The art of writing phonetically does not come of itself any more than the art of painting. The average English person can’t paint his own portrait, nor can he write his own language phonetically or anything approaching it . . . he is always thinking he says one thing when he really says something else; besides which . . . he is always wanting to write down what he thinks people ought to say (which has nothing to do with the case).

This chapter provides a whirlwind tour of the consonants of the world. The full IPA chart is introduced in Section 3.1, and Sections 3.2 and 3.3 work through it systematically, discussing the factors that combine to create each sound, as well as the combinations of factors that do not work. We begin with place and manner of pulmonic consonants, and then discuss non-pulmonic consonants, with notes on diacritics and other symbols included along the way. For most readers, most of the symbols in this chart will be unfamiliar, and the sounds they represent may seem strange. It might be tempting to call some of these sounds “exotic,” to be found, like rare orchids, only in the dense rainforests or remote mountaintops. Our tour of these sounds will in fact take us to every inhabited continent, and to some remote places. You may be surprised, however, to discover how many of these sounds are lurking in common English pronunciations, given the right environments. We finish the chapter, Section 3.4, with a discussion of some details of the pronunciation of English.

3.1 “exotic” sounds and the phonetic environment

Figures 3.1 and 3.2 provide the IPA symbols for consonants. These cover all the consonants of the languages of the world – at least it is the goal of the International Phonetic Association

<table>
<thead>
<tr>
<th>consonants (pulmonic)</th>
<th>bilabial</th>
<th>labiodental</th>
<th>dental</th>
<th>alveolar</th>
<th>post alveolar</th>
<th>retroflex</th>
<th>palatal</th>
<th>velar</th>
<th>uvular</th>
<th>pharyngeal</th>
<th>glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>plosive</td>
<td>p b</td>
<td></td>
<td>t d</td>
<td>t d</td>
<td>c j</td>
<td>k g</td>
<td>q g</td>
<td></td>
<td></td>
<td></td>
<td>?</td>
</tr>
<tr>
<td>nasal</td>
<td>m mj</td>
<td></td>
<td>n</td>
<td>n</td>
<td>n</td>
<td></td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>trill</td>
<td>b r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tap or flap</td>
<td>v r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fricative</td>
<td>φ β f v</td>
<td>θ ð s z</td>
<td>θ z</td>
<td>θ z</td>
<td>θ z</td>
<td>x y</td>
<td>x y</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lateral fricative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>approximant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lateral approximant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

Figure 3.1 IPA symbols for pulmonic consonants. Source: International Phonetic Association (Department of Theoretical and Applied Linguistics, School of English, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece).
that they do so. The chart has several different parts: pulmonic consonants (Figure 3.1) and non-pulmonic consonants, other symbols, and diacritics (Figure 3.2). (Vowels, suprasegmentals, and tone are covered in Chapter 4.) The most extensive is the chart of pulmonic consonants. Again, place of articulation runs across the top of the chart, and manner down the side, so each block represents a particular combination of place and manner. Where there are two symbols in a block, the one on the left is voiceless. Empty blocks indicate combinations that have not been found in any language; shaded empty blocks indicate combinations...
“deemed impossible.” The goal of this chapter is to explain what all of the symbols stand for, and why the shaded blocks are shaded.

Learning the IPA symbols is an important skill for anyone hoping to do linguistics. With these symbols, one has the ability to clearly indicate any speech sound in any language in a way that another linguist can understand. The sounds that these symbols represent, and the relationships between them, will form the subject matter of the rest of this book, so some knowledge of sound symbols is crucial to further progress.

More important than memorizing symbols, however, is understanding the categories of place and manner. Once these are understood, you can use the chart as a reference for unfamiliar (or at least un-memorized) symbols, and be able to figure out the sounds they represent, even if it takes some practice to produce them. But it may be the blank spaces that are most interesting and important of all, since they indicate the limits of the phonetic capabilities of human speakers. The symbols answer the question “What are the occurring sounds of human language?” but the blanks answer the question “What is a possible sound of human language?” Both symbols and blanks will be discussed in turn.

As was noted above, many of these exotic-seeming sounds occur in English, given the right environment. In linguistics, the term *environment* refers to the place in a word where a sound occurs (such as initial or final position), and to the sounds that surround it (between vowels, after an [s], etc.). Languages differ not only in the sounds that they use, but in the environments where sounds may occur. We’ve already seen several examples of this. In English, the velar nasal [ŋ] occurs in medial and final position (“sing,” “singer”), but never in word-initial position. Thai has no such restriction, and words like [ŋaː] (meaning *tusk*) are common. (The colon indicates a long vowel.) The word [ŋaː] may sound “exotic,” but there is no new sound there. You already know how to pronounce [ŋ]; you just need to learn to produce it and recognize it in an unfamiliar environment. The situation is similar with the glottal stop: in English, it may occur in place of [t] after [n]. In Hawai’ian, [ʔ] can occur anywhere [p] or [k] can.

We have also already seen examples of how the environment may introduce variation in the way a sound is pronounced. In English, the voiceless plosives [p, t, k] are aspirated [pʰ, tʰ, kʰ] in word-initial position, but not aspirated when they occur after [s]. Thai is like English in having both aspirated and unaspirated voiceless stops, but in Thai both kinds can occur in word-initial position: [kʰː:w] means *step*, and [kːː:w] means *rice*. (Again, the colon indicates a long vowel, and the circumflex accent indicates falling pitch.) When two sounds occur in the same environment, and create two different words (as [kʰː:w] and [kːː:w] in Thai), we can be confident that the two sounds are contrastive (a term introduced in Chapter 2). That is, the difference between the two sounds is the only thing that is marking the difference (that is, the contrast) between the two words. Aspiration is not contrastive in English: there are no two words that differ *only* in that one has an aspirated stop and one has a non-aspirated stop. There is always some other difference, like an initial [s], that provides the environment for aspiration to occur or not. The ideas of contrast and variation are central to the study of phonetics and phonology, and are discussed throughout this book. (They are particularly important for phonology, and are given a more formal, extensive treatment in Chapter 6.) In this chapter, we will often note the fact that sounds that are positional variants in English are contrastive in some other language.

The IPA contains only enough symbols for the sounds that are used contrastively in at least some language. No language uses all of the sounds in the IPA, but for any two symbols, there is some language that uses both to create a contrast between words. There are not enough symbols to represent all the details of positional variation. Additional details of positional or dialectal variation (such as partial voicing, slight raising or lowering, or more or less rounding) can often be indicated by *diacritics*: superscripts, subscripts, or modifications to the letters to indicate the desired change.
One final note of caution: IPA symbols are an extremely useful way to refer to sounds, but they are not themselves the “atoms” of linguistic representation. It should always be kept in mind that a phonetic alphabet is a system of labels for recurring combinations of articulatory gestures and the acoustic correlates of those gestures. The symbol [p] is a “cover term” for bilabial + plosive + voiceless, the symbol [b] is a cover term for bilabial + plosive + voiced, etc.

### 3.1 In Focus

The largest consonant systems (which include Caucasian languages such as Kabardian and Southern African languages such as Xhosa and Zhu) may contrast 80–100 or more consonants, but these include many secondary and complex articulations, combinations of the simpler symbols, along with contrasts in voice quality and airstream; see Sections 3.2.3 and 3.3. It can be difficult to get an accurate count because it is sometimes not very clear whether a particularly complex combination of constrictions should count as one consonant or a sequence of two. On the other end of the scale, consonant systems may be as small as six or eight in the languages of the Pacific Islands: for example, six in Rotokas (essentially [p, t, k, b, d, g], with nasals and fricatives occurring as positional variants) and 8 in Hawaiian [p, k, m, n, h, l, w]. For the record, the prize for the overall largest sound inventory probably goes to Xhosa, which in addition to its 80+ consonants has more than 20 different vowels, and the prize for the smallest goes to Rotokas, with just five vowels in addition to its six consonants.

### 3.2 Pulmonic Consonants

#### 3.2.1 Stops, Nasals, and Fricatives

As is evident from the IPA chart, fricatives can be made at every place of articulation, and stops and nasals at most. There are a lot of symbols to learn, and nothing for it but to work through them systematically. I strongly recommend trying each one out as you read. While learning symbols and definitions is probably not the most interesting part of linguistics for most students, I hope that trying out some challenging new sound combinations will be rewarding. Try the sounds out by yourself in front of a mirror, or with a friend: see if you can correctly transcribe the distinctions your friend is producing, and vice versa.

At the bilabial place, the oral and nasal stops (as in [pa], [ba], [ma]) are familiar and straightforward. The voiceless bilabial fricative [f] consists of air passing through a constriction between the lips: the same vocal tract configuration as blowing out a candle. This fricative occurs in Japanese words such as “futon” [futon] and “fugo” [fugo]. These words are borrowed into English with an [f], but try pronouncing them without bringing the lower lip back to touch the teeth. A bilabial fricative can also be made with the lips spread and pressed together rather than pursed. Figure 3.3 shows an image of the lip position used for a bilabial fricative in Setswana (a language of Southern Africa). I think most English speakers produce this type of fricative in the exclamation usually written “phew!” [ϕ], which is not identical to “few” [fju]. (Remember that [j] stands for the palatal glide, so that [ju] is pronounced “you.”) The voiced bilabial fricative [β] is made in the same way as [b], but with the lips held more loosely, allowing air to escape. In most dialects of Spanish, “b” is
pronounced [β] when it occurs between vowels: “la bola” (the ball) is pronounced [laβola].

The labiodental fricatives are [f] and [v]. There are no labiodental plosives: a language that had them would require all its speakers to have completely perfect teeth. For real people, air is always able to escape through gaps between the teeth. It is possible, however, to articulate a labiodental nasal [m]. When the velum is open, the constriction of the lower lip against the upper teeth is sufficient to divert most of the airflow through the nose. Although no language is known to use [m] and [ŋ] contrastively, labiodental nasals occur in English and other languages when a nasal precedes [f] or [v], in words like inverse, infamous, comfort, symphony, etc. Anticipating the labiodental fricative, speakers will also make the nasal at the labiodental place. You should be able to feel (and see, if you look in a mirror) that the lips never fully close in a word like “comfort.” Such variation, where two sounds that are adjacent to one another become more similar, is known as assimilation. Assimilation is the most common kind of positional variation, and a number of examples from English will be discussed in this chapter. Assimilation is also discussed at greater length in Chapters 5 and 11.

Languages generally choose either bilabial or labiodental fricatives as part of their inventory. English chooses [f] and [v], for example, while Japanese and Setswana choose [ɸ]. Few languages contrast these types of fricatives, probably because the distinction between them is hard to hear (though not hard to see). One language that includes all four labial fricatives in its inventory is Ewe, a language of West Africa.

The dental, alveolar, post-alveolar, retroflex, and palatal consonants are all coronals, that is, made with the tongue tip or blade. The IPA gives only a single symbol [t, d, n] for the stops at the dental, alveolar, and post-alveolar places of articulation. This is because languages will generally choose only one of these places to use contrastively, and the same (simple) symbol can be used in each case. English makes these consonants at the alveolar place of articulation, while Russian and French make them dental, for example. If necessary, a dental articulation can be indicated by placing a “tooth mark” under the more general symbol [t ɹ d], and a post-alveolar constriction can be indicated by a line under the letter [t ɻ d].

Not many languages are known to contrast dental and alveolar place of articulation. Examples include Malayalam (Dravidian) and some of the languages of New Guinea. In each of these cases, however, the distinction in location of the passive articulator is enhanced by a further distinction in the active articulator: alveolar consonants are made with the tongue tip, and dental consonants with the tongue blade. The phonetic term for tongue-tip consonants is *apical*; tongue blade consonants are *laminal*.

While English coronal stops and nasals are alveolar in the default case, dental and post-alveolar versions occur as positional variants. Repeat the following phrases, paying attention to the position of the tongue front in the final consonant of the first word of the phrase:

(1) Coronal assimilation in English

<table>
<thead>
<tr>
<th>original</th>
<th>assimilated</th>
<th>original</th>
<th>assimilated</th>
</tr>
</thead>
<tbody>
<tr>
<td>ten</td>
<td>ten times</td>
<td>ten things</td>
<td>ten roads</td>
</tr>
<tr>
<td>in</td>
<td>in time</td>
<td>in this</td>
<td>inroads</td>
</tr>
<tr>
<td>made</td>
<td>made time</td>
<td>made things</td>
<td>made rings</td>
</tr>
<tr>
<td>eight</td>
<td>eight times</td>
<td>eight things</td>
<td>eight rows</td>
</tr>
</tbody>
</table>

Figure 3.3 Lip position for the voiceless bilabial fricative in Setswana.
These examples illustrate another case of assimilation. In the first and second columns, the word-final \( [t, d, n] \) are alveolar. In the third column, however, the final consonant becomes dental \( [t, d, n] \) before the dental fricatives \( [\theta, \delta] \), and in the fourth column the final consonants become post-alveolar \( [t, d, n] \) before post-alveolar \( [r] \).

For the coronal fricatives, different symbols are used for the dental, alveolar, and post-alveolar places of articulation, because these sounds are often used contrastively, as in English. The difference between the fricatives is much more noticeable than the subtle sound difference between \( [t], [\theta], \) and \( [\delta] \). One reason is that \( [s, z, \mathcal{f}, \mathcal{s}] \) are grooved fricatives, and \( [\theta, \delta] \) are not. The groove in the tongue creates a strong channel of airflow that produces a loud, high-pitched noise when it hits the teeth. (Again, imagine the splash when you spray water from a hose against a wall.) Try articulating these sounds a few times in succession: \( [s] \; [\mathcal{f}] \; [s] \; [\mathcal{f}] \; [s] \; [\mathcal{f}] \). Can you feel that the sides of your tongue are raised for these fricatives while the center of the tongue is lowered, creating the channel? In contrast, the tongue is flat for \( [\theta, \delta] \), and the noise is softer and lower-pitched. Referencing this difference in pitch and volume, the alveolar and post-alveolar fricatives are termed sibilants, while the dentals are non-sibilants. Distinguishing among the sibilants, the post-alveolars have a somewhat lower pitch than the alveolars. Did you also notice a difference in lip rounding between \( [s] \) and \( [\mathcal{f}] \) ? Rounding the lips for \( [\mathcal{f}] \), by creating a bigger space in front of the teeth, lowers the pitch of the \( [\mathcal{f}] \) even further, magnifying the difference between the alveolars and post-alveolars.

