The Sounds of Language

An Introduction to Phonetics and Phonology

Elizabeth C. Zsiga

WILEY-BLACKWELL

The Sounds of Language

Linguistics in the World

Linguistics in the World is a textbook series focusing on the study of language in the real world, enriching students' understanding of how language works through a balance of theoretical insights and empirical findings. Presupposing no or only minimal background knowledge, each of these titles is intended to lay the foundation for students' future work, whether in language science, applied linguistics, language teaching, or speech sciences.

What Is Sociolinguistics?, by Gerard van Herk
The Sounds of Language, by Elizabeth Zsiga

Forthcoming

Second Language Acquisition, by Kirsten Hummel
An Introduction to Language, by Kirk Hazen
The Nature of Language, by Gary Libben
Sociolinguistics in Language Teaching, by Gary Barkhuizen

The Sounds of Language

An Introduction to Phonetics and Phonology

Elizabeth C. Zsiga

WILEY-BLACKWELL
A John Wiley & Sons, Ltd., Publication

This edition first published 2013

Blackwell Publishing was acquired by John Wiley & Sons in February 2007. Blackwell's publishing program has been merged with Wiley's global Scientific, Technical, and Medical business to form Wiley-Blackwell.

Registered Office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial Offices
350 Main Street, Malden, MA 02148-5020, USA
9600 Garsington Road, Oxford, OX4 2DQ, UK
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

For details of our global editorial offices, for customer services, and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell.

Library of Congress Cataloging-in-Publication Data
Zsiga, Elizabeth C.
The sounds of language : an introduction to phonetics and phonology / Elizabeth C. Zsiga.
p. cm. – (Linguistics in the world)
Includes bibliographical references and index.
ISBN 978-1-4051-9103-6 (pbk. : alk. paper) – ISBN 978-1-4051-9104-3 (alk. paper)
1. Grammar, Comparative and general–Phonology. 2. Phonetics. I. Title.
P217.Z75 2013
414–dc23
2012026660

A catalogue record for this book is available from the British Library.

Cover image: © Andrejs Pidjass / iStockphoto
Cover design by Nicki Averill

Set in Minion 9.5/11.5 pt by Toppan Best-set Premedia Ltd

1 2013

Contents

Preface

to the student: what this book is about

This book is about the sounds of speech. Communication sometimes takes place without sound, of course: with a smile or a raised fist, a nod or a wave, a photograph or a drawing. There can even be language without sound: those who cannot hear use languages based on gestures instead. Yet for most of us most of the time, getting our message across involves encoding it in sounds. Even when we write, we use symbols that are based on speech (though, in English, sometimes not very directly).

The study of the sounds of speech is often divided into the disciplines of *phonetics* and *phonology*. *Phonetics* studies speech sounds as physical objects. Phoneticians ask questions such as:

- How are speech sounds made?
- How many different sounds do languages use?
- How does sound travel through the air?
- How is it registered by the ears?
- How can we measure speech?

Questions concerning how sounds are made fall under the domain of *articulatory phonetics*; questions concerning how sound propagates and how it can be measured fall under the domain of *acoustic phonetics*.

Phonology studies how languages organize sounds into different patterns. Phonologists ask questions such as:

- How do languages organize sounds to distinguish different words?
- What sorts of restrictions, or constraints, do languages put on sequences of sounds?
- What sorts of changes (alternations) do sounds undergo if illicit sequences arise?
- How are sounds organized into larger constituents (syllables, stress feet, words, phrases)?

Patterns relating to linear strings of segments are the concern of *segmental phonology*; patterns relating to larger hierarchically-organized constituents are the concern of *suprasegmental* phonology.

In addition, phonologists and phoneticians may study how sound patterns arise, change, and vary, asking questions such as:

- How do languages change over time?
- Why are there different dialects?
- How do children learn to speak?
- Why is it hard to learn a second language as an adult?

This book begins to address each of these questions, providing a basic introduction to what the linguistic study of sounds is about. While the book follows the traditional divisions of phonology and phonetics, it is hoped that this overview will help students to see the sub-disciplines in relation to each other, identifying areas of overlap and mutual concern.

to the teacher: how to use this book

This book is designed as an introduction to the linguistic study of speech sounds, that is, the disciplines of phonetics and phonology, in either undergraduate or beginning graduate classes. The book is divided into five sections, each of which deals with one of the traditionally-recognized divisions of the discipline: Articulatory phonetics, Acoustic phonetics, Segmental phonology, Suprasegmental phonology, and Variation and change. As noted above, while the divisions are traditional, the approach seeks to present the linguistic study of sound patterns as a unified endeavor, with different sub-disciplines approaching the object of study from different points of view, asking different kinds of questions about the same subject. Phonetics and phonology inform one another, but are not conflated.

