

Sounds and Languages

1.1 Languages Come and Go

Once upon a time, the most important animal sounds were those made by predators and prey, or by sexual partners. As mammals evolved and signaling systems became more elaborate, new possibilities emerged. Nowadays undoubtedly the most important sounds for us are those of language. Nobody knows how vocal cries about enemies, food, and sex turned into language. But we can say something about the way the sounds of languages evolved, and why some sounds occur more frequently than others in the world's languages.

Although we know almost nothing about the origins of language, we can still consider the evolution of languages from a Darwinian point of view. Remember that Darwin himself did not know anything about the origin of life. He was not concerned with how life began but with the origin of the various species he could observe. In the same spirit, we will not consider the origin of language; but we will note the various sounds of languages, and discuss how they got to be the way they are. We will think of each language as a system of sounds subject to various evolutionary forces.

We will begin by considering why people speak different languages. There are many legends about this. Some say it was because God was displeased when the people of Babel tried to build a tower up to heaven. He smote them so that they could not understand each other. Others say that the Hindu God Shiva danced, and split the peoples of the world into small groups. Most linguists think that languages just grew apart as small bands of people moved to different places. We know very little about the first humans who used language. We do not even know if there was one origin of language, or whether people started talking in different parts of the

world at about the same time. The most likely possibility is that speech developed in one place, and then, like any wonderful cultural development, it spread out as the advantages of talking became obvious.

People speak different languages because languages change, often quite rapidly. Elderly people cannot readily understand what their grandchildren are saying, and the same is true in reverse. My granddaughter does not know what I'm talking about when I mention a *jalopy* (old car) or a *davenport* (couch). New words are always needed for new things like *email*, *texting*, and *television*, which my grandfather did not know. Any change in living conditions will bring changes to the language. Time itself is often sufficient to bring about changes. When people are isolated from their neighbors, living in places where travel is difficult, they develop new ways of speaking. Even when travel was comparatively simple, as it is along rivers in many tropical areas, prehistoric groups became independent. If the land provided sufficient food, they had no need to trade or interact with their neighbors. When a small group lives by itself it develops its own way of speaking after only a generation or so, producing a new dialect that its neighbors will understand only with difficulty. In a few hundred years the group will have a new language which is different from that of their ancestors and of everybody else around them.

Languages come and go. The language you are reading now, English, did not exist 1,500 years ago. The people in England spoke a Celtic language at that time. Then the Angles and Saxons and other tribes invaded, bringing with them their own Low German dialects, which gradually evolved into English. English may last another 1,500 years, but, like Latin, it may disappear as a spoken language more quickly.

Historical forces have produced about 7,000 languages in the world today, but in 100 years or so there may be less than half that number. It is worth examining why languages are disappearing and considering whether it is a good or bad thing. At the moment nearly half the people in the world (actually 44 percent) speak one of the 10 major languages: Standard (Mandarin) Chinese, Spanish, English, Arabic, Hindi, Bengali, Portuguese, Russian, Japanese, and German. The first three of these languages are each spoken by more than a quarter of a billion people. (See "Acknowledgments" at the beginning of the book for the basis of numbers such as these.) But most of the world's languages are spoken by comparatively few people. Fifty-one percent of all languages are spoken by fewer than 10,000 people, the number in a small town in industrial countries. Nearly a quarter of all languages have fewer than 1,000 speakers, about the number of people in a village. It's only due to accidents of history that Chinese, Spanish, and English are so widely spoken. If other circumstances had prevailed, the Eskimo might have become a dominant power, and we might all be speaking a language with only three vowels.

More than half the 7,000 languages are spoken by small tribes in three tropical areas, one in the rain forests of South America, another in the equatorial regions

of Africa, and the third centered on Papua New Guinea. In these areas there is ample rainfall and the people have been relatively self-sufficient for thousands of years. As a result, many of them have lived in small groups for untold generations. Until recently they had no great need to talk to people outside their group. They had the resources to live and talk in their own way. Quite often they developed new sounds, constrained only by the general pressures that affect all human speech.

Now that governments, schools, radio, and even television are spreading into remote regions on every continent, the smaller tribes may want to learn more of the language of their dominant neighbors. Most of the world's population is at least partially bilingual, and a high proportion speaks three or more languages. With the advent of more trade and businesses the smaller groups are becoming part of larger communities and their languages are becoming endangered. Generally a language dies because mothers do not speak it to their children and the language of the home changes. Parents learn a new language that their children are learning in school. Soon the children no longer speak their parents' first language and their mother tongue is lost.