The post-alveolar fricatives \( [\mathcal{f}, \mathcal{s}] \) are laminal, made with the blade of the tongue. In order for the tongue tip to make a constriction in the post-alveolar region, it must be curled back. Such a configuration (apical post-alveolar) is known as retroflex (literally, backward-turning). As was noted in Chapter 2, the only retroflex sound in English is the rhotic, and even that may not have much if any tongue-tip raising for many speakers. Other languages, however, notably Hindi and other languages of India, have a full set of retroflex stops, fricatives, and nasals. The IPA symbols for these sounds have a little hook under the symbol, recalling the curling back of the tongue. A telltale sign of an Indian accent is substituting retroflex stops for English alveolars. Figure 3.4 shows an MRI image of a subject producing a retroflex stop. In other languages the retroflexion may not be so extreme.

Modeling your own articulation on the picture, try producing \( [\mathcal{f}a\mathcal{f}], \; [\mathcal{f}u\mathcal{f}], \; [\mathcal{f}a\mathcal{g}], \; [\mathcal{f}a\mathcal{s}], \; [\mathcal{f}a\mathcal{z}] \). The large space under the tongue that is created by the tongue tip moving up and back (the sub-lingual cavity) gives this set of sounds their distinctive quality.

Palatal articulations are usually also considered coronals, although they involve raising of both the tongue front and tongue body up to the hard palate. The symbols are \( [c, j] \) for the stops, \( [c\mathcal{j}, j] \) for the fricatives, and \( [n] \) for the nasal. Be careful of the differences in the direction of the hook, and where it is attached, in distinguishing palatal \( [n] \), retroflex \( [n]\mathcal{f} \), and velar \( [\mathcal{n}] \). For palatal \( [n] \) the hook resembles a \( [j] \), for velar \( [\mathcal{n}] \) the hook resembles a “\( [g] \)”, and for retroflex \( [n]\mathcal{f} \) the hook recalls the turned-up tongue.

English does not use any of palatal stops, fricatives, or nasals contrastively, although the palatal nasal may be familiar from languages such as Spanish and French: Spanish “año” (year) and French “agneau” (lamb) are both \( [a\mathcal{p}a] \), though with a difference in stress. The tongue positions for \( [c\mathcal{j}] \) are very similar to that of \( [j] \) (as in “you”). If you start with a \( [j] \) and then make it voiceless, the resulting increased airflow from the open vocal folds is likely to be sufficient to cause frication: a word such as “hue” may be transcribed \( [\mathcal{f}u\mathcal{u}] \). To

Figure 3.4 MRI image of a retroflex stop. The horizontal lines are a feature of the imaging technology, designed to track changes in muscle shape.

create a voiced palatal fricative, again start with [j], but then raise the tongue slightly to create a narrower constriction: you will be able to hear the point at which frication begins. From [ç] and [ʝ], it is a simple matter to press the tongue up further, against the palate, to close off the airflow and create the palatal stops [c] and [ʝ]. Palatal stops, as opposed to affricates, are cross-linguistically marked, though they do occur in languages such as Akan (West Africa) and Hungarian (Eastern Europe). Because the closure involves a long portion of the tongue, it is rather difficult to release the stop closure quickly enough to avoid frication, and palatal stops often evolve over time into affricates. Many languages, like English, have alveolar and velar stops, and palatal or alveopalatal affricates. (More on the IPA symbols for affricates in Section 3.2.3.) No language contrasts palatal place and (laminal) post-alveolar place for stops or nasals, but a number of languages, including Polish and Cantonese, contrast palatal and post-alveolar fricatives.

There is actually one more possible coronal articulation. It is possible to make a stop by extending the tongue out to make a closure against the upper lip. **Linguo-labial** stops, fricatives, and nasals have been documented in languages spoken on the islands of Vanuatu, in the South Pacific. Being exceedingly rare, linguo-labial place is not accorded its own column in the IPA consonant chart, but can be indicated by placing a “seagull” diacritic (evoking the upper lip) under the corresponding alveolar symbol, as in Tangoa [ŋata] eye.

The next place of articulation in the IPA chart, velar, is used contrastively in just about every language. Velar consonants use the tongue dorsum as the active articulator, moving up and back to make a constriction at the velum. The velar stops [k, ɡ, ŋ] have already been discussed. Velar fricatives [x, ɣ] do not occur in English, but they are not hard to make. On the principle that a fricative is a stop with just slightly less constriction, you can make a [x] by beginning with a [k] and then lowering the tongue just enough to allow some turbulent airflow to pass through. This is the sound at the end of the German (or pretentious English) pronunciation of the name of the composer Bach. Add in voicing (and a slightly tighter constriction, to make up for the reduced airflow) to produce [ɣ].

The tongue body can also make constrictions further back, at the uvular place of articulation. To make the uvular stops [q] and [ʁ], begin with a [k] or [ɡ], then move the tongue a few centimeters back. Practice contrasting [aka], [aqa], [aga], [aqa] and then [ana], [ança]: same tongue movement each time, but with different settings for voicing and nasality. Uvular stops are common in many Native American languages. Also, the name of the Middle Eastern country Qatar is pronounced exactly the way it is spelled (as are most of the famous “q-without-u” words in the Scrabble dictionary). Once you have a feeling for the place of the stops, it is not hard to produce the fricatives [χ] and [ɣ] by loosening the constriction without changing the place. In many dialects of French, an initial “r” as in “rouge” is a voiced uvular fricative: [ʁɔʒ].

Constrictions can also be made deep in the throat, with the tongue root moving back toward the pharyngeal wall. Voiced and voiceless **pharyngeal** fricatives are found in Arabic and other Semitic languages. These sounds are among the most difficult consonants for English speakers to learn to make, since they involve an active articulator, the tongue root, that is not used for English consonants at all. To make the pharyngeal fricatives [ʕ] and [h], I find that it’s easiest to start with an [h], the laryngeal fricative, and then try to tighten the throat until a fricative sound is heard and felt. This is much easier with an open jaw and a low vowel: [ɑˈha], [ɑˈsʔa]. In fact, many of the languages that use pharyngeal fricatives do not allow high vowels to occur adjacent to them.

The tongue root itself cannot be moved far enough back to make a complete closure, but epiglottal stops and fricatives, with the epiglottis touching the pharyngeal wall, have been reported in the Caucasian language Agul. Symbols for these sounds are included in the “other symbols” section of the IPA. Only a voiceless epiglottal plosive is included, presumably because combining that much pharyngeal constriction with a laryngeal configuration
consistent with voicing is not possible. Note that it is also impossible to produce a pharyngeal or laryngeal nasal stop. Knowing what you do about the anatomy of the vocal tract (look at Figure 1.8 on p. 10 if you need to) can you figure out why? (Answer at the end of this section.)

Our tour of the places of articulation finishes with the laryngeals: voiceless [h], breathy voiced [ʔ], and glottal stop [ʔ]. Note that there are only three laryngeal consonants, because there are so few degrees of freedom for laryngeal movement. Opening the vocal folds produces [h], closing them produces [ʔ], and an intermediate position, with enough airflow to cause frication but also enough tension to cause some vocal fold vibration, produces breathy voiced [ʔ]. This last sound may be produced when English [h] occurs between vowels, as in “ahead”: [æʔeɪd]. We already noted in Chapter 2 that the glottal stop is a positional variant of [t] in many dialects of English, frequently occurring before [n] or in word-final position: [ɪmporʔnʔ]. It should now be noted that, in other languages, the glottal stop is used as a regular consonant. For example, in Hawai’ian, the sequence [poʔo] means head, and [ʔoʔo] is a type of bird. In the word “Hawai’i” the apostrophe stands for a glottal stop.

We have seen in this section that fricatives can be made at every place of articulation: bilabial, labiodental, dental, alveolar, post-alveolar, retroflex, palatal, velar, uvular, pharyngeal, and laryngeal, even lingualabial. Plosives can be made at all of these but two: labiodental stops do not work because our teeth have gaps in them, and pharyngeal stops (with the one reported exception of epiglottal stops in one language) do not work because our tongues do not go back that far. For a similar reason, there is not even a column for places of articulation like “retroflex velar,” if you can imagine what that would involve. (In the IPA chart, the block for labiodental stops is left blank, not shaded, because it is not physically impossible to make such a stop, just highly improbable; similarly the blocks for voiceless pharyngeal stops and for pharyngeal fricatives are left open to allow for the possibility of epiglottals.)

All the fricatives and plosives come in voiced and voiceless pairs, except for the glottal stop, which is necessarily voiceless. As was noted in Chapter 1, however, voiced and voiceless fricatives and plosives are not equal: voiced obstruents are more marked than voiceless obstruents. Voiced obstruents are difficult to produce because the vocal tract obstruction inherent in a stop or fricative necessarily interferes with the airflow necessary for vocal fold vibration. A language may have both voiced and voiceless members of a pair (like English), or just the voiceless members (like Finnish or Hawai’ian), but no language has voiced stops but not voiceless. Thus the concept of linguistic markedness may be stated in the form of an implication (an “if–then” statement): If a language has voiced obstruents, then it will also have voiceless obstruents. In more general form, “If a language has the marked option, then it will also have the unmarked option.” Even in English, which contrasts words like [pap] and [bɒb] or [kɪk] and [ɡɪɡ], voicing tends to be slow in starting, or quick to die out, in word-initial and word-final position. A consonant that is thus at least partially “devoiced” may be indicated with an open circle under the symbol for the usually voiced sound: [b互通] and [ɡ互通], for example.

Returning to the IPA chart, we see that nasal stops occur at all places of articulation except pharyngeal and laryngeal. The reason for this is that the pharynx and larynx are lower in the throat than the velum. If airflow is cut off at the larynx, there can be no airflow through the nasal passages, and thus no nasality.

Every language has at least one or two nasal stops, and nasalized vowels and approximants (written with a tilde, [o] or [w]) are also common. But nasalized fricatives, while possible, do not work so well: venting a volume of airflow through the nose leaves a much weaker airstream with which to create a fricative. Try pronouncing [z] and see how fast you run out of air, if you can manage it at all.

The next sections turn to sounds whose articulatory requirements limit the places at which they can be realized.
3.2.2 laterals, trills, taps, and other approximants

A lateral sound, by definition, stops the airflow down the center line of the vocal tract, but allows air to escape along one or both sides of the tongue. The most common lateral is the alveolar lateral approximant, exemplified by English [l]. To make an [l], the tongue tip moves forward to contact the alveolar ridge, while the tongue body moves back and down, stretching out the tongue. Just as stretching a piece of gum makes it thinner, stretching the tongue also makes it thinner, so much so that the sides of the tongue pull away from the sides of the palate or the upper teeth, allowing the air to escape. This backing and lowering movement of the tongue body can be heard when an /l/ follows a high vowel: “feel” is pronounced [flɛl] and “pool” is pronounced [pʊoʊl]. A lateral for which this tongue backing gesture is timed so as to be particularly evident may be termed a “dark l,” or “velarized l” and symbolized [ɻ]. In English, [l]s at the end of a syllable are dark, while those at the beginning of a syllable are light: listen to difference between “leaf” and “feel.”

Voiced and voiceless lateral fricatives [ɭ, ɭ] are made with the tongue in approximately the same configuration as for [l], but with a tighter constriction at the sides and/or more airflow, so that the airflow causes frication. These fricatives are common in Navajo and related languages. Note that the symbols for the “dark l” and the lateral fricative are very similar, differing only in the diacritic that runs through the center of the symbol: the “dark l” is written with a tilde [ɭ], the lateral fricative with a loop [ɭ]. Be careful not to confuse them.

Many languages also have post-alveolar laterals, symbolized [ɭ]. This is the sound written “gli” in Italian, as in “famiglia” [famiˈʎa] family and “meglio” [meʎjo] better. Languages that have retroflex stops and nasals may also have a retroflex lateral [ɭ]. The tongue mechanics are the same for these sounds as for [l], except that the forward contact is made in the post-alveolar region (with the tongue blade for [ɭ] and the tongue tip for [l]), rather than at the alveolar ridge.

It is also possible to make a velar lateral, symbolized [ɭ]. For this sound, the tongue dorsum is raised toward the velum, but with greater constriction on one side than on other. Dorsal laterals contrast with alveolar laterals in Melpa (Papua New Guinea) and in Zulu (Southern Africa). They may also occur as positional variants of [l] in some dialects of English, in words that end in “lk.” Try pronouncing a word like “milk.” Does your tongue tip go up to the alveolar ridge for the [l]? Even if it does, try pronouncing the word with your tongue tip held down behind your lower teeth. If you can succeed in producing a lateral sound without involving the tongue front, it will be a velar lateral.

Labial laterals, on the other hand, are impossible. Our facial muscles just do not permit an approximant articulation in which the lips are closed in the middle but open at the sides. (Try it.) Even if you could do it, such an articulation would not sound any different from [w], so it would not be worth the trouble. There are no pharyngeal or laryngeal laterals either, for obvious reasons. The IPA leaves open the possibility that a uvular lateral might exist, though none has been attested.

The next sets of sounds to be considered are the taps and flaps. These articulations are termed ballistic: they consist of “throwing” the active articulator against the passive articulator, so that it either strikes the passive articulator in passing (a flap) or bounces off it (a tap).