Though the topics and presentation are theoretically grounded, it is not a goal of the book to present the details of particular theories. Rather, the goal is to present: (1) the range of data that any theory must account for; (2) important concepts and constructs that emerge from the data; and (3) some critical overviews of different approaches that have been taken to tackling the issues, with opportunities for the students to practice data analysis and hypothesis testing. Many examples will be from English (as the one language accessible to all readers), and several sections are devoted to the detailed description of English phenomena (as this is of particular interest to many), but there is also broad coverage of typologically diverse systems.

The book follows a modular design, allowing maximum freedom for the instructor. There is enough material for a two-semester course, covering both phonetics and phonology. It will be an advantage that the two subjects are covered from a consistent point of view with consistent terminology, avoiding the gaps and contradictions often encountered when different texts are used.

The book can also be used, however, in a one-semester course that covers both. The first two chapters of each section present a generally non-technical, data-oriented overview, with emphasis on English, raising topics which the following chapters treat in more detail or more formally. Chapter 2, for example, describes the sounds of English, while Chapters 3 and 4 delve into the full IPA. Chapter 11 surveys the types of alternations common in the languages of the world, with a rich set of examples, and Chapters 12–14 discuss more formal theoretical approaches to accounting for them. A one-semester class covering both phonetics and phonology might be created using the first two chapters of each section, with the instructor choosing other chapters according to the topics deemed most important. The final section of the book discusses historical change, first and second language acquisition, and sociolinguistic variation, topics that are often of great interest to students of speech, but are not often covered in introductory texts.

Finally, fresh exercises are included at the end of every chapter.

ad majorem dei gloriam

1 The Vocal Tract

In all things of nature there is something of the marvelous.

Aristotle, *Parts of Animals*

Parts is parts.

Wendy's commercial

Chapter outline

The Sounds of Language: An Introduction to Phonetics and Phonology, First Edition. Elizabeth C. Zsiga.

We begin our study of the sounds of speech by surveying the parts of the body used to make speech sounds: *the vocal tract*. An understanding of how these parts fit and act together, the topic of Section 1.2, is crucial for everything that comes later in the book. Before we dive into the study of human anatomy, however, Section 1.1 considers some of the tools that speech scientists have used or currently use to do their work: How can we "see" inside the body to know what our vocal tracts are doing?

1.1 seeing the vocal tract: tools for speech research

The vocal tract is comprised of all the parts of the body that are used in the creation of speech sounds, from the abdominal muscles that contract to push air out of the lungs, to the lips and nostrils from which the sound emerges. We sometimes call this collection of parts "the organs of speech," but there really is no such thing. Every body part that is used for speech has some other biological function – the lungs for breathing, the tongue and teeth for eating, the larynx to close off the lungs and keep the two systems separate – and is only secondarily adapted for speech.

We're not sure at what point in time the human vocal tract developed its present form, making speech as we know it possible; some scientists estimate it may have been 50,000 to 100,000 years ago. And we don't know which came first, the development of a complex brain that enables linguistic encoding, or the development of the vocal structures to realize the code in sound. While hominid fossils provide some clues about brain size and head shape, neither brains nor tongues are well preserved in the fossil record. We do know that no other animal has the biological structure needed to make the full range of human speech sounds. Even apes and chimpanzees, whose anatomy is generally similar to ours, have jaws and skulls of very different shape from ours, and could only manage one or two consonants and vowels. That's why scientists who try to teach language to chimps or apes use manual sign language instead: chimps are much better at manipulating their fingers than their mouths. There are birds that are excellent mimics of human speech sounds, but their "talking" is really a complex whistling, bearing little resemblance to the way that humans create speech. (The exact mechanism used by these birds is discussed in Chapter 9.)

But probably for as long as people have been talking, people have been interested in describing how speech sounds get made. Linguistic descriptions are found among the oldest records of civilization. In Ancient India, as early as 500 BCE, scribes (the most famous of whom was known as Pāṇini) were making careful notes of the exact articulatory configurations required to pronounce the Vedic Scriptures, and creating detailed anatomical descriptions and rules for Sanskrit pronunciation and grammar. (The younger generation, apparently, was getting the pronunciation all wrong.) Arab phoneticians, working several centuries later but with many of the same motivations as the Indian Grammarians, produced extensive descriptions of Classical Arabic. Al-Khalil, working in Basra around 100 CE, produced a 4,000-page dictionary entitled *Kitab al ʻayn*, "The Book of the Letter ʻAyn." The ancient Greeks and Romans seemed to be more interested in syntax and logic than in phonetics or phonology, but they also conducted anatomical experiments, engaging in an ongoing debate over the origin of speech in the body. Zeno the Stoic argued that speech must come from the heart, which he understood to be the source of reason and emotion, while Aristotle deduced the sound-producing function of the larynx. The Greek physician Galen seems to have settled the argument in favor of the Aristotelian view by noting that pigs stop squealing when their throats are cut. Medieval European linguists continued in the Greek tradition, further developing Greek ideas on logic and grammar, as well as continuing to study anatomy through dissection.