When a language disappears much of the culture often disappears with it. In the face of socioeconomic changes, small groups are forced into a choice that can seldom be fully satisfactory. They have to choose whether to keep at least partially to themselves, maintaining their traditional lifestyle, or to gain the benefits of belonging to a larger group by joining the world around them. They *may* (and it's a big *may*) gain access to schools, health care, and a higher standard of living. But they may lose many aspects of life that they hold dear. However, outsiders should note that the choice between remaining apart and assimilating can be made only by the members of the group themselves. No one except the speakers should decry the loss of a particular language. Others who do so are being paternalistic and assuming they know what is best for other people. Some groups may be better off if they retain their language and culture, which would require some degree of separateness. Others might find it preferable to change, and allow the loss of their mother tongue. As a linguist I am sad that many languages are disappearing so that I will no longer be able to study their wonders. But it is not up to me to decide whether efforts should be made to keep a particular language alive. The speakers may find the costs too great, and the benefits small in comparison with becoming potentially equal members of a larger group.

Studying endangered languages (as I have done for more than 10 years) may reveal previously unreported sounds. It's not that endangered languages are more likely to have unusual sounds. These languages are not endangered because of their sound systems, but because of socioeconomic changes. Many well-known languages have unusual sounds, but that does not make them likely to become endangered. American English has a rare vowel, the one in words such as *her*, *sir*, *fur*. This r-colored sound occurs as a vowel in less than 1 percent of the world's languages, but it won't cause the death of American English. Some endangered

languages have complex sets of sounds and others do not. But they are all a joy to phoneticians because there is always the chance that some of their unstudied sounds are unusual. By investigating little-known languages we get a further glimpse into the range of human phonetic capabilities, and the constraints on possible speech sounds.

1.2 The Evolving Sounds of Languages

The sounds of languages are constrained, first by what we can do with our tongues, lips, and other vocal organs, and second by our limited ability to hear small differences in sounds. These and other constraints have resulted in all languages evolving along similar lines. No language has sounds that are too difficult for native speakers to produce within the stream of speech (although, as we will see, some languages have sounds that would twist English-speaking tongues in knots). Every language has sounds that are sufficiently different from one another to be readily distinguishable by native speakers (although, again, some distinctions may seem too subtle for ears that are unfamiliar with them). These two factors, articulatory ease and auditory distinctiveness, are the principal constraints on how the sounds of languages develop.

There are additional factors that shape the development of languages, notably, from our point of view, how our brains organize and remember sounds. If a language had only one or two vowels and a couple of consonants it could still allow words of half a dozen syllables, and make a vast number of words by combining these syllables in different orders. But many of the words would be very long and difficult to remember. If words are to be kept short and distinct so that they can be easily distinguished and remembered, then the language must have a sufficient number of vowels and consonants to make more than a handful of syllables.

It would be an added burden if we had to make a large number of sounds that were all completely different from one another. It puts less strain on our ability to produce speech if the sounds of our languages can be organized in groups that are articulated in much the same way. We can think of the movements of our tongues and lips as gestures, much like the gestures we make with our hands. When we talk we use specific gestures – controlled movements – to make each sound. We would like to use the same gestures over and over again. This is a principle that we will call gestural economy. Typically, if a language has one sound made by a gesture involving the two lips such as **p** as in *pie*, then it is likely to have others such as **b** and **m**, as in *by* and *my* made with similar lip gestures. If you say *pie*, *by*, *my*, you will find that your lips come together at the beginning of each of them. If a language has *pie*, *by*, *my*, and also a sound made with a gesture of the tongue tip such as **t** in *tie*, then it is also likely to have other sounds made with the tongue tip, such as **d** and **n** in *die* and *nigh*. You can feel that your tongue makes a similar gesture in each of the words *tie*, *die*, *nigh*. The sounds that evolve in a language form a pattern; and there is a strong pressure to fill gaps in the pattern.

Societies weight the importance of the various constraints – articulatory ease, auditory distinctiveness, and gestural economy – in different ways, producing mutually unintelligible languages. But despite the variations that occur, the sounds that all languages use have many features in common. For example, every language uses both vowels and consonants to produce a variety of words. All languages use outgoing lung air in all words (though some may use ingoing air in parts of a word). And all languages use variations in the pitch of the voice in a meaningful way.