American English speakers produce a tap instead of a plosive for “t” and “d” in words like “city,” “pretty,” “shady,” “Eddie,” “felicity,” “automatic,” “ladder” (essentially, these consonants become taps when they occur between any two vowels, the second of which is unstressed). The IPA symbol for this tap is [ɭ]. You may also see it written as [D]. Almost all of the [d]s and [t]s are tapped in the phrase: “Dead-headed Ed edited it” (attributed to Peter Ladefoged). Try this phrase out and try to feel what your tongue is doing. In a tap, the tongue tip moves up to tap briefly against the alveolar ridge, but the two articulators do not stay in contact long enough for any pressure to build up. (The closure for typical tap is about 20 milliseconds,
or $\frac{1}{50}$th of a second, long.) Because there is no pressure build-up, a tap counts as a sonorant. Taps are so brief that the vocal folds do not have time to pull apart, so taps are also voiced. In a more careful and slow pronunciation, [t] is voiceless and [d] is voiced, but if they become taps, the distinction is lost: both are voiced. Thus you often see children spell words like “pretty” as “priddy.” We once had a long discussion at my house over whether the toys Santa brings in the song “Santa Claus is Coming to Town” are cars designed for cats or for children (that is, kitty cars or kiddie cars). If all you hear is the tap, and you haven’t been taught how to spell the word, the choice between “t” and “d” can be a toss up. (The medial sound in words like “Santa” and “winter” can be a nasal tap.)

3.2 Note that tapped alveolar stops mark a particularly American or Australian accent. Other English dialects will tend to use glottal stops or full alveolar stops where Americans and Aussies have taps. On the other hand, some dialects of British English use a tapped variant of [r]. The Scottish pronunciation of “pearl” is very like the American pronunciation of “pedal.”

The words tap and flap are often used interchangeably: one often hears about “the American flapping rule,” for example. However, to a phonetician, there is a difference. Technically, a tap involves an up and down motion: the tongue tip bounces off the alveolar ridge. A flap is a one-way motion, like a swinging door. In this text, we’ll maintain the distinction, but keep in mind that it is common to call the sounds of English flaps.

Retroflex flaps [t] occur in Hindi: the tongue tip curls back as for a retroflex stop, and then drops down, hitting the alveolar ridge with the underside of the tongue on the way down. The only other kind of flap that has been identified is a labiodental flap. This symbol [v] is the newest addition to the IPA (as of this writing), having been added in 2005. A labiodental flap is made by drawing the lower lip behind the upper teeth, and then pulling it forward. Labiodental flaps occur in Banda and neighboring languages of Central Africa. (See: http://www.sil.org/~olsonk/research.html).

While taps and flaps are ballistic, trills are aerodynamic. The active articulator is held in the correct position with the correct tension, so that the flow of air between active and passive articulators sets one or both into vibration. The best known trills are the tongue-tip “trilled rs” of Spanish and French, symbolized [r]. In a uvular trill [r], it is the passive articulator, the uvula, that vibrates. You already know how to make a bilabial trill [b]: as a motorboat sound or raspberry. (Just lips – no tongue involvement. Sticking the tongue out would make a linguo-labial trill – certainly doable, but not known to be used as speech sounds in any language.) Bilabial trills, however, function as regular consonants, parallel to [r] or [n] or [b], in languages such as Kele and Titan (Papua New Guinea).

The last set of consonants to be considered are those listed simply as “approximants” in the IPA chart. These sounds involve a constriction in the vocal tract, so that the active and passive articulators are brought near each other, but not so much that any frication is caused. Non-lateral approximants may be labiodental [v], dental/alveolar/post-alveolar [j], retroflex [r] palatal [j], velar [w], or labiovelar [w]. The labiodental approximant may be thought of as a severely weakened [v], and the velar as a severely weakened [g]. Bilabial approximants may also occur, though not contrastively: they may be symbolized [β], the diacritic standing for “a lowered version” of [β]. The glides [j] and [w] are well known to English speakers. This leaves [j] for last. The sound that begins the English word “road” is an alveolar or post-alveolar approximant: a sound that is cross-linguistically much rarer than either a trill or a tap. Note also that, like [j], [i] is made with rounded lips.
3.2.3 contour and complex segments

You may have noticed that the IPA chart contains no symbols for affricates. It is not that affricates are not common or contrastive — they are both — but that an affricate can always be represented by a sequence of two symbols, a stop followed by a fricative at the same place of articulation. English has only post-alveolar affricates ([tʃ] and [dʒ]) but affricates can be made at any place a stop and fricative can be made: bilabial as in German [ʃtʃ] pepper, alveolar as in Japanese [tsuʃi] tsunami, palatal as in Thai [kʰɔ̂pʰaː] blank, velar as in Setswana [xɔ̂pʰaː] smear with dung, uvular as in Oowekyala [qɔ̂] powder or [dʑiɛ̂] “to stop making vocal noise, e.g., stop crying.” The last has both an alveolar and uvular affricate.

The real question to ask about affricates is why they “count” as just one segment, not two. You cannot necessarily tell by listening: an affricate might not be any shorter, or articulated any differently, than a sequence of two segments. The argument for their unity is not based on their physical, phonetic, characteristics, but on their phonological patterning. Edward Sapir, one of the most influential linguists of the first half of the twentieth century, answered by saying that affricates “feel” like one segment to the speakers of the language. English speakers have a sense of a “ch-sound” that is parallel to their sense of a “b-sound” or “s-sound.” There’s certainly something to be said for taking native speaker intuitions into account, but we might go further in finding that affricates behave like single segments: they appear in contexts where only single segments, not sequences, should be allowed. English generally does not allow sequences of stop + fricative at the beginning of a word: we simplify “psychology,” whose root was pronounced with [ps] in Greek, to just [s], for example. But the affricates [tʃ] and [dʒ] are fine. Setswana goes even further, never allowing two obstruents in a row at all, but [tʃ] and [dʒ] are fine in Setswana too. Additionally, the two parts of an affricate always share the same place and laryngeal configuration, contributing to the sense of sameness. If [dʒ] or [ʃtʃ] occurred, for example, they would not count as affricates. Segments like affricates are called contour segments: they start off as stops, but finish as fricatives.
(Arguably, there are no reverse affricates: single segments that start off as fricatives and end as stops.)

Segments can have contours in other dimensions as well. Pre-nasalized stops [n̥d, n̥b] start off nasal and end oral: the velum closes part-way through. There can also be stops with lateral release: [t̥, d̥']. The segment begins as a standard stop, but the closure is released from the side rather than from the center.

Contour segments like affricates and pre-nasalized stops have two different articulations in sequence. Complex segments have two different articulations at the same time. We have already discussed one complex segment: [w], the labio-velar glide, which combines an approximant constriction at both the velum and the lips. Labio-velar stops, which combine a [k] and a [p], or a [g] and a [b], at the same time, are common in many West African languages – in the name of the language Igbo, for example, or the Igbo word [kpa], to weave by hand. While double articulations at various places of articulation are not hard to make (given the independence of the active articulators), they can be hard for the ear to distinguish, so double articulations other than labio-velars are vanishingly rare. A doubly articulated fricative combining [ʃ] and [χ] is reported for Swedish, and has been recognized with the IPA symbol [ʃ]; another fricative combination, [ʃ] and [ϕ], has been reported for dialects of Setswana.

If a sound has two different articulations, but one has a greater degree of constriction than the other, the more constricted articulation is the primary articulation, and the more open is the secondary articulation. Secondarily articulations are indicated with superscripts. They include rounding (as in [kʷa], [sʷa]), palatalization ([kʻa], [sʻa]), velarization ([lʻa], and pharyngealization ([sʻa]). In each case the main constriction is accompanied by an approximant-like constriction at the lips, palate, velum, and pharynx, respectively.

3.3 non-pulmonic consonants

There are three kinds of non-pulmonic consonants: implosives, ejectives, and clicks. Each uses a source of airflow other than the lungs. Each creates an enclosed body of air within the oral tract, and then manipulates the pressure by enlarging or compressing the enclosed space. When the pressure is released, a distinctive sound is made. Implosives and ejectives create a space between the larynx and another place of articulation; clicks create a space between the body of the tongue and another constriction further forward (lips or tongue front). We consider each in turn.

3.3.1 implosives

Implosives may also be called glottalic ingressive stops. For these consonants, the main source of change in air pressure is movement of the larynx, and (to the extent that air is moving) the primary movement of air is into, rather than out of, the mouth. Implosives probably arose in languages as a way to keep voiced stops voiced. Think about what happens during a pulmonic voiced stop, such as [b]. A closure is made with the lips; by definition, no air is allowed to escape from the mouth. Meanwhile, however, air is continuing to flow out of the lungs and into the oral cavity, passing over the vocal folds and causing vibration. It does not take long, however, for the mouth to fill up with air. Within a few hundredths of a second, air pressure above the larynx is equal to the air pressure below the larynx, and at that point airflow, and voicing, will cease. Try pronouncing a syllable like [ab], and try to keep voicing going into the [b] for as long as possible, without opening the lips. Voicing will stop very quickly.
How can a speaker keep voicing going, without opening the lips? One way, of course, would be to vent the airflow through the velar port, turning the voiced oral stop into a nasal stop. If that is undesirable (since [m] and [b] are probably contrastive) another way is to somehow make the oral cavity bigger, to allow more air to fit in. The walls of the oral cavity, especially the cheeks, are somewhat pliable, and they do expand a bit in voiced stops, but the speaker can also take action by lowering the larynx. If the larynx moves down in the throat, the space in the oral cavity expands by just that much. Recall Boyle's law: if the volume of an enclosed space increases, air pressure within decreases. The decrease in air pressure, plus the action of physically pushing the vocal folds down over the air in the trachea, can keep the vibration going a bit longer.

You will probably get larynx lowering in your [ab] experiment. If you place your finger on your larynx while trying to draw out the [b], you should feel the larynx moving down in an effort to expand the oral cavity. You may also feel, for a moment, the lips being slightly drawn in by the lowered air pressure. If you open your lips at exactly the moment you feel the lips being drawn in, you will have created an implosive. Air does not exactly rush into the mouth, but the interaction of the differing pressures inside and outside the oral cavity creates a distinctive sound.

It seems to be easiest to learn to make implosives at the bilabial place of articulation, since the longer distance from larynx to lips makes for maximum oral cavity size, but languages use implosives at all places of articulation. They are especially common in languages of West Africa. As shown in Figure 3.2, the IPA symbols use a hook ascending from the symbol for the voiced plosive: \([b, d, g]\). Because the aerodynamics of creating an implosive sound require both a vocal tract closure and vibrating larynx, implosives are always voiced stops. There are no voiceless implosives, and no implosive fricatives or approximants.

### 3.3.2 ejectives

Ejectives are the inverse of implosives. They may be called glottalic egressive: as with the implosives a change in air pressure is caused by movement of the larynx, but in this case direction of airflow is outward. Ejectives are written in IPA with an apostrophe following the symbol for the voiceless plosive at the same place of articulation: \([p’, t’, k’, q’]\).

Ejectives begin with two closures, one at the larynx (a glottal stop) the other at a different place of articulation, such as velar. The two closures enclose a body of air. In the case of the ejectives, the larynx is raised rather than lowered. Raising the larynx makes the space between the closures smaller, compressing the air and raising the pressure. Then, the forward closure is released, causing a distinctive pop. A few milliseconds later, the glottal closure is opened, and the free flow of air (and voicing) for the vowel is resumed.

In order to make an ejective, you have to first be comfortable with making a glottal stop. Practice “uh-oh” until you are familiar with the muscle configuration required. Try a glottal stop between various vowels. From there, you should be able to make a velar and glottal closure at the same time: a glottalized \([k’]\). With ejectives, it’s easier to make them if the oral closure is further back, since it’s easier to increase the pressure if the volume between larynx and closure is small. Once you can make a \([k’]\), you need to add in larynx raising. This is the hardest part. It may help to get a feel for what’s involved if you try (silently) singing a very low note and then a very high note. For the high note, the larynx rises in the throat. Try to feel the muscular sensations of larynx raising, and get it under conscious control. Finally, sequence all the parts:

1. Simultaneous closure at larynx and in the oral cavity (for example, \([k’]\)).
2. Larynx raising.
Again, it can help to place a finger on your throat so that you can feel if your larynx is moving. You’ll know you are producing an ejective if you can feel the larynx rise, and if you hear a short period of silence between the release of the stop and the beginning of the vowel, when the mouth is open but the larynx is still closed. That period of silence is an important cue to distinguishing this kind of stop.

It is also possible to make ejective fricatives, affricates, and laterally-released stops. Although the fricative or lateral portion will be short, larynx raising can generate sufficient airflow for these articulations. They are found in Native American languages, and in Amharic (Ethiopia). Because the larynx must be closed, however, all ejective sounds are voiceless.

### 3.3.3 clicks

The third kind of non-pulmonic consonants are the clicks, or velaric ingressive consonants. You probably already know how to make clicks: the “kiss-kiss” and “tsk-tsk” sounds are clicks. The trick is to incorporate them into the speech stream as consonants.

Clicks begin with a closure of the tongue body against the velum or uvula. Then a second closure is made further forward, for example, the tongue front against the teeth. A small body of air is trapped between the two closures. The tongue body is then slid backward, without breaking the seal against the upper surface of the vocal tract, and the middle of the tongue is lowered. These actions make the pocket of air larger, and thus the air pressure in the pocket becomes lower. When the forward closure is released, a clicking sound is produced.

Clicks can be made at the labial, dental, post-alveolar, and retroflex places. The labial click [ʘ] is the “kiss-kiss” sound, though it is generally made with the lips pressed flat together rather than pursed. (The bilabial is the rarest of the clicks.) The dental click (ǀ) is the “tsk-tsk” sound. Once you’ve practiced that a few times, try the alveolar (!) and palato-alveolar (ǂ) places, moving your tongue tip further back each time. You should be able to notice that as your tongue tip moves back, and the space under the tongue (the sub-lingual cavity) gets bigger, the pitch of the click sound gets lower. If you articulate the click slowly, you may also be able to feel the tongue body moving back and the resulting pressure change pulling back on the tongue front. The lateral click [ǁ] is made by sliding the side of the tongue down the side of the teeth. The lateral click has been described as the sound used to call horses. Once you’re confident in making the different clicks in isolation, try putting them between vowels, and then into longer words.