The main obstacle in studying speech is that the object of study is for the most part invisible. Absent modern tools, studies of speech production had to be based on either introspection or dissection. (According to one history of speech science, it didn't occur to anyone until 1854 that one could use mirrors to view the living larynx in motion.) Of course some speech movements are visible, especially the lips and jaw and sometimes the front of the tongue, so that "lip reading" is possible, though difficult. But most of speech cannot be seen: the movement of the tongue in the throat, the opening of the passage between nose and mouth, sound waves as they travel through the air, the vibration of the fluid in the inner ear. The experimental techniques of modern speech science almost all involve ways of making these invisible movements visible, and thus measurable.

One obvious way is to take the pieces out, at autopsy. Dissection studies have been done since antiquity, and much important information has been gained this way, such as our knowledge of where muscles and cartilages are located, and how they attach to each other. But the dead patient doesn't speak. Autopsy can tell us about the *anatomy* of the vocal tract, that is, the shape and structure of its parts, but it cannot tell us much about *physiology*, that is, the way the parts work together to produce a specific sound or sound sequence.

The discovery of the X-ray in 1895 was a major advance in speech science, enabling researchers to "see" inside the body. (The mysterious "X" ray was discovered by physicist Wilheim Conrad Röntgen, who received the first Nobel Prize in physics for his work.) X-rays are not necessarily great tools for visualizing the organs of speech, however, for two reasons. X-rays work because they pass through less dense, water-based soft tissue like skin, but are absorbed by denser materials like bone, teeth, and lead. Thus, if an X-ray is passed through the body, the bones cast a white "shadow" on a photographic film placed behind the subject. The first problem with the use of X-rays in speech science is that muscles, like the tongue, are more like skin than like bones. The tongue is visible on an X-ray, but only as a faint cloud, not a definite sharp outline.

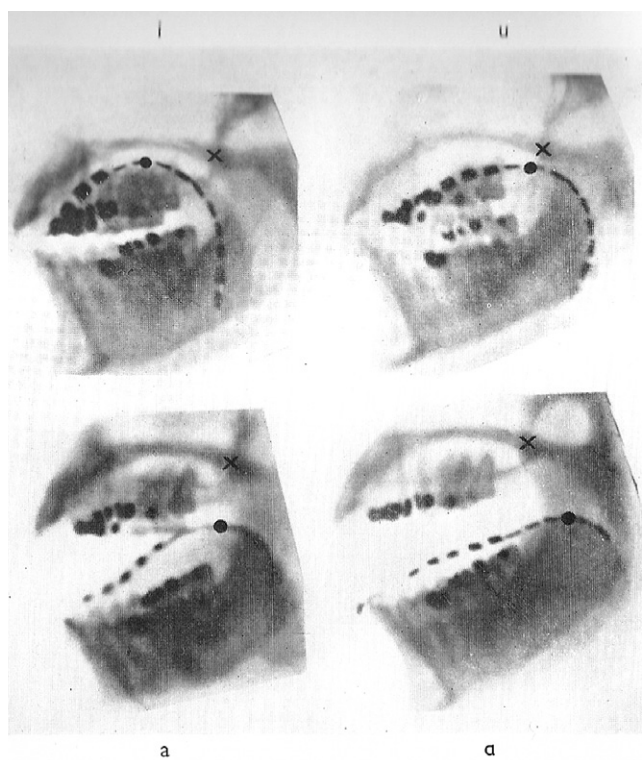

Figure 1.1 X-rays from the lab of Daniel Jones. Source: Published in Jones (1966) *The Pronunciation of English.*

Figure 1.1 shows the results of one experiment where researchers tried to get around this problem in an ingenious way. These images were made by the British phonetician Daniel Jones, in 1917. Jones was very interested in vowel sounds, and is famous (among other achievements) for devising a system for describing the sound of any vowel in any language (the "cardinal vowel" system, which is discussed in Chapter 4). But nobody really knew exactly how the tongue moved to make these different sounds, since we cannot see most of the tongue, and we do not have a very good sense of even where our own tongues are as we speak.

To create these images, Jones swallowed one end of a small lead chain, holding on to the other end, so that the chain lay across his tongue. He then allowed himself to be X-rayed while articulating different vowel sounds, and these pictures are the result. The images, beginning at the upper left and going clockwise, show vowels similar to those in the words "heed," "who'd," "hod," and "had." ("Hod" rhymes with "rod", and was a common word in 1917. It refers to a bucket or shovel for carrying coal.) The tongue itself does not image well, but the lead chain shows up beautifully, indicating how the tongue is higher and more towards the front of the mouth for the vowel in "heed" than the vowel in "hod." (The large dot showing the high point of the tongue was drawn in by hand, and the other black dots are lead fillings in the subject's teeth.)