1.3 Language and Speech

The main point of a language is to convey information. Nowadays a language can take various forms. It can be spoken or written, or signed for those who cannot hear, represented in Braille for the blind, or sent in Morse code or semaphore or many other forms when necessity arises. Speech is the most common way of using language. But speech is not the same as language. Think of what else you learn just by listening to someone talking. There are all sorts of non-linguistic notions conveyed by speech. You need only a few seconds to know something about a person talking to you, without considering the words they use or their meaning. You can tell whether they come from the same part of the country as you. You know the social group they belong to, and you may or may not approve of them. Someone talking with a so-called Harvard accent may sound pretentious. In Britain the differences between accents may be even more noteworthy. As Bernard Shaw puts it in the Preface to *Pygmalion*: “It is impossible for an Englishman to open his mouth without making some other Englishman despise him.” The accent someone uses conveys information about what sort of person they are, but this is different from the kind of information conveyed by the words of the language itself.

Another aspect of speech that is not part of language is the way speech conveys information about the speaker’s attitude to life, the subject under discussion, and the person being spoken to. We all know people who have a bright, happy way of talking that reflects their personalities – perhaps someone like Aunt Jane, who was always cheerful, even when she was dying of cancer. Of course, people who sound happy may be just putting on a brave front. But, true or false, their speech conveys information that is not necessarily conveyed by their words. Whenever someone talks, you get an impression of their mood. You know whether they are happy, or sad, or angry. You can also assess how they feel about whatever you are discussing. They may sound interested or indifferent when they reply to your comments. In addition you can tell from their tone of voice what they think about you. They may be condescending, or adoring, or just plain friendly. All these attitudinal aspects of speech are wrapped up together in information conveyed by speech. You may be wrong in whatever inferences you make, but, true or false, whenever someone talks, their speech is conveying information of this sort.

The final kind of non-linguistic information conveyed by speech is the identity of the speaker. You can often tell the identity of the person who is speaking without looking at them. Again, you may be wrong, but when someone telephones and simply says ‘Hi’, you may be able to say whether it is a member of your family or a friend you know. You can get this kind of information from the aspects of speech we have just been discussing, the regional accent and the attitude that the speaker has. But there is often something more. You can tell which person it is from that region, and you can say who they are, whatever their current emotions. I know my wife’s voice on the phone, even when I am expecting a call from one of her relatives, and irrespective of whether she is cross because she has just had her purse stolen, or delighted because she has won the lottery. She still sounds like Jenny. (To be truthful, I have never had the opportunity to test the last part of this observation; I don’t really know whether I could identify her voice when she has won the lottery.)

Courts of law have traditionally granted considerable leeway to earwitnesses, even more so than to eyewitnesses. But experience suggests that we may not be quite as skilled at earwitness identification as we think we are. Have you never been stumped when someone telephones and simply says “Hi”? A typical reaction might be to keep the conversation going, guardedly, in hopes that the caller might mention a shared experience that hints at his or her identity, or that enough of the caller’s individual linguistic characteristics will eventually emerge from the speech stream to permit an identification.

1.4 Describing Speech Sounds

In this book we will refer to the sounds of languages in three different ways. We will describe the sound waves in acoustic terms; we will note the gestures of the vocal organs used to produce them (the articulations); and we will label them using the symbols of the International Phonetic Alphabet (the IPA). We will discuss the latter two possibilities, the articulations and symbols, later, although we can note here that sounds illustrating all the IPA symbols are on the website in Recording 1.1. When you listen to the sounds illustrating the IPA symbols, remember that they are the typical sounds represented by these symbols in many different languages. They may not be equivalent to any of the sounds of your accent of English or any other particular language.

Phonetic symbols should not be confused with letters of the alphabet used in spelling words. I always put phonetic symbols in bold, thus: **a**, **e**, **i**, **o**, **u**. Whenever I am referring to letters I put them in quotes: ‘a’, ‘e’, ‘i’, ‘o’, ‘u’. Another convention adopted in this book is to put the written form of the word under discussion in italics. I might, for example, refer to the consonant **p** as in *pie*. Finally, the English translation of a foreign word is also put in quotes, to distinguish it from the foreign word itself. Thus, *padre* ‘father’.