The velaric ingressive sound is made completely in the front part of the mouth: from the tongue body forward. Therefore, clicks can be “accompanied” by other articulations going on at the back of the mouth. Clicks can be nasal, with air flowing through the open velum behind the velar closure while the click sound is being made in front of the closure. They can be voiced, voiceless, aspirated, creaky voiced or breathy voiced. They can also be affricated on release or not. For example, the ǃXóõ language (spoken in Namibia and Botswana) is reported to have 83 different varieties of clicks (five places of articulation and 17 different possible accompaniments). The accompaniments can be indicated by linking the symbol for the appropriate velar stop to the click symbol: [ǃ] is a nasal dental click, [gǁ] is a voiced lateral click, etc.

While clicks occur as expressions, interjections, and mimetics throughout the world, they occur as regular speech sounds mostly in a cluster of languages in Central and Southern Africa: the Khoesan languages and their neighbors. Given their versatility and clear auditory distinctiveness, it is not really clear why click consonants do not occur more widely. They are “difficult” to make, certainly, requiring a high degree of articulatory coordination, but no more so than other sounds that are much more widely distributed, such as ejectives.
3.4 positional variation in English

As promised at the beginning of this chapter, we have found a number of “exotic” sounds hidden in American English pronunciations, not only the post-alveolar approximant, but species such as bilabial fricatives, glottal stops, voiceless sonorants, labiodental nasals, and even velar laterals. For those with a special interest in the pronunciation of English, for teaching, learning, or speech therapy, we pause to collate these examples into a list of some positional variants of English consonants.

In this section, because we are concerned with predictable details of pronunciation, we use narrow transcription rather than broad. In a broad transcription, transcribing the English word “pan” as [pæn] is perfectly correct, since the word does indeed consist of a voiceless bilabial plosive followed by a low front vowel followed by an alveolar nasal stop, and the symbols are sufficient to distinguish this word from all others in the language, including “ban” [bæn], “pad” [pæd], and “pin” [PIN]. But many details of English positional variation are left out of the broad transcription: for example, initial voiceless stops are aspirated, vowels are nasalized before a nasal consonant, vowels are lengthened before a voiced consonant, and (at least in my dialect) the low front vowel has a high front onglide before [n]. A narrow transcription includes all these details: pʰæn. Of course, since sounds that are positional variants in one language may be contrastive in another, the question of how much detail to include in a broad transcription will depend on the language: aspiration will be included in a broad transcription of Thai, but not of English.

Linguists often state positional distributions in the form of rules: A is realized as B in environment Y. (For example: voiceless stops are aspirated in word-initial position.) Part of knowing how to speak English is following these rules. It is important to remember, however, that these rules are descriptive, not prescriptive. They describe how people actually do speak, not how people ought to speak. They are more like the laws of physics (if you let go of a rock, it will fall to the ground; water freezes at 0° C) than they are like the rules of “good English” you may have been taught in school (don’t say “ain’t”; a preposition is a bad thing to end a sentence with). It is not right or wrong to produce an intervocalic tap; it is just a fact that American English speakers do it regularly and British English speakers do not.

Chapter 10 further investigates the question of what it means to “know” a rule: even though you follow these rules (if you are a native speaker), you don’t know them in the way you know your phone number, for example. Here, our goal is a description of some of the positional variants of American English consonants. You can think of them as rules for converting broad into narrow transcription.

3.4.1 laryngeal configuration

i. Voiceless stops are aspirated in word-initial position, or beginning a stressed syllable.
   ii. Voiceless stops are unaspirated after [s]:

<table>
<thead>
<tr>
<th>Word</th>
<th>Narrow Transcription</th>
</tr>
</thead>
<tbody>
<tr>
<td>pit</td>
<td>[pʰt]</td>
</tr>
<tr>
<td>tall</td>
<td>[tʰæt]</td>
</tr>
<tr>
<td>coop</td>
<td>[kʰæp]</td>
</tr>
<tr>
<td>apart</td>
<td>[apʰæt]</td>
</tr>
<tr>
<td>retain</td>
<td>[rætʰæn]</td>
</tr>
<tr>
<td>potato</td>
<td>[pʰætʰæro]</td>
</tr>
<tr>
<td>accord</td>
<td>[skʰæd]</td>
</tr>
<tr>
<td>spit</td>
<td>[spit]</td>
</tr>
</tbody>
</table>
iii. Sonorants are devoiced when they follow an initial voiceless stop:

```
play  [plei]
pray  [prei]
tray  [trei]
cray  [krei]
crate  [kreit]
```

iv. Voiced plosives are (at least partially) devoiced, unless surrounded by voiced segments:

```
bad times  [bad'ta:mz]
bad guys  [bad gaiz]
good news  [ɡud'nus]
gag  [ɡæɡ]
```

v. Sonorants are produced with creaky voice when they precede a glottal stop:

```
important  [ɪmpɔr'tænt]
can't  [kænt]
```

vi. Intervocalic [h] is produced with breathy voice at the beginning of a stressed syllable, and is deleted before an unstressed syllable:

```
ahead  [əhed]
rehabilitate  [rɪhəbɪlɪteɪt]
rehearse  [rɪˈheərəs]
can he go?  [kæn ɪ goʊ]
I like him  [aɪ laɪk ɪm]
```

vii. Stops are produced without audible release when they precede another consonant:

```
begged  [bɛd ˈdʒl]
stopped  [stɑpˈtʃl]
stop that  [stɑpˈðæt]
make noise  [meɪk ˈnoiz]
met someone  [met ˈsʌməm]
read books  [riːdˈbʊks]
```

### 3.4.2 change in place

viii. [n] assimilates in place of articulation to a following consonant. [m] becomes labiodental preceding [f]:

```
in place  [ɪmpləs]
in case  [ɪnkeɪs]
in time  [ɪntˈaɪm]
in fact  [ɪnfækkt]
in there  [ɪnˈðeər]
comfort  [kɒmфɜːt]
symphony  [sɪnfəni]
```
ix. Alveolar [t, d] become dental before [θ, ð], and become post-alveolar before [ʃ]:

- width [wɪðθ]
- hide this [hæɪðəs]
- hit them [hɪθəm]
- get rich [gɛtɪŋf]
- not really [nætəli]
- bad roads [bædɹəʊdz]

x. [t] becomes a glottal stop preceding a nasal:

- button [bʌʔn]
- hit me [hɪʔmi]
- litmus [lɪʔmʌs]

xi. Alveolars become postalveolars preceding [j]. (This rule is optional, but usually applies before “you” and “your,” and in common phrases.)

- miss you [mɪʃju]
- this year [ðɪʃjɪə]
- bless you [blesjju]
- made you look [meɪdʃjʊlʊk]
- meet you [miʃju]
- those years [ðəʊʃjɪəz]

3.4.3 change in manner

xii. [t] and [d] are tapped in between two vowels, the second of which is unstressed:

- city [sɪri]
- atom [ærəm]
- photograph [foʊˈɡræf]
- electricity [ɪˈlektrɪsɪri]
- edited [ɛdɪtəd]
- lady [leɪri]
- daddy [dæri]
- pedestal [pəˈrɛstəd]

3.4.4 other changes

xiii. [l] is velarized (“dark”) at the end of a word or syllable:

- feel [fɪl]
- pool [pʰʊl]
- call [kʰæl]
- cold [kʰɔld]
- melted [meɪltəd]
- billboard [ˈbɪltboʊrd]
- leaf [lif]
- loop [lʊp]
- lock [lɔk]
xiv. The American English rhotic is a post-alveolar approximant \[\mathcal{R}\]. (Post-vocalic \[\mathcal{R}\] is not pronounced in some dialects: see Chapters 4 and 19 for more details).

\[
\begin{array}{ll}
\text{wreck} & [\mathcal{R}ek] \\
\text{ripe} & [\mathcal{R}eip] \\
\text{grow} & [\mathcal{R}au\mathcal{u}] \\
\text{pray} & [\mathcal{R}e\mathcal{e}] \\
\text{core} & [k\mathcal{R}e\mathcal{u}] \\
\text{harm} & [h\mathcal{R}om] \\
\end{array}
\]

xv. In a sequence of three consonants the second is deleted, except in very careful speech. (This rule depends to some extent on the consonants involved. It’s most likely to apply if the first consonant is [s], the second is [t] or [d], and the third is a stop. The rule does not apply if one of the consonants is \[\mathcal{R}\].)

\[
\begin{array}{ll}
\text{best man} & [b\mathcal{R}sm\mathcal{e}n] \\
\text{desktop} & [dest\mathcal{R}p] \\
\text{coastline} & [k\mathcal{R}ous\mathcal{l}m] \\
\text{hand-picked} & [h\mathcal{R}m\mathcal{p}\mathcal{k}\mathcal{t}] \\
\text{best friend} & [b\mathcal{R}f\mathcal{n}\mathcal{d}] \\
\text{pumpkin} & [p\mathcal{R}m\mathcal{k}\mathcal{m}] \\
\end{array}
\]

\text{but}

\[
\begin{array}{ll}
\text{hard line} & [h\mathcal{R}d\mathcal{m}] \\
\text{straight} & [\mathcal{R}t\mathcal{e}] \\
\end{array}
\]

---

\text{chapter summary}

- The symbols of the IPA chart represent all the (known) contrastive consonants in the languages of the world.
- Sounds that are contrastive in one language may be positional variants in another. The most common type of positional variation is assimilation: adjacent sounds become more similar.
- Fricatives may be produced at the bilabial, labiodental, linguolabial, dental, alveolar, post-alveolar, retroflex, palatal, velar, uvular, pharyngeal and laryngeal places of articulation. Plosives may be produced at all of these except labiodental. Nasals may be produced at all of these except pharyngeal and laryngeal.
- Lateral approximants and fricatives may be made with the tongue front or tongue body. Laterals are defined by a side-channel of airflow.
- Taps and flaps are ballistic articulations: the active articulator either bounces off the passive articulator (taps) or strikes it in passing (flaps). A trill is aerodynamic, with the movement driven by airflow.
- Contour segments combine two articulations in sequence; complex segments combine two articulations at the same time.
- The non-pulmonic consonants are ejectives (change in air pressure caused by larynx raising), implosives (change in air pressure air caused by larynx lowering) and clicks (air trapped between two tongue closures).
- Section 3.4 lists some of the positional variants of American English consonants.
further reading

No linguist describing the articulation of the sounds of the world’s languages can do so without an acknowledgement of the influence of, and a debt of gratitude to, the phoneticians Peter Ladefoged and Ian Maddieson. The sounds described in this chapter and the next are covered in much more detail in the books *Sounds of the World’s Languages* by Ladefoged and Maddieson, and *Vowels and Consonants* by Ladefoged. Soundfiles of example words for the languages cited can be found on the web archive compiled by Ladefoged, Maddieson, and colleagues at UCLA: http://www.phonetics.ucla.edu/.


review exercises

1. List the active and passive articulator for each of the places of articulation listed in the Chapter Summary.
2. Define the following terms:
   - environment
   - variation
   - contrastive
   - diacritic
   - assimilation
   - apical/laminal
   - sibilant
   - tap/flap
   - ballistic articulation
   - aerodynamic articulation
   - contour segment
   - complex segment
   - primary articulation
   - secondary articulation
   - implosive
   - ejective
   - click
   - sub-lingual cavity
   - descriptive vs. prescriptive rule
3. How is it possible for a click to be nasal?
4. The following steps are involved in making a [k']. Number the steps in the correct order. (One step is not used: Cross it off.)
   ______ make a closure at the larynx.
   ______ close the lips.
   ______ release the labial closure.
   ______ lower the vibrating larynx.
5. The following steps are involved in making a [k']. Number the steps in the correct order. (One step is not used: Cross it off.)
make a closure at the velum and the larynx at the same time.
release the velar closure.
open the larynx.
raise the larynx.
lower the larynx.

6. The following steps are involved in making a dental click. Number the steps in the correct order. (One step is not used: Cross it off.)

make a closure of the tongue front against the teeth.
release the dental closure.
make a closure of the tongue body against the velum.
release the velar closure.
lower the larynx.
lower the middle of the tongue.

7. Add three more example words or phrases to each of the sets in Section 3.4.

8. Transcribe these English sentences, first in broad transcription, then in more narrow transcription, indicating the positional variation in consonant pronunciation. (Positional variation in vowels is covered in Chapter 4.)

a. Would you meet me at the station at five?
b. My husband should have bought potatoes.
c. Will he be kept back?
d. They all wanted to be in the pictures.
e. In fact, I hated those times.

**Further analysis and discussion**

9. Fill in the following table to describe each consonant symbol.

<table>
<thead>
<tr>
<th></th>
<th>Airstream</th>
<th>Place</th>
<th>Manner</th>
<th>Voiced?</th>
<th>Nasal?</th>
<th>Lateral</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>l</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>j</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>q'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ð</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>θ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
10. For each description, write the IPA symbol in the left column. One description is not articulatorily possible: draw an X through that square.

