3.3: Vowels (Part 1)

The goals of linguists, if you remember, are to describe what people know about their language and to figure out how languages are similar and different. For word forms, specifically for phonemes, this means that we must describe both how speakers and hearers distinguish phonemes within a given language and how individual phonemes and systems of phonemes differ between languages. To satisfy both of these goals, we will be looking for ways to describe the variation between speech sounds. Just as we saw that categories of things could be described in terms of values along different dimensions, we will be looking at dimensions of sound, dimensions that allow us to make distinctions between phonemes within and between languages.

Phonetic Symbols

Exercise \(\PageIndex{1}\)

If we want to write down how words in different languages are pronounced, why not just use the letters of English to do this? Could we use the letters of English to write down the pronunciation of English words? What if we want to distinguish the pronunciation of words in different English dialects?

Before we discuss the specific phonemes of languages, we need to decide how to represent the sounds. Let's briefly consider English. Like other spoken languages, English has a set of vowel and consonant phonemes that its speakers use to make the words of their language. Like other written languages, English also has a way of graphically representing spoken word forms. The English writing system is an example of an alphabetic writing system in which phonemes are represented by characters or combinations of characters. But, for various historical reasons, the English writing system does this very imperfectly. Consider the words way, weigh, wait, and wake. These words share the same vowel phoneme, but it is spelled in four different ways. That is, a single phoneme may be represented using different letters or combinations of letters. Now consider the words bother, brother, border, and voter. These words share the
letter o, but it represents four different vowels, each a different phoneme. That is, a single letter may represent multiple phonemes. We can conclude two things from this.

- We must be careful not to confuse sounds with letters; the letter o is not a vowel, though it is used to represent vowels.
- We cannot rely on English spelling when we are concerned with the pronunciation of English words.

Because we will need a way to represent the phonemes of English and other languages unambiguously, we must rely on a set of symbols for this that are not used quite like the alphabets of any alphabetic writing systems. Symbols representing the basic sounds, or phones, of spoken languages, are called phonetic symbols. Linguists use a set of phonetic symbols called the **International Phonetic Alphabet (IPA)**. The symbols in the IPA are based on the characters in the Roman alphabet, which is also the basis for the writing systems of many languages, including English, Spanish, Lingala, and Tzeltal. The IPA is overseen by the International Phonetics Association. Here is a full chart of the IPA symbols that you can click on to hear some of the sounds. Note that you probably won't understand this table until after you've studied this section, the next one on English consonants, and the one after that on consonants in other languages. I will use a subset of the IPA symbols in this book. (To view many of these symbols, you will need Unicode support in your browser. Go to this appendix to see if your browser has this support.)

### Symbols for Phonemes vs. Symbols for Phonetic Details and Language Differences

I will use phonetic symbols in two ways. Much of the time they will be used to distinguish the phonemes within a language. The important thing will be to make sure each phoneme has a unique symbol; which symbol we use is not as important. When the symbols are used in this way, they will be enclosed by slashes. As an example, my pronunciation of the word phonemes is /'fonimz/. I will also be using symbols in another way: to represent the actual sound that is produced in a given situation, rather than a category of sounds (a phoneme, that is), and to represent sound differences between different dialects and different languages. For this purpose we will have to be more careful about which symbols are used for which phones. When the symbols are used in this way, they will be enclosed in square brackets. As an example, my (more detailed) pronunciation of the word phonemes is ['founimz]. It is important to note that the symbols are at best only an approximation to the actual phone that is used in any given situation because the space of possible phones is enormous, possibly infinite, while we only have a relatively small, finite set of phonetic symbols.

### Vowel Features

Exercise 1

Notice what you do with your mouth when you pronounce the names of the letters "a", "e", "i", and "o". What features of your mouth seem to distinguish these vowels from one another?

A complete account of the sounds of a spoken language (or of spoken language in general) would have to make reference both to the way sounds are produced, articulation, and the way they are perceived, their **auditory** properties. In this book we will not have time to go into the auditory properties of speech in any detail, but we cannot neglect them completely. Differences in articulation are obviously pointless if they aren't reflected in the way the results sound to hearers. In fact, two quite different ways of articulating a sound can sometimes produce the same auditory effect. From the perspective of a hearer, those would have to belong to the same phoneme.
In order to understand how vowels and consonants work, we need to know a little bit about the physical apparatus that is used to produce them. The figure below shows a side view of the vocal tract, with labels for some of the parts that we'll be discussing in this section and the next sections.