The second problem with X-ray technology, of course, is that we eventually learned that absorbing X-rays into your bones, not to mention swallowing lead, is dangerous for the subject. Prof. Jones' experiments would never make it past the review committees that every university now has in place to protect subjects' health.

A safe way to get pictures of parts of the vocal tract is through *sonography*. This technology, based on the reflection of sound waves, was developed in World War II, to allow ships to "see" submarines under the water. Most of us are familiar with this technology as it is used to create images of a fetus *in utero*. The technology works because sound waves pass harmlessly and easily through materials of different kinds, but bounce back when they hit a surface of different density from what they are traveling through. (Transmission of sound waves is covered in detail in Chapter 6.) So sound waves travel through the air, but bounce back when they hit a mountainside, creating an echo. They travel through the water, but bounce back when they hit the ocean bottom (or a submarine), creating a sonar image. They travel through the amniotic fluid, but bounce back when they hit a body part. A transducer receives the echoed signal and calculates the time delay between transmission and reception. The time measurement is converted into distance between the transducer and the reflecting object. Graphing these distances creates an outline, resulting in an image of the object being studied.

Figure 1.2 Ultrasound image of tongue shape for a vowel similar to the one in "heed." Image courtesy of the author.

In speech science, the sonar probe is held under the chin, so that the sound waves travel up through the tongue. They bounce back when they hit the border between the tongue and the air in the mouth, creating an image of tongue shape. Such an image is seen in Figure 1.2: The shape of the tongue for the vowel similar to the one in "heed" is seen as a bright line. (In this image, the subject is facing to the right.)

Figure 1.3 MRI mid-sagittal section of the tongue, showing the same vowel as in Figure 1.2. Source: Courtesy of Maureen Stone, Vocal Tract Visualization Laboratory, University of Maryland. http://speech.umaryland.edu/MICSR.html.

Finally, Figures 1.3 and 1.4 show images of the head made by *magnetic resonance imaging* (MRI). The phenomenon of magnetic resonance was discovered in 1946 by Felix Bloch and Edward Purcell (working independently). (As with X-rays, the discovery resulted in a Nobel Prize.) Applications for imaging the body began to be used around 1977. This technology is in some ways the reverse of an X-ray: because it is based on the response of hydrogen atoms to a magnetic field, it works best on soft tissue that is mostly H_2O. For MRI imaging, the subject is placed in a very large electro-magnet. If you've ever had an MRI scan, you know this is like being slid into a small plastic tunnel. When the magnet is turned on, all the hydrogen atoms in the watery parts of the body line up with the direction of the magnetic field. Then a radiowave pulse of energy is sent to a specific part of the body, and this pulse knocks the atoms out of their alignment. When the pulse passes, the atoms snap back into position, but they give off energy as they do so. This energy is detected and measured by the MRI technology. Depending on the density of the hydrogen atoms in a given material, more or less energy is given off, and thus the presence and shape of different kinds of tissue can be detected. The technology is excellent, for example, at differentiating a tumor from normal brain tissue. It also creates beautiful, crisp pictures of the inside of a subject's head. It's especially useful because it can image a "slice" at any depth into the body. Image acquisition is rather slow in speech terms, however: while MRI "movies" are possible, at this writing, the technology is mostly used for still pictures.

The image in Figure 1.3, a view that's often used in speech science, is called a *mid-sagittal section* – that is, cutting halfway through, and looking from the side. It's basically a profile, but cut down the middle of the head. You can clearly see the nose and chin in outline, as well as the spinal cord. Our interest will be in the structures in the mouth. The mid-sagittal section is especially useful because the different parts of the vocal tract are clearly outlined. Over-reliance on this way of picturing the vocal tract, however, can give a false impression that our mouths are two-dimensional and that the only differences that matter are front-to-back. Figure 1.4 shows a different MRI view, called a *coronal section*: this is a view from the front, again slicing the head in half, this time from ear to ear. The coronal section in Figure 1.4 shows the whole head: the brain and eyes are unmistakable. The dark open spaces are the sinus cavities. The solid gray structure under the nose is the tongue. Notice how the top of the tongue is high and arched from side to side, while the muscle mass fills the floor of the mouth and extends down to the jawbone. (While the source does not indicate what, if any, sound was being pronounced when this picture was taken, the high position of the tongue looks consistent with that of the previous figures.)

While the coronal section reminds us not to be trapped in two dimensions, the mid-sagittal section is the most useful view to begin our discussion of the names and locations of the different parts of the vocal tract.

Figure 1.4 MRI coronal section of the head. Source: Image by Patrick J. Lynch, medical illustrator; C. Carl Jaffe, MD, cardiologist. http://en.wikipedia.org/wiki/File:Head_mri_coronal_section.jpg. Creative Commons.

