʌmpθŋ], but transcription is only a tool and should not be thought of as necessarily portraying the units used in the production of speech.

The next statement accounts for the shortening effects that occur when two identical consonants come next to one another, as in *big game* and *top post*. It is usually not accurate to say that one of these consonants is dropped. There are two consonantal gestures, but they overlap considerably. Even in casual speech, most people would distinguish between *stray tissue*, *straight tissue*, and *strait tissue*. (Try saying these in sentences such as *That's a stray tissue* and see for yourself.) But there clearly is a shortening effect that we can state as follows:

- (17) A consonant is shortened when it is before an identical consonant.



TABLE 3.2 Some diacritics that modify the value of a symbol.

°	h									
Voiceless	Aspirated	Dental	Nasalized	Velarized	Syllabic	ɰ	ɸ	ɪ	kʰɪk, pɪeɪs	quick, place
						ɸ	ɪ	kʰ	tʰeɪp, kɪs	tap, kiss
						ɸ	ɪ	ɰ	æɪ(ð), heɪθ	at the, health
						ɸ	ɪ	æ̃	mæ̃n	man
						ɸ	ɪ	ɸ	pʰɪɪ	pill
						ɸ	ɪ	ɪ	ˈmɪ ɹŋ	mitten

We can describe the overlapping gestures that result in more front articulations of /k/ in *cap*, *kept*, *kit*, *key* [kæp, kept, kɪt, ki] and of /g/ in *gap*, *get*, *give*, *geese* [gæp, get, gɪv, gɪs]. You should be able to feel the more front position of your tongue contact in the latter words of these series. We can say:

(18) Velar stops become more front before more front vowels.

Finally, we need to note the difference in the quality of /l/ in *life* [laɪf] and *file* [faɪl], or *clap* [klaɪp] and *talcl* [tæɪk], or *feeling* [fiɪŋ] and *feel* [fi].

(19) The lateral /l/ is velarized when after a vowel or before a consonant at the end of a word.

Note that there are clearly distinct gestures required for /l/ in the different circumstances. These are not differences that can be ascribed to overlapping gestures.

### DIACRITICS

In this and the previous chapter we have seen how the transcription of English can be made more detailed by the use of diacritics, small marks added to a symbol to narrow its meaning. The six diacritics we have introduced so far are shown in Table 3.2. You should learn the use of these diacritics before you attempt any further detailed transcription exercises. Note that the nasalization diacritic is a small wavy line above a symbol, and the velarization diacritic is a similar wavy line through the middle of a symbol. Nasalization is more common among vowels, which will be discussed in the next chapter.

### EXERCISES

(Printable versions of all the exercises are available on the CD.)

A. The sequence of annotated diagrams below illustrates the actions that take place during the consonants at the end of the word *bench*. Fill in the blanks.