The heart of the system is the larynx, located in the throat under the Adam's apple; the larynx contains the vocal cords, two pieces of flesh that can be loosened or tightened. When air from the lungs passes through the larynx, it may be allowed to pass unhindered, as when we are breathing out. Or the vocal cords may be tightened and brought together so the air causes them to vibrate. This vibration, or voicing, is what distinguishes normal speech from whispering and certain speech sounds from others. Here is a slow-motion movie, produced at the UCLA Phonetics Laboratory, of the vocal cords vibrating during speech.

### How the Vocal Tract is Superior to a Trumpet

Given an outward airstream and vibrating vocal cords, we have a device that is a little like a brass musical instrument. The vocal cords are like the vibrating lips of the musician, and the region of the vocal tract between the larynx and the opening of the mouth (the yellow region in the figure), the oral cavity, is like the body of the horn. What makes the vocal tract far more versatile than a trumpet or a trombone, however, is that speakers can change the shape of the instrument as they are playing it, producing a great variety of sounds.

What distinguishes vowels from each other auditorily is the precise shape and volume of the oral cavity, and the main organ involved in adjusting the shape and volume is the tongue. Given the constraints on the shape of the vocal tract and the way the tongue is manipulated, there are three extreme positions the tongue can take that lead to vowel sounds. If the body of the tongue is pushed forward and toward the roof of the mouth, we get this vowel, something like the vowel in the word *beat* and written with the symbol [i]. If the body of the tongue is lowered while the back of the tongue is pushed toward the roof of the mouth, we get this vowel, something like the vowel in the word *boot* and written with the symbol [u]. Note that in both cases the tongue cannot approach the roof of the mouth too closely, or we get the sound of
friction near the contact, that is, a consonant rather than a vowel sound. Finally, if the body and back of the tongue are both moved away from the roof of the mouth, at the same time narrowing the region that is between the larynx and the mouth (the "pharynx") we get this vowel, something like the vowel in the word hot (as pronounced by most Americans) and written in this book with the symbol [ɑ]. (The IPA symbol for this sound is the variant of a in the Roman alphabet that resembles this Greek letter.) The configurations of the vocal tract for these three vowels are shown in the figures below.

How can we turn this informal description of what is going on in the vocal tract into a more compact description that makes it clear how vowels differ from one another? First, [i] and [u] share one property; they are both associated with a
relatively narrow gap between the tongue and the roof of the mouth (and a relatively wide pharynx). They differ in this way from [ɔ], which has the widest possible gap between the body of the tongue and the roof of the mouth (and a relatively narrow pharynx). Just as we spoke of dimensions along which categories of meaning varied, we will speak here of dimensions along which these sound categories vary. The dimension that distinguishes [i] and [u] from [ɔ] we'll call height. So far we have seen two different values on this dimension, one for [i] and [u] and another for [ɔ].

Another, equivalent way to talk about the difference between these vowels is in terms of features; [i] and [u] have the feature high, while [ɔ] has the feature low. A feature always corresponds to some value on a dimension. For example, with the domain of personal pronouns we saw that the gender dimension had possible values of masculine, feminine, and neuter; each of these can also be seen as feature. All of the things having a certain feature can also be viewed as a category; thus there is the category of masculine pronouns and the category of high vowels.

**Turning a Huge Number of Mouth Configurations into a Small Set of Categories**

Dimensions can be discrete, that is, with a finite set of possible values, or continuous, that is, with an open-ended and possibly infinite set of possible values. Gender is a discrete dimension — there are only three possible values — while vowel height is a continuous dimension because we can theoretically put the highest part of the tongue at any point between the lowest and highest possible points. The fact that height is a continuous dimension means that values high and low identify approximate points along the dimension and that a given vowel is only relatively high or low. But note that the fact that a dimension is continuous does not imply that people can distinguish all of the possible values; they cannot. For example, speakers can probably reliably produce and hearers can probably reliably distinguish no more than about five different vowel heights. Note also that what is a continuous dimension in the world may be treated as discrete within language (or elsewhere in cognition). Thus within a given language, as we'll see soon, only a small number of height distinctions are made.