1.2 the parts of the vocal tract

As stated above, the vocal tract comprises all the structures of the body that are used to create speech sounds, from the lungs to the nose. It is useful, however, to divide the overall structure into a number of sub-systems, which are diagrammed in Figure 1.5. (This view, which is not actually biologically possible, might be called "Egyptian": The head is in profile, but the shoulders are square to the reader.)

The dividing point of the subsystems is the larynx, a valve of cartilage and muscle in the throat that sits on top of the trachea (windpipe) at the point where the passage to the lungs separates from the passage to the stomach, and that can be felt in the front of the neck as the "Adam's apple." The parts are the *sub-glottal* or *sub-laryngeal system* (everything below the larynx), the *laryngeal system* (the larynx itself), and the *supra-laryngeal system* (everything above the larynx). The supra-laryngeal system can be further divided into the *oral tract* (the mouth) and the *nasal tract* (the nose). In this section, we'll consider each of these systems in turn, describing each of the structures and the role they play in speech production, starting at the bottom, with the lungs, and working our way out. In this chapter, the perspective is functional: The structures and systems are described in just enough detail that the job that they play in making speech sounds can be understood. Further details of anatomy and physiology are covered in Chapter 5.

1.2.1 the sub-laryngeal vocal tract

Breathing is generally pretty quiet, snoring aside. The act of speaking is the act of making the movement of air out of the mouth and nose audible, using a code that associates different

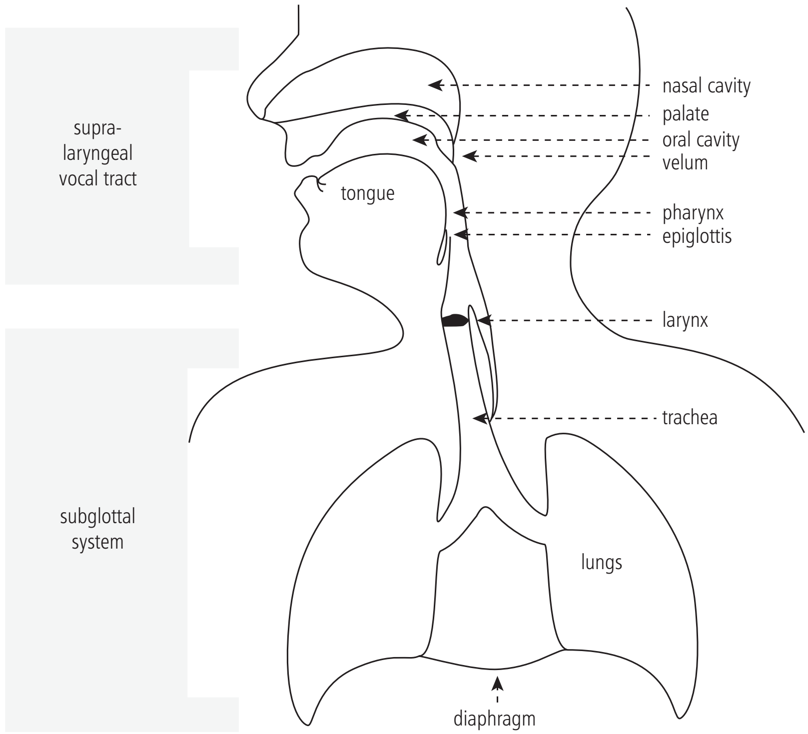

Figure 1.5 The sub-systems of the vocal tract. Source: Philip Rubin and Eric Vatikiotis-Bateson, Haskins Laboratories. http://www.haskins.yale.edu/featured/heads/MMSP/figure1.html.

sound combinations with different meanings. The movement of air can be made audible by making it fast and turbulent, like water rushing through rapids. It can be made audible by building up pressure that is suddenly released, like the pop of a champagne cork. Or air can be made audible by making the air molecules vibrate (or *resonate*), like a bell or like the air inside the body of a guitar or clarinet. Each of these methods, alone and in combination, are used in making speech sounds. And all involve first getting the air moving, and then shaping that movement either by making a narrow channel, by closing off the airflow momentarily so that pressure builds up, or by changing the shape of the resonating chamber.

In most cases, the movement of the air is powered by the lungs: the *pulmonic airstream mechanism*. It is possible to get air moving in other ways, by movements of the tongue and larynx, and many languages use speech sounds created in these ways. These other airstream mechanisms are covered in Chapter 3. But all languages use the pulmonic airstream mechanism, some (like English) exclusively, so we begin with that.

Speech begins when air is drawn into the lungs, and then forced out. The lungs, of course, are air-filled sponges, covered with a membrane, that fill the ribcage. The lungs have no muscles of their own, but expand or contract as the muscles of the abdomen and ribcage squeeze or pull on them.