<table>
<thead>
<tr>
<th>Airstream</th>
<th>Place</th>
<th>Manner</th>
<th>Voiced?</th>
<th>Nasal?</th>
<th>Lateral?</th>
</tr>
</thead>
<tbody>
<tr>
<td>pulmonic egressive</td>
<td>velar</td>
<td>stop</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>pulmonic egressive</td>
<td>interdental</td>
<td>fricative</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>pulmonic egressive</td>
<td>velar</td>
<td>approximant</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>velaric ingressive</td>
<td>alveolar</td>
<td>stop</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>pulmonic egressive</td>
<td>post-alveolar</td>
<td>fricative</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>pulmonic egressive</td>
<td>uvular</td>
<td>trill</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>glottalic ingressive</td>
<td>bilabial</td>
<td>stop</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>pulmonic egressive</td>
<td>labio-velar</td>
<td>glide</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>glottalic ingressive</td>
<td>alveolar</td>
<td>stop</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>glottalic ingressive</td>
<td>alveolar</td>
<td>stop</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

11. Draw a mid-sagittal diagram (or series of diagrams) illustrating the crucial steps in producing each of the following sounds. Annotate the diagrams to describe the required positions and movements of the tongue, lips, velum, and/or larynx, and to indicate the order in which the steps must be carried out.

a. [q’]
b. [ŋ!]
c. [d’]

12. Explain why the following articulatory combinations are impossible:

a. labiodental plosive
b. pharyngeal nasal
c. retroflex uvular
d. bilabial lateral
e. velar tap
f. nasal ejective

Go online Visit the book’s companion website for additional resources related to this chapter at: http://www.wiley.com/go/zsiga.

references

Most of the example words in this chapter are from the author’s own fieldwork. Additional examples (from Agul, Kele, Melpa, Tangoa, Titan, and Zulu) are from the database compiled at http://www.phonetics.ucla.edu/. Other sources include:

Oowekeyala

ǃXoõ


The instructions for making non-pulmonic consonant are based on those given by Ladefoged in *A Course in Phonetics*. I don’t know how I could improve on them.
Open wide, and say “ah.”

Chapter outline

4.1 The landscape 56
4.2 Cardinal vowels 57
4.3 Building inventories: dimensions of vowel quality 59
   4.3.1 Height and backness 59
   4.3.2 Tense/lax 59
   4.3.3 Rounding 61
   4.3.4 Central vowels 62
   4.3.5 Contrasts among the low vowels 64
4.4 Nasality and voice quality 66
4.5 Length and diphthongs 67
4.6 Tone 68
4.7 Positional variants of the vowels of English 70

Chapter summary 71
Further reading 71
Review exercises 72
Further analysis and discussion 73
Further research 74
References 74
Vowels, by definition, are open articulations. Because the active and passive articulators never actually make contact, the dimensions of “place” and “manner” are less useful for vowels than for consonants. Instead, vowels are conceived of as occupying points in a multidimensional space, whose coordinates may be either articulatory or acoustic. In this chapter, we map out the dimensions of the vowel space. We begin in Section 4.1 with a broad view of the landscape, considering how vowels and consonants relate to each other, and how they are organized into larger constituents of syllables and stress. Section 4.2 provides our reference points: the “cardinal vowels.” Sections 4.3 through 4.6 then describe the different dimensions that languages may use to define their vowel systems: height, backness, tenseness, rounding, nasality, voice quality, length, and tone. Special attention is given to central and low vowels, which are important in the English system. Section 4.7 concludes with a discussion of positional variation in English vowels.

4.1 the landscape

Vowels are the open-mouth periods separating the constrictions of the consonants. All speech is organized around this repeated closing–opening, consonant–vowel sequence, and all languages make this basic distinction. While it would be theoretically possible, no language is composed of all consonants or all vowels. Certain words might be all one or the other, like English [ɑː] or Oowekyala [qɜː] powder, so one could imagine a whole language made up of such words, but no natural language works that way.

The repeated closing–opening sequence forms the basis of syllable structure. The most basic syllable type is CV: a consonantal constriction released into a vowel. Children’s first words tend to be of the form [papa] or [mama], and all languages have syllables of CV shape. (Some, like English, allow syllables that are quite a bit more complicated, but all languages allow at least CV, and some allow only CV.) The vowels, in a way, form the background against which the consonantal constrictions can be heard. During the closure for a voiceless stop, after all, there is nothing but silence: we only perceive that silence as a consonant because of its contrast with the vowels around it.

A syllable can be defined as a grouping of segments. A more open articulation, usually a vowel, forms the syllable nucleus. More constricted consonantal articulations associated to the nucleus constitute the preceding onset and following coda. In the syllable [læf], [æ] is the nucleus, [l] is the onset, and [f] is the coda. Sometimes the more open articulation can be another consonant, as in Czech [vɪk] wolf; or the “opening” can occur through the nasal passages, as in English “prism” [prɪzmn]. In these cases, the less constricted consonant forms the syllable nucleus. These are termed syllabic consonants, and their status as a syllable nucleus is indicated by a vertical line beneath the symbol.

Syllables, in turn, often alternate between more prominent and less prominent, or stronger and weaker. The prominent syllables are termed stressed: they tend to be longer, louder, higher-pitched, and/or more clearly articulated than less prominent unstressed syllables. Not all languages make stress distinctions: in Korean, for example, all syllables are equally prominent. Those languages that do use stress vary in which phonetic parameters are used to indicate prominence. Thai speakers, for example, rely on differences in length to signal stress differences, while Russian speakers rely more on loudness. English speakers use a combination of cues, but seem to rely most on pitch. Whatever the cue, the prominent stressed syllables alternate with less prominent unstressed ones. Syllables and stress are discussed in detail in Chapters 15 and 16; they are mentioned here because they are part of the landscape within which vowels and consonants are articulated, and because syllable structure and stress patterns can often have an influence on the way segments are pronounced.
All languages make distinctions among the vowels, as they do among the consonants. Different vowel qualities are created as the tongue and lips take different positions. The vibrating vocal folds create a complex mix of frequencies. Depending on the shape of the mouth cavity, different frequencies present in the laryngeal vibration are amplified, and these differences are heard as different vowel qualities. (See Chapters 6 through 8 for details.)

As was discussed in Chapter 2, vowels can be described, rather roughly, in terms of the highest point of the tongue body: the tongue moves up for the high vowels, down for the low vowels, forward for the front vowels and back for the back vowels. (Since English has one of the larger vowel systems in the languages of the world, the discussion of the vowels of English has covered much of the ground.) The descriptors high/low and front/back have turned out to be very useful descriptions, but a number of caveats must be kept in mind when mapping from those descriptors to actual vowel articulation.

First, the shape of the vocal tract is not symmetrical. There is more room for tongue movement along the palate than there is in the pharynx. Thus there is more room for distinctions among the high vowels than among the low vowels. The vowel “space” therefore, is usually described as a quadrilateral, wider at the top than at the bottom. Second, we do not have great proprioceptive feedback in the back of our mouths: we do not have an accurate “feel” for where our tongues are, especially since, by definition, vowels do not make contact at a specific point along the surface of the vocal tract. We are much better at hearing vowel distinctions than feeling them. If this were not the case, it would be easy for people who are deaf to learn to speak; in reality, it is very hard. Third, the mapping from articulation to perception is not linear: a different amount of upward motion is required to make a front vowel sound high than to make a back vowel sound high, for example.

Therefore, when we use terms like “high” or “back,” or map a vowel as a point in the vowel space, we are more accurately describing what a vowel sounds like than where our tongues really are, though as an approximation of relative tongue position, it’s not that far off.

4.1 In Focus

According to Ladefoged and Maddieson in Sounds of the World’s Languages, a very few languages (Margi in West Africa, Eastern Arrente in Australia, and Abkhaz in Central Asia) contrast only two vowel qualities, essentially high vs. low tongue position. A larger number (including Classical Arabic, North American Aleut, South American Garawa, and Australian Yidiny) have three (usually high front [i], low central [a], and high back [u]). At the other end of the spectrum, the larger vowel systems (including !Xóõ, which also has a champion-size consonant system, and West Germanic languages such as German, Norwegian, and English) contrast more than a dozen. When distinctions in length, tone and voice quality are thrown into the mix, as well as the contour vowels known as diphthongs, distinctions multiply. Hanoi Vietnamese, for example, is reported to contrast 11 different vowel qualities, six tones, and five different ways of combining vowels into diphthongs.

4.2 Cardinal vowels

The idea of “mapping” vowels is due to the British phonetician Daniel Jones (1881–1967). Recall that the X-rays of tongue position shown in Figure 1.1 were done in his lab. You may want to consult that figure again, and note that what counts as “high” for [i] and [u] is not
the same, nor “front” for [i] and [æ]. Realizing this problem, and knowing the impracticality of making X-rays of speakers of all the languages linguists would hope to describe, Jones developed a system, called the cardinal vowel system, which would allow linguists in the field to describe the vowels of any language in terms that other linguists would understand.

Just as the cardinal points of a compass (N, S, E, W) provide reference points for orientation in geography, the cardinal vowels provide reference points for orientation in the vowel space. Jones defined eighteen cardinal vowels, using a mix of articulatory and perceptual properties. The eight primary cardinal vowels are graphed in Figure 4.1. Cardinal vowel #1 (symbol [i]) was defined as having the tongue body as high and forward in the mouth as possible, short of creating a fricative, with the lips spread as wide as possible. Cardinal vowel #5 (symbol [u]), positions the tongue as low and as far back as possible with the jaw wide open (you have to nearly swallow your tongue to make this one); while cardinal vowel #8 (symbol [u]) moves the tongue as high and back in the mouth as possible, with the lips tightly pursed. These are extremes of articulation, and don’t correspond to actual vowel qualities in any particular language. They define the extreme corners of the vowel space – high front, low back, and high back – with respect to which other vowel qualities can be defined. The other cardinal vowels between 1 and 8 were described in terms of how they sounded relative to the three extremes. Cardinal vowel #4, for example, was described as sounding “three-quarters of the way between numbers 1 and 5.” Numbers 9 through 16 (not shown in the figure) were defined as having the exact same tongue positions as numbers 1 through 8, but with opposite lip position. So cardinal vowel #9 has the tongue body as high and forward in the mouth as possible, just like #1, but with lips pursed instead of spread. Vowels 17 and 18 add in a central (neither front nor back) tongue position.

Jones trained his students to perceive and produce the cardinal vowels exactly as he had defined them, and they trained their students, and they theirs, until it was possible to write of an unknown language in the *Journal of the International Phonetic Association* that “the height of this vowel is about halfway between cardinal vowels #7 and #8,” and readers would know exactly the quality of sound referred to. Vowel charts, which graphed the positions of the vowels of a language relative to the cardinal vowels, were often used.

With the advent of audio recorders and digital means of measuring sound quality, the cardinal vowel system has fallen out of use. Vowel charts made today are usually based on measurements made by a computer from a recorded sound. The cardinal vowels are still a useful system of reference when a computer isn’t handy, however. You can still access a recording of Jones’ pronunciations on the web (see references on the website at http://www.wiley.com/go/zsiga).

The IPA vowel chart, reproduced in Figure 4.2, is similar to Jones’ cardinal vowel chart, but with some extra symbols added. (Jones was a leading member of the International Phonetic Association.) You can see the eight primary cardinal vowels around the edge of the chart. The secondary cardinal vowels are paired with them: as noted in the chart, pairs of vowels are identical in tongue position, and the vowel with
lip rounding is on the right. Jones’ original cardinal vowels line up with the horizontal lines. Additional vowels have been placed in the spaces between. As with the consonant chart, this represents a lot of symbols to process all at once, and some are more common and useful than others.

For our discussion of the vowels, we’ll begin with the most common vowel system: [i, e, a, o, u], and the basic dimensions of height and backness. From there, we’ll see how more complex systems may increase their inventories.

### 4.3 building inventories: dimensions of vowel quality

#### 4.3.1 height and backness

The most common number of vowels for a language to have is five. Languages as diverse as Modern Standard Arabic, Hawai’ian, Japanese, Spanish, Swahili, Tagalog, and Zulu all have five distinct vowel qualities, approximating the qualities of [i, e, a, o, u]. (Latin also had this basic inventory, giving us the five vowel letters of the Roman alphabet.) This system exploits three degrees of height (high vs. mid vs. low) and two degrees of backness (back vs. front). (Note that the IPA chart uses the term “close” for “high” and “open” for “low”.) The smaller vowel systems use some subset of these. A typical five-vowel system is graphed in Figure 4.3. Example words from Spanish and Hawai’ian are shown in Table 4.1.

Note in Figure 4.3 that the vowel symbols are not graphed in the exact same location as on the cardinal vowel chart. The fact that many diverse languages have vowel systems that can be transcribed [i, e, a, o, u] does not mean that the vowels have the exact same quality as Jones’ cardinal vowels. It would be unusual for a language to have an [i] or [u] as extreme as cardinal vowels #1 and #8. As is shown in Figure 4.2, the vowel transcribed as [e] is generally somewhere between cardinal vowels #2 and #3, [o] is between #6 and #7, and [a] is between #4 and #5. That is, the vowels are generally distributed fairly evenly in the vowel space, making maximal use of the space available. Nor does the transcription mean that vowel quality is exactly the same in each language. When the vowel system is small, the “area” occupied by each vowel can be rather large: there is room for quite a bit of variation before the different vowels “trespass” on one another’s space and cause confusion.

Thus, the symbol [e] does not always refer to the precise same vowel quality: There are not enough symbols to cover every detail of difference in every language, let alone speaker-to-speaker variation. It is useful here to remember, however, a point that was made in Chapter 3: the IPA symbols are really just “cover terms” for combinations of articulatory gestures. So [i] refers to “high, front (unround) vowel” and [u] to “high, back (round) vowel” even if the exact degrees of height, backness, and rounding are somewhat different.

#### 4.3.2 tense/lax

Enlarging the vowel system beyond five, languages will often exploit the tense/lax distinction, contrasting lax [i, e, a, u] with tense [i, e, o, u] as shown in Figure 4.4. Examples of
the tense/lax contrast in English and in the West African language Akan are given in Table 4.2.

Generally, when the terms “tense” and “lax” are used of vowels, the tense vowels are somewhat higher and longer than their lax counterparts, and they are said to have more “muscular tension,” though that term is seldom if ever quantified. In addition, the tense vowels of English are diphthongized, with upward movement of the tongue body over the course of the vowel: [ij], [eɪ], [oʊ], [u] (see Section 4.5 for further discussion). MRI and X-ray studies have shown that in at least some languages that make a distinction between vowels such as [i] and [ɪ], the tongue root is pulled forward for the tense member of the pair, widening the pharynx. Figure 4.5 shows a vocal tract tracing, based on MRI images, illustrating the difference in tongue root position between [i] and [ɪ] in Akan.