The dimension of height distinguishes [i] and [u] from [ɔ], but it does not distinguish [u] from [i]. Looking at the diagrams of the vocal tract again, we can see that these two vowels differ mainly in terms of where the narrowest gap between the tongue and the roof of the mouth is. For [i] it is as far forward as it is possible for the tongue body to move and still produce a vowel sound; for [u] it is as far backward as it is possible for the tongue body to move and still produce a vowel sound. This dimension is called backness; again we have seen (so far) two possible values on this dimension: front, as for [i], and back, as for [u]. But note that, like height, backness is a continuous dimension, so front and back are labels for approximate regions along the dimension. For [ɔ], the closest point of approach between the tongue body and the roof of the mouth is near the back of the mouth, so this is also considered a back vowel. Summarizing, the two articulatory dimensions of height and backness represent the location in the mouth of the closest point of contact between the tongue body and the roof of the mouth.

It turns out that these two articulatory dimensions correspond roughly to two fundamental auditory dimensions. That is, as a speaker changes the height of a vowel, the way that vowel sounds to a hearer changes along one auditory dimension, and as a speaker changes the backness of a vowel, the way that the vowel sounds to a hearer changes along another auditory dimension. However, some recent research has shown that the correspondence between the two articulatory and the two auditory dimensions is only approximate. For vowels other than the extreme ones we've been discussing, there are often multiple articulatory ways of achieving the same auditory effect; precisely adjusting the position of the closest approach between the tongue body and the roof of the mouth is only one of these ways. Thus in a
sense it is a bit of an over-simplification to characterize vowels in terms of their precise height and backness. However, height and backness still seem to provide the simplest articulatory description of vowels, and I will follow this traditional approach in what follows.

If every vowel has a height value and a backness value, then we can visualize all of the possible vowels in a two-dimensional space known as vowel space. Given the constraints of the mouth and tongue, it turns out that vowel space has roughly the shape of a trapezoid; this is shown in the figure below, along with some of the terms used to refer to different regions within vowel space. Recall that these terms are only for convenience; there are really very many, maybe even an infinite number of, possible positions within this space.

![Vowel Space Diagram](https://socialsci.libretexts.org/Bookshelves/Anthropology/Linguistics/Book%3A_How_Language_Works_(Gasser)/3%3A_Word_Forms_-_Units/3.3%3A_Vowels_(Part_1)

Vowels of Spanish

**Spanish has Five Vowels, None Quite like English.**

We’ll start with a relatively simple vowel system, that of Spanish. Spanish has five vowel phonemes, represented in the Spanish alphabet by the letters a, e, i, o, and u. The same characters are conventionally used to represent the vowels themselves. Listen to the vowels in these five Spanish words, and try to pronounce them yourself (here and elsewhere in this book, we will focus on Western Hemisphere, rather than European (Castilian), Spanish): piso, puso, peso, poso, paso. You may not be aware of the position of your tongue with respect to the roof of your mouth, but you should be aware of how the position of your tongue changes as you change from one vowel to another and how similar particular pairs of vowels are to each other in terms of tongue position. The figure below shows the approximate positions of the Spanish vowel phonemes in the vowel trapezoid. The positions are approximate because each phoneme is a category with a fairly wide range of possible pronunciations, depending in particular on which consonants appear before or after the vowel. You can click on each vowel in the figure to hear its pronunciation.
We can describe each of the Spanish vowels along the two dimensions of height and backness. For Spanish we need three values of height and two values of backness to do this. There is a further dimension that by itself does not distinguish any of the vowels, but it will be useful for us in discussing how languages and dialects differ from one another. This is the degree of lip rounding that accompanies the vowel. In Spanish /u/ and /o/ are the only vowels accompanied by significant lip rounding. For /i/ and /e/ the lips take the opposite position: they are spread. I will treat lip rounding and spreading as a single dimension. The figure below shows roughly where each of the Spanish vowels falls along the three dimensions of height, backness, and rounding. Each vowel is represented by three small circles of the same color, one for each dimension.

Another way to depict the differences is in a table that gives a label (feature name) for the value of each vowel on each dimension. I will use the symbol "∅" to indicate that a phone has no value or a default value for a dimension.