The largest muscle used in speech is the *diaphragm*, which is a large dome-shaped muscle that runs through the middle of the body and separates the chest cavity from the stomach. When the diaphragm contracts, the dome flattens out, causing the chest cavity, and thus the lungs, to enlarge. Remember Boyle's law from high-school physics? When the volume of an enclosed space increases, air pressure decreases. So when the volume of the lungs is increased the air pressure goes down, and air rushes in from outside the body to equalize the difference. Then, as the diaphragm is slowly relaxed, and the muscles running inside the ribcage slowly contract, the lungs are squeezed, and air is forced out in a measured stream, up the trachea, through the larynx, and into the mouth and nose.

Note that, generally, speech takes place on the exhalation only. That is, the airstream is pulmonic *egressive*. It is possible to speak on the indrawn breath – in gasps, for example, or when children want to count quickly to 100 for a game of hide and seek – but no language uses the pulmonic *ingressive* airstream in any systematic way.

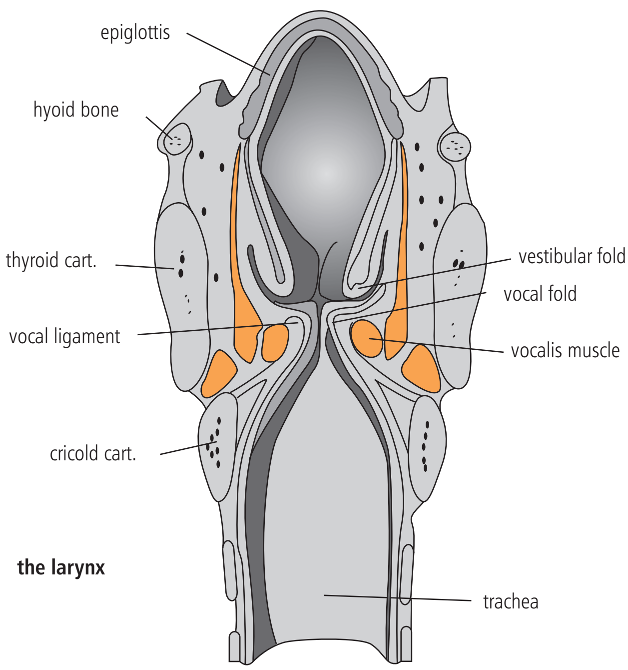

Figure 1.6 The vocal folds, coronal section.

1.2.2 the larynx

The *trachea*, or windpipe, is the tube that connects the mouth and lungs. It consists of rings of cartilage (horseshoes, really, open to the back) connected by muscle fibers and lined with smooth moist tissue. The top ring of the trachea, the *cricoid cartilage*, is thickened, and instead of being open at the back is closed with a flat plate: the shape is often compared to a signet ring, with the plate of the ring toward the back of the neck. (*Cricos* is Greek for ring.) Just above the cricoid cartilage, the tissue inside the trachea thickens into two folds or flaps: the *vocal folds*, which can potentially cover the opening to the trachea. (They were erroneously called the vocal cords by a Medieval anatomist, and the term stuck, but they are not strings.) The folds are actually composed of multiple layers of tissue: they are thin and somewhat stiff on their inside edge, at the *vocal ligament*, thick and muscular at the sides. The muscle running inside the vocal folds is called the *vocalis muscle*. Figure 1.6 shows a diagram of the vocal folds and surrounding structures (in coronal section). Note that above the true vocal folds is another set of flaps of tissue, the *vestibular folds*, often called the *false vocal folds*. Vibration of the false vocal chords may play a role in various forms of "throat singing," but they are not active for speech.

A set of cartilages surrounds, protects, and manipulates the vocal folds. Figure 1.7 shows the cartilages of the larynx as viewed from the side; other views are shown in Chapter 5 (Figures 5.2 and 5.3).

In Figure 1.7, one can see the thickened cricoid cartilage sitting on top of the trachea, with its signet-like plate in the back. The large *thyroid* cartilage sits over the front of the cricoid like a shield, folded in the middle. On top of the cricoid plate, hidden behind the thyroid and thus not visible in Figure 1.7, are the two *arytenoid cartilages*, shaped approximately like triangular pyramids. The vocal folds are attached to the thyroid cartilage in the front and the arytenoid cartilages in the back. These four cartilages – the cricoid, the thyroid, and the two arytenoids – along with the vocal folds and the muscles that connect them all make up the *larynx*.

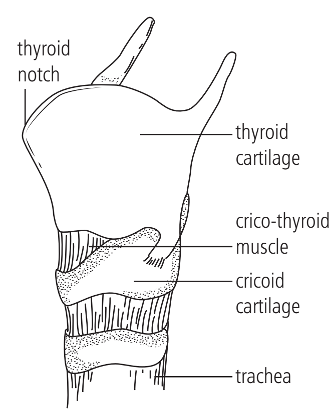

Figure 1.7 The cartilages of the larynx, viewed from the side.