Table 4.2 Tense/lax vowel contrasts in Akan and English. Akan examples from Stewart (1967) and Tiede (1996). *See Section 4.3.5 for discussion of the low vowels of English.

<table>
<thead>
<tr>
<th></th>
<th>Front</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tense/Advanced</td>
<td>Lax/Retracted</td>
</tr>
<tr>
<td>Akan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>[di] to eat</td>
<td>[di]</td>
</tr>
<tr>
<td>Mid</td>
<td>[efie] home</td>
<td>[efie]</td>
</tr>
<tr>
<td>Low</td>
<td>[kasa] language</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General American English</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>[kijd] keyed</td>
<td>[kid]</td>
</tr>
<tr>
<td>Mid</td>
<td>[bet] bait</td>
<td>[bet]</td>
</tr>
<tr>
<td>Low</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

4.2 Native speakers of languages that do not have the tense/lax distinction (such as Spanish and Japanese) often have trouble learning to make the distinction in a language that does (such as English). I remember being shocked to hear a Spanish-speaking actor, on being asked in an American radio interview what his next project would be, reply that he was planning “to just leave this life.” It took me a moment to realize he was planning to “live” and not to “leave”: he just didn’t distinguish [i] and [ɪ] the way a native English speaker would.
Based on such findings, some linguists have argued that the tense/lax dimension would be more accurately described as advanced/retracted tongue root. However, in different languages the contributions of tongue root movement, tongue body movement, length, and perhaps larynx lowering, may vary. For example, the study illustrated in Figure 4.5 found that an English-speaking subject had less of a difference in pharynx width and more of a difference in tongue dorsum height than did the Akan-speaking subject.

However, crucial involvement of the tongue root may explain why all vowel heights are not equally amenable to a tense/lax (or advanced/retracted) distinction. Pulling the tongue root forward will tend to push the tongue body up, so that tongue root advancement and tongue body raising tend to go together. It is rare for a language to make a tense/lax distinction among the low vowels, presumably because tongue root advancement is not compatible with tongue body lowering. Some languages, such as Italian and Yoruba (which is related to Akan), make a tense-lax distinction only among the mid vowels. High vowels will tend to be tense, and low vowels will tend to be lax.

### 4.3.3 rounding

In the systems examined so far, rounding is predictable from backness and height: all and only the non-low back vowels are round. The incompatibility between rounding and low vowels is probably articulatory: it is harder to round the lips when the jaw is open (try it). The reason for the compatibility of rounding and backing, however, is perceptual. There is nothing difficult in articulating the front round counterpart of [i], which is symbolized [y], nor the back unround counterpart of [u], which is [ɯ]. Lips and tongue move independently: simply start with the more common vowel, and then switch lip position without moving your tongue. However, you will quickly hear that [i] and [u] sound much more distinct than do [y] and [ɯ].

The reason for this is that protruding the lips and backing the tongue have a similar effect on vowel acoustics. Rounding and backing both work to enlarge the space at the front of the mouth, while fronting the tongue and spreading the lips both work to make the front cavity smaller. Because rounding and backing reinforce each other, back round vowels sound maximally different from front unround vowels. Again, the concept of markedness may be invoked here. Back round and front unround vowels are unmarked: they are more common and the distinctions among them are easier to perceive. Back unround and front round vowels are marked: they are less common and harder to perceive. In general, languages will not choose a marked vowel unless the unmarked counterpart is also part of the inventory: The presence of [y] implies the presence of [i]. Thus, the most common five-vowel inventory consists of [i, e, a, o, u] while a system of just [y, o, u] is unattested.

Larger vowel inventories, however, may add the more marked front round vowels or back unround vowels. French and German have front round vowels in addition to front unround, Vietnamese has back unround vowels in addition to back round. Turkish has all four possibilities. The full set of symbols for front round and back unround vowels can be seen in
the IPA chart in Figure 4.2. Examples of vowel contrasts in French, Vietnamese, and Turkish are shown in Tables 4.3, 4.4, and 4.5.

Note that French and Vietnamese have a tense/lax contrast only for the mid-vowels. (One language that does contrast [i] and [u] is German: [bitən] to ask vs. [bitən] tubs.) While Turkish does contrast the tricky [kys]/[kus], it contrasts only two vowel heights. Turkish also has a system of vowel harmony, such that the vowels of a word will tend to agree in both backness and rounding (see Chapter 11). Such harmony not only makes articulation easier, but reinforces perception, since a separate decision on rounding and backness does not have to be made for every vowel in the word: the listener knows that if one vowel in the word is front, all the others will be as well. (Akan and Yoruba have tongue root harmony: the vowels in a word agree in whether the tongue root is advanced or retracted.)

### 4.3.4 central vowels

Languages tend to make their vowels as distinct as possible. If all the vowels of a language were similar, confusion could easily result. Thus, vowels tend to be peripheral: that is, around the edges of the vowel space, either back or front but not central. Some languages do make use of this central area, however.

The meaning of the central vowel symbols in Figure 4.2 should at this point be clear: tongue height comparable to that of the other vowels on the same horizontal line, but with a tongue body position intermediate between back and front. Lips are unrounded for symbols on the left, rounded for those on the right.

A few of the small-inventory languages are argued to be vertical, that is, to make no use of the front/back distinction and have all central vowels. Ladefoged and Maddieson report, for example, that the two vowels of Margi are [i] and [a] and the three vowels of Kabardian are [i, o, a]. However, languages seldom, if ever, contrast back and central vowels without also making a contrast in rounding. That is, no language is definitely known to distinguish different words on the basis of [i] vs. [ɯ] or [ʌ] vs. [ɤ].

There are a few possible cases for which a three-way front/central/back contrast might be needed. Ladefoged and Maddieson suggest that Nimboran, a language of Papua New Guinea, can be described as having three high unround vowels: [kip] fire vs. [kip] lime vs. [pakup] lid. The transcriptions are

---

### Table 4.3 Vowel contrasts in French. (http://www.phonetics.ucla.edu/vowels/chapter14/french.html).

<table>
<thead>
<tr>
<th></th>
<th>Front</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>French</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>Unround</td>
<td>Round</td>
</tr>
<tr>
<td>Mid Tense</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[le] the, pl.</td>
<td>[la] the, m. sg.</td>
<td>[lo] prize</td>
</tr>
<tr>
<td>Mid Lax</td>
<td>Unround</td>
<td>Round</td>
</tr>
<tr>
<td>[le] ugly</td>
<td>[le] their</td>
<td>[br] during</td>
</tr>
<tr>
<td>Low</td>
<td>Unround</td>
<td>Round</td>
</tr>
<tr>
<td>[la] there</td>
<td>[la] tired</td>
<td></td>
</tr>
</tbody>
</table>

### Table 4.4 Vowel contrasts in Vietnamese. Unmarked vowels have mid tone; acute accent = high tone. (http://www.phonetics.ucla.edu/course/chapter11/vietnamese/vietnamese.html).

<table>
<thead>
<tr>
<th></th>
<th>Front</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vietnamese</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>Unround</td>
<td>Round</td>
</tr>
<tr>
<td>[ti] bureau</td>
<td>[tu] fourth</td>
<td>[tu] to drink</td>
</tr>
<tr>
<td>Mid Tense</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[te] numb</td>
<td>[te] silik</td>
<td>[to] soup bowl</td>
</tr>
<tr>
<td>Mid Lax</td>
<td>Unround</td>
<td>Round</td>
</tr>
<tr>
<td>[te'] to fall down</td>
<td>[ə] favor</td>
<td>[tɔ] large</td>
</tr>
<tr>
<td>Low</td>
<td>Unround</td>
<td>Round</td>
</tr>
<tr>
<td>[aeŋ] to eat</td>
<td>[tə] we/our</td>
<td></td>
</tr>
</tbody>
</table>

### Table 4.5 Vowel contrasts of Turkish, Istanbul dialect. (http://www.phonetics.ucla.edu/appendix/languages/turkish/turkish.html).

<table>
<thead>
<tr>
<th></th>
<th>Front</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Turkish</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>Unround</td>
<td>Round</td>
</tr>
<tr>
<td>Non-high</td>
<td>Unround</td>
<td>Round</td>
</tr>
</tbody>
</table>
based on a written description, however, and it is notable that [u] is not described as contrasted with [i]. Another candidate is Norwegian, which is described as having three high round vowels (in addition to unround [i]): [byː] *town* vs. [buː] *shack* vs. [buː] *live*. However, the lip position is also different for [buː] vs. [buː]. The lips are more compressed for the former and more protruded for the latter. Thus, it is not clear if this is a case of a three-way distinction in backness or a three-way distinction in rounding.

In practice, especially in broad transcription, the symbols for the central and back vowels that match in rounding are often used interchangeably. The non-front, non-round high vowel of Turkish is sometimes transcribed as [i] and sometimes transcribed as [u], for example. The distinction between a central and back symbol should be observed when a precise narrow transcription is called for.

English is one of those languages that make use of the central region of the vowel space, so a few notes on the transcription of English central vowels are called for. Some example transcriptions are given in Table 4.6.

The transcription of the English central vowels is complicated by a number of factors. The first is the influence of stress: the symbols used for stressed and unstressed vowels are different. The second is dialectal and individual variation: from place to place and person to person, pronunciation may differ. The difference is especially acute for *r-ful* vs. *r-dropping* dialects. In *r*-ful dialects, typified by General American, [i] is pronounced at the end of a word, so that “beer,” for example, is pronounced [bɪə]. In *r*-dropping dialects, typified by the British accent used by BBC reporters, final [ɪ] is deleted or replaced by a vowel, so that “beer” is [bɪə]. The dialects of Boston, New York City, and some areas of the Southeastern US are also *r*-dropping, as is Australian English. A different set of central vowels is used in *r-ful* vs. *r*-dropping dialects. In addition, transcriptions will differ depending on how narrowly precise the linguist chooses to be. (Finally, the exact details of the circumstances under which [ɪ] is pronounced or not are complex. For our purposes in this chapter, words are assumed to be spoken in isolation.)

To begin with the simplest case, the non-front, non-round, mid, stressed vowel of English (as in “mud”) is usually transcribed as [ʌ]. (The symbol can be called “wedge.”) It represents a neutral tongue position, somewhat back of central: the tongue body moves neither up, down, back (much), or forward. It is no accident that the English “filler syllable”, when you open your mouth but don’t quite know what to say, is “uh”; that is [ʌʌʌ].

In the transcription of English words, the wedge symbol is only used for stressed syllables. In unstressed syllables, as in “panda” or “about,” the symbol [ə], *schwa*, is used for a vowel with neutral tongue position. Thus the word “abut” is transcribed [əbʊt] and “adjust” is [ədʒʊst]. (The teaching of phonetics must have been in vogue when I was in elementary school: I distinctly remember a page in my third-grade Language Arts workbook entitled “Schwa” with a picture of a panda bear on it. I guess I was daydreaming during the lesson that day, however, because for many years after I was convinced that a “schwa” was a black and white bear.)

There is quite a bit of variation in the way English unstressed vowels are pronounced, however. One source of variation is the surrounding consonants. Because unstressed vowels

<table>
<thead>
<tr>
<th>symbol</th>
<th>example</th>
<th>spelling</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>ʌ</td>
<td>[mʌd]</td>
<td>mud</td>
<td>Used in stressed syllables</td>
</tr>
<tr>
<td>ə</td>
<td>[pəˈændə]</td>
<td>panda</td>
<td>Used in stressed syllables</td>
</tr>
<tr>
<td>ə</td>
<td>[bəˈliv]</td>
<td>believe</td>
<td></td>
</tr>
<tr>
<td>ə</td>
<td>[əˈbat]</td>
<td>abut</td>
<td></td>
</tr>
<tr>
<td>ə</td>
<td>[əˈouzəz]</td>
<td>Rosa’s</td>
<td></td>
</tr>
<tr>
<td>ɨ</td>
<td>[ˈəouzəz]</td>
<td>roses</td>
<td>Used in unstressed syllables</td>
</tr>
<tr>
<td>ɨ</td>
<td>[ˈəarid]</td>
<td>rotted</td>
<td></td>
</tr>
<tr>
<td>ɜ</td>
<td>[fɜ]</td>
<td>fur</td>
<td>In “r-dropping” dialects, including Boston and BBC English</td>
</tr>
<tr>
<td>ɜ</td>
<td>[ˈbɜdən]</td>
<td>burden</td>
<td></td>
</tr>
<tr>
<td>ɜ̯</td>
<td>[ˈfɜ]</td>
<td>fur</td>
<td>In “r-ful” dialects, including General American; stressed syllables</td>
</tr>
<tr>
<td>ɜ̯</td>
<td>[ˈbɜdən]</td>
<td>burden</td>
<td></td>
</tr>
<tr>
<td>ɔ</td>
<td>[nɛvə]</td>
<td>never</td>
<td>GA pronunciation, unstressed syllables</td>
</tr>
<tr>
<td>ɔ</td>
<td>[ˈhelpə]</td>
<td>helper</td>
<td></td>
</tr>
<tr>
<td>ɔ</td>
<td>[ˈspiɾə]</td>
<td>supper</td>
<td></td>
</tr>
<tr>
<td>ɔ</td>
<td>[fəˈθə]</td>
<td>further</td>
<td></td>
</tr>
</tbody>
</table>
are short, the tongue may not move very far from the positions required by the surrounding consonants. So, for example, unstressed vowels surrounded by coronal consonants such as [t, d, s, z] may tend to be fronted and raised. A reduced vowel with a higher tongue position than [a] may be transcribed [i]. The difference between [a] and [i] is contrastive for some English speakers in at least some words. Listen carefully to your own pronunciation of pairs like “Rosa’s/rozes”, “Lisa’s/leases”, “Nida’d/needed” (where “Nida’d” is the contracted form of “Nida would”). For many English speakers, the vowel [i] is used in unstressed plural and past-tense suffixes (“rozes” and “needed” are [ɹoːzɪ], [ˈnɪdɪd]), while the vowel [ə] is used for all other unstressed syllables (“Rosa’s” and “Nida’d” are [ɹəʊzɪ], [ˈnɪdəd]). Some speakers may also use [i] in initial position, especially before a coronal. A word like “edition” may thus be pronounced [ədiʃən] (homophonous with “addition”) or [ɪdiʃən] with a higher vowel. Or some speakers may give the first vowel a little bit of stress and say [ɛdiʃən].