<table>
<thead>
<tr>
<th></th>
<th>/i/</th>
<th>/e/</th>
<th>/u/</th>
<th>/o/</th>
<th>/a/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height</td>
<td>high</td>
<td>mid</td>
<td>high</td>
<td>mid</td>
<td>low</td>
</tr>
<tr>
<td>Backness</td>
<td>front</td>
<td>front</td>
<td>back</td>
<td>back</td>
<td>back</td>
</tr>
<tr>
<td>Rounding</td>
<td>spread</td>
<td>spread</td>
<td>rounded</td>
<td>rounded</td>
<td>∅</td>
</tr>
</tbody>
</table>
From the table, we can see that height is a **contrastive dimension** for Spanish vowels. That is, if we change nothing but the height of a vowel phoneme, we get a different phoneme: lowering the height of /i/ gives us /e/; raising the height of /o/ gives us /u/. The situation is a little more complicated for the other two dimensions because they tend to change together. Changing only the backness of a Spanish vowel phoneme does not give us a Spanish vowel at all. If we make /u/ a front vowel, we also have to make it spread to get /i/. If we make /e/ a back vowel, we also have to make it rounded to get /o/. If we make /a/ a front vowel, we also have to make it mid to get /e/. The same holds if we attempt to change only the rounding of a Spanish vowel. To distinguish all of the Spanish vowels, we actually need only *either* backness or rounding, but not both; that is, all we’d need to know about a Spanish vowel would be its values on two dimensions to know which vowel it is (verify this for yourself from the table). So we can call backness contrastive or rounding contrastive, or we can call the combination of backness and rounding contrastive. The point is that the system is **redundant**; it provides more information than is actually needed to make the distinctions.

**Redundancy Can Come in Handy in Case the Hearer Misses Something.**

Redundancy is common in language. We will see further examples with English vowels and in the grammar of various languages. Redundancy is a Hearer-oriented feature of language. When two or more different dimensions agree with one another, the Hearer can still extract the meaning even if they fail to note what is going on one of the dimensions. So with Spanish vowels, if the Hearer detects that a vowel is high and back but misses the fact that it is rounded, they can still know which vowel it is (/u/). The same would be true if they only detected that the vowel was high and rounded, but not that it was back.

**Vowels of Japanese**

The vowels of Standard Japanese are quite similar to those of Spanish. There are five phonemes centered on roughly the same positions in the vowel trapezoid as for Spanish. There are two differences. First, the high back vowel, /u/ in Spanish, is not normally rounded in Japanese; in Japanese it sounds like this. When we need to make it clear that it's an unrounded high back vowel we are discussing, we use the symbol [ɯ], but /u/ is conventionally used to represent this Japanese vowel phoneme.

**Unlike in English, in Japanese Long Vowels are Just like Short Vowels, Only Longer.**

A second difference between Spanish and Japanese is more significant. In Spanish, vowels tend not to vary too much in length, and how much they vary depends on the dialect. In any case, if the length of a vowel is changed in a word, we get the same word. So if a Spanish speaker pronounces the word *peso* with an extra-long /e/, the difference will probably be noticeable to a hearer, who might find it a bit odd but would not interpret it as a different word. In other words, vowel length is *not* contrastive in Spanish. Japanese differs in this regard. If we take a Japanese word with a short vowel and lengthen that vowel, we get potentially a different word. For example, *su* means 'nest', *suu*, with the same vowel, roughly twice as long, means 'number'; *koke* means 'moss', *kokee* means 'solid', *kookee* means 'succession'. (Note that ee and oo here represent long versions of e and o, not the sounds these letters would represent in English.) That is, vowel length is contrastive in Japanese. Note that this does not mean that changing the length of any vowel in a Japanese word must yield a different Japanese word, only that it potentially does. Thus *mise* 'store' is a word, but *miise* is not; *miira* 'mummy' is word, but *mira* is not.
There are two ways to analyze Japanese long vowels. We could consider each to be a separate phoneme, making ten vowel phonemes all together. Alternatively we could consider each long vowel to be two short vowels in succession. For our purposes, the choice doesn't really matter, and I will go with the second alternative (which is the usual way to treat Japanese vowels).