You can locate your own larynx by tilting your chin up and running your finger down the underside of your chin to your throat. You should be able to feel the forward point of the thyroid cartilage as a bump under the skin, with a small notch at the top. (This thyroid notch is usually called the "Adam's apple," and it is larger in men than in women; more on sex-specific characteristics of the larynx in Chapter 5.) The whole larynx will move up and down if you swallow. You should also be able to feel some of the other tracheal rings below the larynx along the front of your throat.

During speech, the positions of the vocal folds can be changed in various ways by adjusting the position of the thryoid and arytenoid cartilages, and by increasing or decreasing the tension in the vocalis and other muscles. (Details are given in Chapter 5.) The vocal folds can be pulled together tightly and clamped down to temporarily stop the airflow, or they can be pulled apart to allow air to flow freely. If they are held in an intermediate position with just the right tension, touching but not tightly closed, the vocal folds will vibrate as the air passes over them, with a movement similar to a flag undulating in the breeze. This vibration produces a hum called *voicing*, and the vibration of voicing is the basis of many of the sounds of speech.

You can feel this vibration by placing your fingers on the side of your larynx, and alternately pronouncing a sustained [zzzzzz] and [ssssss]. The vibration you feel during the *voiced* [z] will cease during *voiceless* [s]. (Conventionally, symbols for sounds are written within square brackets. There will be much more information on symbols for sounds in the coming chapters.)

The faster the vibration of the vocal folds, the higher the pitch of the voice. To some extent, the pitch of your voice is determined biologically: large vocal folds vibrate more slowly than small vocal folds, producing a lower range of sound. Men, on average, have vocal folds 50% longer than women's (hence the more prominent Adam's apple); male vocal folds are also thicker. Thus, men tend to have lower voices. The average rate of vocal fold vibration for adult males is about 120 times per second, for adult females, about 220 times per second, and for small children, as high as 300 to 400 times per second. Part of this difference is due to the fact that men on average have larger bodies than women, and both of course are larger than children. But males will tend to have larger vocal folds even when matched with females for overall body size. In adolescence, young males undergo a hormone-driven laryngeal growth spurt, during which the size of the vocal folds and thyroid cartilage rapidly increase. They may find that their voice "cracks" until they become used to manipulating their newly-found deeper voices. The growth of female vocal folds is steady and proportional to body size, with no disproportional growth spurt in adolescence. There is presumably a selective advantage for males, but not females, to sound bigger (and thus stronger and more fearsome) than they actually are.

So the overall range of your voice – whether you'll be a soprano, alto, tenor, or bass – is a biological given. (Modulo human intervention, of course. In seventeenth- and eighteenth-century Italy, there was a tradition of male *castrati* undergoing surgery to maintain their operatic high voices throughout life.) But within the biologically-given range, the actual note that is sung is under individual speaker control. Pitch is adjusted by changing the tension of the vocal folds. Pulling forward on the thyroid cartilage stretches the folds and raises the pitch. Contraction of the vocalis and other muscles can also be directly controlled. You can feel the laryngeal control of pitch if you once again place your finger lightly on your larynx and then hum a series of notes from the lowest in your range to the highest. Can you feel the thyroid cartilage move up in your throat as pitch rises? Such speaker-controlled changes in pitch can be very important to the linguistic message.

Thus, the linguistic function of the larynx is (mainly) to control pitch and voice. Its biological, non-speech, function is to serve as a valve that separates the lungs and the stomach. Other animals that use their mouths to both breathe and eat have valves that serve the same non-speech function as the human larynx, but no other animal has such fine-tuned laryngeal control.

The human larynx is also nearly unique in being placed so low in the throat. In adult humans, the trachea (the passage to the lungs) and esophagus (the passage to the stomach) do not split off until halfway down the throat. In addition, the trachea is in front of the esophagus, so that food and water must pass over the top of the trachea in order to get to the stomach – a surprisingly dangerous arrangement. (Your mother was right – don't try to talk and eat at the same time.) In other primates and almost all other mammals the trachea extends much higher, up through the back of the mouth so that the trachea can connect directly to the nasal passages.

Thus a deer can drink and breathe at the same time: the air going through the nose and down the trachea, with water going around the sides of the trachea rather than over the top. Humans cannot.

In order to be able to eat at all without getting food into the lungs, the human must close off the trachea by tightly closing the larynx when food or drink is being swallowed. The closure is aided by the *epiglottis*, a flap of tissue attached to the base of the tongue, which folds down over the top of the larynx, which rises to meet it, during swallowing. Try swallowing a few times, paying attention to the muscular sensations. Can you feel the larynx rising up, and the tension in the throat that corresponds to the lowering of the epiglottis?

1.1 In Focus

Human infants, who have very poor muscle control and spend a lot of time drinking while lying down, are born with the larynx high in the throat, similar to the arrangement typical of other primates. Thus, the liquid the baby ingests passes around the sides of the trachea rather than over the top, preventing choking. The larynx lowers to the normal human position over the first months of life, as the child gains control over the muscles of the neck and head, learns to sit up, and begins to eat solid food and to babble.