Things get even more complicated when transcription of the “er” sound is considered. For r-dropping dialects like BBC English, the stressed vowel in words like “bird,” “fur,” and “herd” is transcribed [ɜː]. As indicated by its position in the vowel chart, this vowel is intermediate between [ɛ] and [ə]. Unstressed “er” is simply [ə], so that “murder” is [mʌðə] and “pander” and “panda” are homophonous as [pʰændə].

For r-ful dialects like GA English, the first thing to realize is that although a word like “fur” consists of three letters, there are only two sounds. The “er” sound is not a sequence of vowel+r, but a single rhotic sound: the tongue body is in a mid-central position but the tongue front is raised (and for some speakers curled back), the same as or similar to the [i] in “run.” Try drawing out the word (if you are a speaker of an r-ful dialect): the sound is not [fʌlʌstæʒə], but [fɪlɪʃən]. So one way to write the word “fur” would be [fɜː], with a syllabic consonant. Most transcribers, however, consider this sound to be a separate, contrastive vowel of English – a mid, central, rhotacized (or r-colored) vowel – and use the symbol [ɜː]. In an unstressed syllable, as for the “-er” suffix, the symbol is [ə]. (Note the little hook reminiscent of a script “r.”) Thus, for an r-ful dialect, “murder” is [mʌðə] and “further” is [ˈfɜːðə]. Ladefoged and Maddieson note that rhotacized vowels are known to occur in only two languages – English and Mandarin Chinese – making these sounds typologically rarer than labiodental flaps.

There is another interesting effect of vowel+r sequences in many dialects of English, including General American. In syllables where a vowel is followed by [i], the 12+ vowels of English are reduced to only six: the prototypical [i, e, a, o, u], plus the syllabic rhotic itself. There is no distinction between tense and lax vowels before [i] (see Figure 4.6). Thus, in most American dialects there is no distinction between “war” and “wore,” and the three words “Mary”, “merry”, and “marry” are homophonous. (One dialect that keeps them distinct is Philadelphia.) Due to the influence of spelling, we often think we are distinguishing these words, much as we may think we are distinguishing “kitty” and “kiddie.”

4.3.5 contrasts among the low vowels

Finishing our discussion of the dimensions of vowel quality, and the symbols of the IPA vowel chart, we turn to contrasts among the low vowels. The IPA chart gives six different low vowel symbols [æ, a, ɛ, ə, o, ʌ], but it is unusual for any one language to contrast more than two.

Most languages, in fact, have only a single low vowel. This is true of the three-vowel systems like Classical Arabic, five-vowel systems like Spanish, seven-vowel systems like Italian, and nine-vowel systems like Akan. As
has already been mentioned, there is less room for movement in the pharynx than along the palate, and it is harder to round the lips when the jaw is lower. Thus, it is common for languages, even those with larger inventories, not to exploit either a backness, roundness or tongue root contrast for low vowels, and to use only a single vowel at this height. The low vowel in these languages is usually articulated with a central tongue position, neither fronted nor backed, but there can be a lot of context-dependent variation. When there is no contrast, the one low vowel is usually transcribed as [a].

In languages with two low vowels, one usually has a more fronted tongue body position than the other. In General American English, the front low vowel in “pat” is usually transcribed [æ], while the backer vowel in “pot” is transcribed with [a] (sometimes [a]). In other languages that have two low vowels, such as French, the more front vowel is not quite as far forward as the English [æ], so the distinction may be transcribed as [a] vs. [la], as in Table 4.3: [la] là “there” vs. [a] las “tired.” The vowel quality in [la] is in between that of the vowels in “lack” and “lock.”

### 4.3 In Focus

Two dialects of English furnish examples of a rare three-way contrast among the low vowels. British English adds a round version of the low back vowel: [ø]. Where General American has [pʰæk], [pʰk], [pʰak], for “pack,” “pock,” “park,” British English has [pʰæk, pʰk, pʰk]. In r-dropping Boston, these three words illustrate what may be the best case for a three-way contrast between a front, central, and back vowel at the same vowel height. The vowel in “park” (as well as in “car,” “Harvard,” and “yard”) is intermediate between [æ] and [ø]. There is thus a contrast between front [æ] in “pack,” central [a] in “park” and back [ø] in “pocket,” all unround. The only language known to me for which a four-way distinction among the low vowels is claimed, as well as the only language to use a front, low, round vowel, is Bavarian German, which is reported to contrast front unround [æ], back unround [a], front round [Œ], and back round [ø].

The symbol [ø], which is the last vowel symbol to be discussed, indicates a raised version of [a]. It may be used as an alternative to the symbol [ə] to indicate a mid, central, unreduced vowel. Japanese speakers of English, for example, tend to use [ø] where native speakers have [ə]. This mid to low central vowel is also argued to occur in Portuguese (see Table 4.7 and Figure 4.7 on p. 66).

### Table 4.7 Oral and nasal vowels of European Portuguese.

<table>
<thead>
<tr>
<th>Oral vowels</th>
<th>Nasal vowels</th>
</tr>
</thead>
<tbody>
<tr>
<td>front</td>
<td>front</td>
</tr>
<tr>
<td>[pipu]</td>
<td>[pêtre]</td>
</tr>
<tr>
<td>cask</td>
<td>[bude]</td>
</tr>
<tr>
<td>(I) owed</td>
<td>[topu]</td>
</tr>
<tr>
<td>[peke]</td>
<td>[pêtu]</td>
</tr>
<tr>
<td>(he) sins</td>
<td>[pati]</td>
</tr>
<tr>
<td>[levê]</td>
<td>[kêtu]</td>
</tr>
<tr>
<td>(he) takes</td>
<td>[pati]</td>
</tr>
<tr>
<td>each</td>
<td>[pati]</td>
</tr>
<tr>
<td>[patu]</td>
<td></td>
</tr>
</tbody>
</table>
4.4 nasality and voice quality

Thus far, we have discussed how languages make use of tongue body position (high/mid/low and front/central/back), tongue root position (advanced/retracted or tense/lax), and lip position (round/unround) to create different vowel sounds. Contrasts can also be made in nasality and in voice quality.

Commonly, vowels will become somewhat nasalized when they precede a nasal consonant. The velum, which moves rather slowly compared to the other articulators, begins opening early, in preparation for the upcoming nasal consonant, so that the vowel becomes gradually more nasal over the course of its duration. In a number of languages, however, including French, Portuguese, Polish, Chinese, Hindi, Gujarati, and others, vowels can be nasal even when no following nasal consonant is pronounced. (A nasal consonant is often written when the word is spelled, however, and may have been present historically.) Typically, the number of nasal vowels is smaller than the number of oral vowels. European Portuguese, for example, has nine oral vowels, but the inventory of nasal vowels is reduced to five (again, versions of [i, e, a, o, u]). Adding the extra resonance of the nasal cavity interferes with the acoustic cues to tongue and lip position, making nasal vowels harder to distinguish, thus leading to a smaller number of vowel contrasts. Words illustrating the oral and nasal vowel inventories of European Portuguese are shown in Table 4.7. (The vowels of Brazilian Portuguese are slightly different.) Figure 4.7 shows MRI images of an oral and nasal vowel in European Portuguese, with velum opening for the nasal vowel clearly evident. In transcription, nasality is indicated by a tilde over the vowel: [ã].

Vowels may also be distinguished by voice quality. In the unmarked case, vowels, which rely on vocal fold vibration to produce vocal tract resonance, are voiced. But in some cases vowels may have different voice qualities: they may be devoiced (or whispered), or produced with creaky voice (tense vocal folds), or breathy voice (lax vocal folds). Devoicing is transcribed with an open underdot ([ã]), creaky voice with a subscript tilde ([ã]), and breathy voice with two filled underdots ([ã]).
Sometimes voice quality differences may depend on the surrounding consonants. Unstressed vowels following aspirated stops in English, for example, may be completely voiceless: [pʰtrəʊ]. In Japanese, high vowels are devoiced when surrounded by voiceless consonants: [kuton]. As with nasality, however, vowels may sometimes contrast in voice quality independent of the consonantal context. For example, Gujarati contrasts plain (or modal) voice vs. breathy voice on vowels, independent of a contrast in aspiration on both voiced and voiceless consonants (see Table 4.8). Mpi (a language of Thailand) contrasts modal vs. “tense” voice: [sī] to roll vs. [sị] to smoke (both with mid-rising tone, see Section 4.6 below). Jalapa Mazatec (Mexico) makes a three-way contrast: [jά] tree vs. [ją] he carries vs. [ją] he wears.

### Table 4.8 Voice quality contrasts in Gujarati. (Ladefoged & Maddieson, p. 315.)

<table>
<thead>
<tr>
<th>Plain vowel</th>
<th>Breathy vowel</th>
<th>Aspirated consonant</th>
</tr>
</thead>
<tbody>
<tr>
<td>[pəɾ] last year</td>
<td>[pəɾ] early morning</td>
<td>[pʰədʒ] army</td>
</tr>
</tbody>
</table>

Vowels may also contrast simply in length. Many factors influence the amount of time it takes to articulate a given segment. A person may speak faster or slower, of course. Some differences are inherent in the articulation: low vowels, for which the mouth has to open wide, take longer than high vowels, for which little movement is necessary. As was noted above, the tense vowels of English are longer than lax vowels, and stressed vowels are longer than unstressed vowels. In some languages, however, two vowels may differ in length alone, without any other difference in stress or tongue position: the long segment and short counterpart are exactly the same, except that the former is held for a longer period of time, an extra “beat.” These long segments may be written with a double symbol ([aa]), or with a colon after the usual symbol ([aː]).

Japanese and Arabic are languages that make length distinctions in both vowels and consonants. In Tokyo, you want to be careful to order [bɨɾu] a beer rather than [bɨɾu] a building or to ask directions to a certain [tɔɾi] road rather than [tɔɾi] bird. A linguist needs to be clear she’s looking for a new [oɾo] sound, not [oɾo] husband. In Arabic, length distinctions often have grammatical import: [kata],[b] is wrote and [kata],[b] is corresponded with; [darasa],[b] is studied and [dar:asa],[b] is taught. English has no true long consonants as found in Japanese and Arabic. English can create long consonants when two words come together – compare “stop Paul” [stəpɔl] to “stop all” [stəpɔl] – but we do not distinguish long and short consonants within words. When double consonants are written, for example in “supper” vs. “super,” they actually tell us about the quality of the vowel, not the length of the consonant.

**Diphthongs** are vowels that require a change in tongue and/or lip position, often a drastic change, over the course of their duration. (The word is from Greek: [di] two, [θονγος] sounds.) They contrast with monopthongs, which hold a relatively steady state. Rarer triphthongs have a sequence of three vowel qualities within a single syllable. English has three true diphthongs: [aɪ] as in “hide” and “buy”, [au] as in “crowd” and “how,” and [ɔi] as in “void” and “boy.” The diphthongs might be thought of as contour vowels, in the way that affricates are contour consonants: they start out in one position and end in another. The English diphthongs move across the whole vowel space: from central back to high front in [aɪ], for example. Any combination of vowel sounds can be made into a diphthong, however. Dutch, for example, has [ɛi, æy, ɑo], in addition to monopthongal [i, y, ı, ɛ, e, o, u, o ɔ, ɑ]. In Vietnamese, a diphthong can begin from just about any of the 11 steady-state vowel qualities (see Table 4.4) and move up and forward to [i], up and back to [u], or from a more extreme position to the center. Vietnamese also has triphthongs, such as [iəʊ].

---

**4.5 length and diphthongs**
As was noted in Chapter 2, the crucial difference between a true diphthong and a diphthongized vowel is not the extent or direction of tongue body movement, but whether or not the movement is contrastive. In a true diphthong, deleting the offglide changes the meaning of the word: \[k^eət\] is distinct from \[k^eət\] and \[waʊnd\] is distinct from \[wand\]. On the other hand, no English words are distinguished by the difference between \[e\] and \[æ\]. That's not to say that the difference isn't noticeable. Pronouncing the tense vowels without an offglide is a mark of a non-native speaker of English, and pronouncing \[e\] and \[o\] with offglides is one of the strongest marks of an English accent in French or Spanish. Changing one's typical vowel pronunciation to match a vowel in another language that is similar but not quite the same is one of the hardest tasks confronting a second-language learner (see Chapter 20). It is easier for native English speakers to learn to correctly pronounce the French word for “read” (\[l\i\]) than it is to learn to stop asking for \[k\æfə\ ou \l\i\]\ when they want \[k\æfə\ ou \l\i\].

4.6 tone

Last, but far from least, vowels may contrast in pitch. The use of pitch differences to create different words is called tone. Strictly speaking, tone is a property of a unit larger than a segment, usually a syllable or perhaps a word, but the differences in pitch are most often produced and perceived during the vowel. Tone and other linguistic uses of pitch are discussed in more depth in Chapter 17; only a few general points, particularly with regard to transcription, are covered here.

The pitch of the voice carries a lot of different kinds of information. It can tell you whether the speaker is a male or female, a large person or small, old or young. High pitch can tell you that a person is frightened; low pitch that she is angry. This sort of information isn’t really linguistic, however, but physical or emotional: you could get the same messages from a dog’s low-pitched growl or a puppy’s high-pitched whimper.