Listen to sound
files online

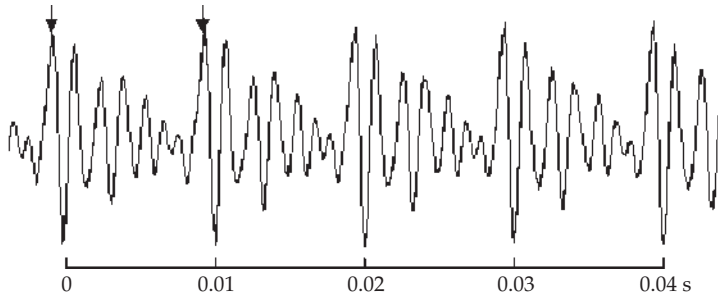


Figure 1.1 Part of the sound wave of the vowel *a* as in *father*. The arrows indicate a section that is repeated every one-hundredth of a second.

We will begin our description of the sounds of languages by considering the sound waves that are produced when we talk. This is a very useful way of quantitatively describing speech. One of my favorite quotations is the statement by the nineteenth-century physicist, Lord Kelvin:

I often say that when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely in your thoughts advanced to the state of Science, whatever the matter may be.

For many years we have been gathering information about the acoustics of speech, based on the precise measurement of sound waves. This constitutes far and away the greatest amount of data on the languages of the world available today. There is also, to be sure, a great deal of ongoing research into the articulation of speech – precise measurements of the positions of the tongue, lips, and other articulators by means of magnetic resonance imaging, ultrasound, and other biomedical imaging techniques. But while these data would satisfy Lord Kelvin, the range of languages for which there is a substantial body of such data is still very limited, compared to audio recordings that allow us to measure the sound waves scientifically.

Whenever you speak you create a disturbance in the air around you, a sound wave, which is a small but rapid variation in air pressure spreading through the air. Figure 1.1 shows part of the sound wave of the vowel *a* as in *father*. During this sound the air pressure at the speaker’s lips goes up and down, and a wave with corresponding fluctuations is generated. When this sound wave reaches a listener’s ear it causes small movements of the eardrum, which are sensed by the brain and interpreted as the sound *a* as in *father*, spoken with a particular pitch and loudness.

We can start thinking about the sound waves that form the acoustic structure of speech by considering the ways in which sounds can differ. Speech sounds such as vowels can differ in pitch, loudness, and quality. You can say the vowel *a* as in *father*

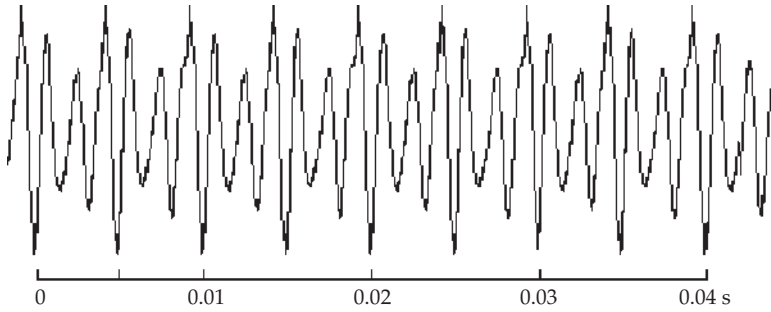


Figure 1.2 Part of the sound wave of the vowel *α* as in *father* said on a pitch corresponding to a frequency of 200 Hz, making it an octave higher than the sound in Figure 1.1.

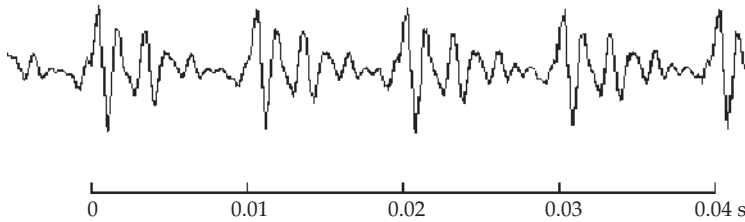


Figure 1.3 Part of the sound wave of the vowel *α* as in *father* produced with approximately half the loudness of the sound in Figure 1.1.

on any pitch within the range of your voice. You can also say it softly or loudly without altering the pitch. And you can say many different vowels, without altering either the pitch or the loudness.