The tradeoff for the lowered larynx position in the human is significant, however. With the trachea down and out of the way, humans have an open space, the *pharynx*, at the back of the mouth behind the tongue. This open space allows greater freedom for movement of the tongue, making a wide range of vowel and consonant sounds, and thus human language, possible.

1.2.3 the supra-laryngeal vocal tract

Thus far, we've seen that the lungs provide the moving air on which speech is based, and the larynx adds (or not) the vibration of voicing and control of pitch. It is the structures above the larynx that move to further shape and constrict the air as it moves out from the lungs, creating distinctions between individual speech sounds. It is useful to divide the structures of the mouth into the *active articulators* and *passive articulators*. The active articulators move toward the passive articulators in order to constrict and shape the airstream. The labels for the active and passive articulators are shown in Figure 1.8. In this chapter, we concentrate just on the names, relationships, and basic functions of the

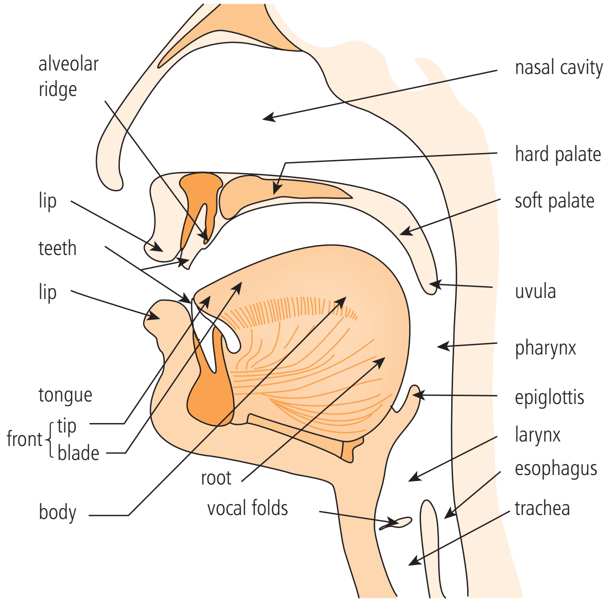

Figure 1.8 The supralaryngeal vocal tract.

different parts. Chapter 2 begins the discussion of how each is involved in creating specific distinct speech sounds.

The active articulators are the lips and tongue. The movement of the lips, being visible, is obvious. In speech, the lips may be closed or open, pursed or spread. Both lips move to some degree, but the lower lip moves more, so it receives the active label.

All that we usually see of the tongue is the small, pink tip, but it is in fact a large mass of interconnected muscles that fills the floor of the mouth. The tongue is all muscle: it has no bones or cartilage, thus no definite shape, though it maintains a constant volume. The technical term for this kind of organ is a *muscular hydrostat*: the closest biological analogs are an elephant's trunk or the tentacle of an octopus. The various muscles of the tongue are discussed in detail in Chapter 5. At this point it suffices to know that three different regions of the tongue can move relatively independently: the *tongue front*, the *tongue body*, and the *tongue root*. The tongue front is made up of the very *tip* of the tongue and the *tongue blade*, which extends a few centimeters behind the tip. The tongue front can be raised or lowered, stuck forward or curled back. The tongue body, or *dorsum*, is the main mass of the tongue. It can move up and down, back and forward. The tongue root is the very back part of the tongue, extending down into the pharynx to the epiglottis. The tongue root can be pulled forward or back, enlarging or constricting the pharynx.

Each of these four active articulators – the lower lip, tongue front, tongue body, and tongue root – can move to create a constriction, that is, a narrowing of the vocal tract, against one or more of the passive articulators that lie along the top of the vocal tract.

1.2 In Focus

The jaw, of course, and the lower teeth embedded in the jaw, also move during speech, but the jaw is not counted as a separate active articulator because the jaw and lower teeth are never the parts that actually make the constriction. It is always the lips or tongue riding on the movement of the jaw: no language makes any speech sounds by grinding the teeth together.

The easily visible passive articulators are the *upper lip* and *upper teeth*. Right behind the upper teeth is the *alveolar ridge*: you can feel this raised ridge with your tongue. If you continue to run your tongue along the top of your mouth, you'll feel the *post-alveolar region* arching from the alveolar ridge toward the *hard palate*, the very roof of the mouth. If you curl your tongue very far back in your mouth, you can feel that the bony structure of the hard palate gives way to softer tissue, which is known as the *soft palate*, or *velum*. The velum is a muscular flap, like a trap door, that regulates the *velar port*, the opening in the back of the mouth that connects the mouth and nose. At the very end of the velum is the *uvula*, the little pink pendulum you can see hanging down in the back of your mouth when you open