English also makes various linguistic uses of pitch. It was noted above that stressed vowels tend to have higher pitch than unstressed. We also use pitch differences to convey different kinds of sentence-level or discourse-level meanings. Yes/no questions will usually have rising pitch, for example, while statements have falling pitch. We can use a more complex pitch pattern to indicate emphasis or incredulity. These discourse-level uses of pitch are called intonation.

Intonation distinguishes different kinds of sentences, or focuses attention on a particular word. As an example of some uses of intonation in English, try reading the following sentences out loud:

“Well, I thought it was just a cat.”

“I thought it was a mountain lion!”

If you read with feeling, you should have noticed that the pitch of your voice moved in different directions on the word “cat.” In the first, pitch goes up, indicating a question. In the second, pitch falls, indicating a statement or confirmation. In the third, a more complicated rise-fall pattern indicates incredulity. (Typographically, we indicate these different “readings” with a question mark, period, and italics, respectively.) In each case, the sequence \[k\æt\] refers to the same object, a feline. The pitch differences indicate only the role that the refer-
ence to the feline is playing in the current conversation: asking for information about the cat, providing it, or expressing disbelief regarding the information offered.

All languages use intonation to some extent, though the patterns and meanings are not exactly the same cross-linguistically (see Chapter 17). But in addition to intonation, most languages also use pitch to distinguish different words. In English, whether you say \([\text{kæt}]\) with a rising pitch or falling pitch, the word still refers to a feline; the only difference may be attitude toward the animal. In Thai, if you say \([\text{kʰaː}]\) with rising pitch, it means \(\text{leg}\); but if you say it with falling pitch, it means \(\text{value}\). The two words are as different as “cat” and “cot” to an English speaker. This use of pitch, to distinguish different words, is tone.

The idea of tone may seem strange to English speakers, and we may associate the term with “exotic” languages. Yet by one estimate 70% of the languages of the world are tonal, including most of the languages of Africa, and many of the languages of East Asia and of the Americas. (The major European languages and their relatives are in fact exceptional in not having tone. No tone languages have been reported in Australia, either.) There are certainly more native speakers of tone languages in the world than there are speakers of non-tone languages, in large part because Mandarin Chinese, native language of approximately one in six humans, is tonal.

Figure 4.8 gives the IPA symbols and diacritics for transcribing tone. As shown in the figure, tone can be transcribed using diacritics over the vowel symbols. Alternatively, if more precision is required, the linguist can use tone letters written after each syllable. A tone letter is a schematic diagram of the pitch pattern: the vertical line indicates the speaker’s overall pitch range, while the horizontal or slanted line indicates the level and movement of the pitch pattern associated with the tone.

Implicit in these descriptions is that the pitch of tones is always relative to a speaker’s overall range. The absolute pitch of a high-toned syllable spoken by a large male might be lower than the absolute pitch of a low-toned syllable spoken by a small female. But each speaker’s high tones will be higher than that speaker’s low tones. The pitch of tones is also sensitive to the context of surrounding syllables and to position in the sentence. Pitch tends to get lower over the course of a sentence, for example: sometimes gradually and sometimes in more dramatic downsteps, so that the absolute pitch of high tone at the end of a sentence might be lower than the absolute pitch of a low tone at the beginning. So the terms “high” and “low” always mean “relatively high” or “relatively low” for that speaker in that position.

In Figure 4.8, tones are divided into level tones and contour tones. Level tones do not necessarily have a perfectly level pitch, but they create contrast based on relative pitch height, not direction of pitch change. Languages that use only level tones are called register tone languages. More complex systems add contrasts that include crucial pitch rises and falls. These are called contour tone languages.

The simplest tonal system is a two-way contrast between high and non-high. Languages such as Margi (West Africa), Setswana (Southern Africa), Navajo (North America) and Mixtec (Central America) are systems with just two contrastive tones. To give just one example, in Margi, \([\text{ʃú}]\) with high pitch, means \(\text{tail}\) and \([\text{ʃù}]\), with lower pitch, means \(\text{to dry up}\). An acute accent indicates a high tone and a grave accent indicates a low tone. (My mnemonic for remembering the difference is that you need to look at where the accent mark ends.) Yoruba (West Africa) is a language that contrasts three pitch levels: \([\text{wá}]\) to come, \([\text{wà}]\) to look, \([\text{wà}]\) to exist. A mid tone may be indicated with a macron over the vowel, or the vowel may be left unmarked.
Table 4.9 exemplifies the tones of Thai, a contour tone language. In addition to three level tones, Thai adds rising and falling pitch patterns. The diacritic for a rising tone is (‘), the symbol for low (‘) followed by high (‘), while the diacritic for the falling tone is (‘) which corresponds to high followed by low.

The complexities of tonal systems and tonal contrasts are addressed at greater length in Chapter 17.

### Table 4.9: The tones of Thai

<table>
<thead>
<tr>
<th>high</th>
<th>mid</th>
<th>low</th>
<th>falling</th>
<th>rising</th>
</tr>
</thead>
<tbody>
<tr>
<td>ná:</td>
<td>ná:</td>
<td>ná:</td>
<td>ná:</td>
<td>ná:</td>
</tr>
<tr>
<td>aunt</td>
<td>rice</td>
<td>custard apple</td>
<td>face</td>
<td>thick</td>
</tr>
<tr>
<td>kʰáː:</td>
<td>kʰáː:</td>
<td>kʰàː:</td>
<td>kʰáː:</td>
<td>kʰàː:</td>
</tr>
<tr>
<td>to trade</td>
<td>to be stuck</td>
<td>galangal spice</td>
<td>value</td>
<td>leg</td>
</tr>
</tbody>
</table>

4.7 positional variants of the vowels of English

As we did in Chapter 3 for consonants, we end this chapter with a summary of the positional variants of the vowels of English, for those readers with a particular interest in English pronunciation. The rules exemplified here apply in most or all of the varieties of English. Keep in mind, however, that the vowels of English differ much more than the consonants based on dialect: the transcriptions here may not exactly match your pronunciation, though the general rule probably still applies. There is more variation in English vowels than can be squeezed onto the end of a general chapter, and the vowel systems of the major English dialects are discussed in Chapter 19.

i. Vowels are nasalized before nasal consonants

- ban [b ăn]\n- tome [tʰóʊm]\n- sing [sʃɪŋ]

ii. Tense vowels are diphthongized

- beak [b ijk]\n- bake [b ɛɪk]\n- rope [róʊp]\n- soup [suwp]

iii. Vowels are lengthened in open syllables, and before voiced consonants. Note that, with the exception of [o], lax vowels do not occur in open syllables. For diphthongs and diphthongized vowels, most of the lengthening occurs on the offglide.

- beak [b ijk]\n- bead [b ijd]\n- bee [b iː]\n- boost [b uʌst]\n- bood [b uwd]\n- boo [b uː]\n- sought [s oʊt]\n- sawed [s awd]\n- saw [s ɔː]\n- bite [b an t]\n- bide [ba d]\n- buy [ba r]
iv. Stressless vowels ([ə]) may be deleted, especially preceding sonorants

risen [ɹɪzn]
I can go [aɪkŋoʊ]
general [dʒənərəl]
camera [kəˈmɛrə]

v. Rhotic dialects (General American) have rhotacized mid central vowels in “bird” and “fur”: [bɔd̪], [fɔr̪]. Non-rhotic dialects (RP, Australia, New Zealand, South Africa, some parts of the American South and New England) have mid central [ɜ]. In unstressed syllables with orthographic “r”, rhotic dialects have [ə], non-rhotic have [ɜ].

vi. In many rhotic dialects, there is no distinction between tense and lax vowels before [ɪ], though some dialects (Irish, Philadelphia), the distinction is maintained for the mid vowels. Vowel quality preceding [ɪ] is usually in between that of the tense and lax vowel, but may be closer to one or the other depending on dialect. In non-rhotic dialects, the [ɪ] is replaced by [ə], so that “beer” [bɪə] rhymes with “idea.”

---

**chapter summary**

- The constrictions of the consonants are superimposed on the open articulations of the vowels. Consonants and vowels are organized into syllables, which may be more or less prominent (stressed or unstressed).
- Daniel Jones created the cardinal vowel system as a map for comparing vowel qualities across languages.
- Vowels may contrast in height, backness, tense/lax, rounding, nasality, voice quality, length, and tone. Different vowel systems choose different dimensions of contrast.
- In the unmarked case back vowels are round, front vowels are unround, low vowels are lax, high vowels are tense, and vowel systems tend to be peripheral. However, languages may choose the more marked options, such as front round vowels, high lax vowels, and central vowels, in addition to the unmarked ones.
- Section 4.7 describes some of the positional variants of the dialects of English, including rhotic and non-rhotic pronunciations.

---

**further reading**

The books mentioned at the end of Chapter 3 discuss vowels as well as consonants.

For more information on linguistic tone, see Chapter 17 and Yip, M. 2002. *Tone*. Cambridge: Cambridge University Press.
A MAP OF THE VOWELS

review exercises

1. Define the following terms:
syllable
nucleus
onset
coda
syllabic consonant
stressed syllable
vowel harmony
peripheral vowel
vertical vowel inventory
r-colored vowel
tone letters
downstep

2. Compare and contrast each of the following pairs (or triplets) of terms:

   primary cardinal vowel  secondary cardinal vowel
advanced tongue root retracted tongue root
r-ful dialect r-dropping dialect
modal voice breathy voice  creaky voice
monophthong diphthong  triphthong
tone intonation
level tone contour tone

3. Why was the cardinal vowel system an important advance in cross-linguistic description? Why has it fallen out of use?

4. Why is it unusual for a language to make a tense/lax distinction among the low vowels?

5. Explain why the inventory \([i, e, a, o, u]\) is extremely common, while \([\text{y, ø, ð, ʏ, ʊ}]\) is unattested.

6. Add three more example words or phrases to each of the sets in Section 4.7.

7. Revisit your narrow transcriptions of the following sentences. Add in any positional variation in the vowels.

   a. Would you meet me at the station at five?
   b. My husband should have bought potatoes.
   c. Will he be kept back?
   d. They all wanted to be in the pictures.
   e. In fact, I hated those times.

8. Make a narrow transcription of your own pronunciation of the following sentences, indicating positional variation of both vowels and consonants. Is your dialect r-ful or r-dropping?

   a. Your brother heard the concert.
   b. Would you prefer beer or ale?
   c. I was shocked to find a shark in the shack.
   d. Her mother is searching for a cure for cancer.
   e. The senator from Ohio is never boring.
further analysis and discussion

9. Fill in the descriptions for each of the following vowels.

<table>
<thead>
<tr>
<th></th>
<th>Height</th>
<th>Backness</th>
<th>Rounding</th>
<th>Tense or lax</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ø</td>
<td></td>
<td></td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>i</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>o</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

10. Provide the IPA symbol that matches each description.

<table>
<thead>
<tr>
<th></th>
<th>Height</th>
<th>Backness</th>
<th>Rounding</th>
<th>Tense or lax</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>front</td>
<td>unround</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>high</td>
<td>back</td>
<td>round</td>
<td>tense</td>
<td></td>
</tr>
<tr>
<td>mid</td>
<td>front</td>
<td>unround</td>
<td>tense</td>
<td></td>
</tr>
<tr>
<td>high</td>
<td>back</td>
<td>unround</td>
<td>tense</td>
<td></td>
</tr>
<tr>
<td>mid</td>
<td>back</td>
<td>round</td>
<td>lax</td>
<td></td>
</tr>
<tr>
<td>high</td>
<td>front</td>
<td>unround</td>
<td>lax</td>
<td></td>
</tr>
<tr>
<td>mid</td>
<td>back</td>
<td>unround</td>
<td>tense</td>
<td></td>
</tr>
<tr>
<td>mid</td>
<td>central</td>
<td>unround</td>
<td>lax</td>
<td></td>
</tr>
</tbody>
</table>

11. The vowels in each of the following pairs differ in only one dimension: height, tense-ness, backness, rounding, nasality, voice quality, length, or tone. Indicate the change. The first two have been done.

a. __i__ to __ɪ__  tense to lax
b. __ø__ to __o__  nasal to oral
c. __i__ to __y__  ____________ to ____________
d. __ε__ to __ʌ__  ____________ to ____________
e. __u__ to __o__  ____________ to ____________
f. __æ__ to __a__  ____________ to ____________
g. __e__ to __ɛ__  ____________ to ____________
h. __ɑ__ to __ɒ__  ____________ to ____________
i. __ɛ__ to __ɛ̃__  ____________ to ____________
j. __u__ to __ɯ__  ____________ to ____________
k. __o__ to __o̞__  ____________ to ____________
l. __ɔ__ to __ɔ̃__  ____________ to ____________

(Continued)
12. Write the symbol that would result from the change. The first two have been done.
   
   a. Make [ə] unround  ___ʌ___
   b. Make [e] lax  ___ɛ___
   c. Make [i] rising tone  ____________
   d. Make [o] creaky  ____________
   e. Make [æ] round  ____________
   f. Make [y] back  ____________
   g. Make [u] tense  ____________
   h. Make [o] front  ____________
   i. Make [a] long  ____________
   j. Make [ã] low tone  ____________
   k. Make [ɛ] high  ____________
   l. Make [ɔ] low  ____________
   m. Make [i] central  ____________

**Further research**

13. Describe the vowel system of a language that was not discussed in this chapter, either one you know or one you learn about in a reference book. Transcribe the vowels using IPA, and create a chart of the vowel space. What dimensions of contrast does the language use?

14. Make a narrow transcription of a passage of English poetry or prose.

15. Ask a non-native speaker, or someone you perceive as having a different accent from yours, to read the sentences in exercises 7 and 8 above. Make a narrow transcription of their pronunciation. What differences do you notice?

**References**


Other sources, listed in alphabetical order by language:

**Akan**

**Bavarian German**

**Dutch**
Portuguese

Thai