The pitch of a sound depends on the rate of repetition of the sound wave. A short section of the sound wave of the vowel *α* as in *father* in Figure 1.1 repeats itself every one-hundredth of a second. The frequency of repetition is 100 times a second, or, in acoustic terms, 100 Hz. (Hz is the abbreviation for hertz.) This particular frequency corresponds to a fairly low pitch in a male voice. Figure 1.2 shows the same vowel with the higher frequency of 200 Hz, which means that it had a higher pitch. This vowel was said on a pitch an octave above the sound in Figure 1.1. It is in the higher part of the male voice range.

The loudness of a sound depends on the size of the variations in air pressure. Figure 1.3 shows the sound wave of another utterance of the vowel *α* as in *father*. On this occasion the sound wave has an amplitude (the size of the pressure variation) which is about half the amplitude of the vowel in Figure 1.1. You can see that the peaks of air pressure in Figure 1.3 are about half the size of those in Figures 1.1 and 1.2. The only difference between Figures 1.1 and 1.3 is in the amplitude (the size) of the wave. In all other respects the waves have exactly the same shape. Differences in amplitude are measured in decibels (abbreviated dB). The difference in amplitude between the sounds in Figures 1.1 and 1.3 is 6 dB, but, because of the complex way

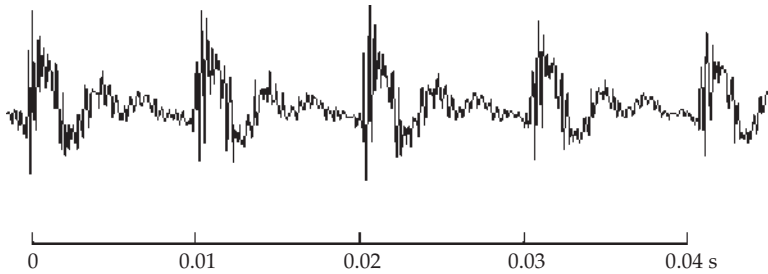


Figure 1.4 Part of the sound wave of the vowel *i* as in *see* said on the same pitch and with approximately the same loudness as the sound in Figure 1.1.

in which dB differences are calculated, there is no easy way of putting an amplitude scale on these figures. The wave in Figure 1.1 sounds a little more than twice as loud as the one in Figure 1.3.

The third way in which sounds can differ is in quality, sometimes called timbre. The vowel in *see* differs in quality from the first vowel in *father*, irrespective of whether it also differs in pitch or loudness. The symbol for the quality of the vowel in *see* is *i*, corresponding to the letter ‘i’ in French, Italian, or Spanish *si*. (In English we rarely use ‘i’ for this sound, but we do in *police*.) As we have seen, the symbol for the quality of the first vowel in *father* is *ɑ*, a script letter ‘a’.

Differences in vowel quality have more complex acoustic correlates, loosely summed up as differences in the shape of the sound wave (as opposed to its repetition rate and size). Figure 1.4 shows the sound wave of the vowel *i* in *see*. This wave, like the waves in Figures 1.1 and 1.3, has a frequency of 100 Hz, in that the wave repeats every hundredth of a second. It also has a slightly smaller amplitude than the *ɑ* in Figure 1.1. The vowel *i* (remember, this is the internationally agreed-upon symbol for the vowel in *see*, not the sound of the word *I*) is usually less loud – has less amplitude – than the vowel *ɑ* because the mouth is less open for *i* than for *ɑ*. In general, the wider you open your mouth, the louder the sound.

The shape of a sound wave is sometimes called the waveform. The waveforms in Figures 1.1 and 1.4 repeat every one-hundredth of a second, so that both sounds have the same pitch. The waveform in Figure 1.4 has a greater number of small variations than that in Figure 1.1. It is the waveform of a sound with a different quality. We will discuss differences in quality in Chapter 4. In the next chapter we will consider how differences in pitch are used in the world’s languages. Then in Chapter 3 we will set the scene for discussing how vowels differ in their acoustic quality.

1.5 Summary

The 7,000 languages in the world today are mainly languages with a small number of speakers. Many of these languages will not be spoken in the near future, usually because they are no longer spoken in the home. The principal constraints on the

evolution of the sounds of the world's languages are ease of articulation, auditory distinctiveness, and gestural economy. The differences between speech and language are in the kinds of information that are conveyed; speech conveys more information about the speaker's background, attitude, and personal identity. The most scientific way of describing speech is in acoustic terms. The main acoustic distinctions among sounds are those corresponding to differences in pitch, loudness, and quality.