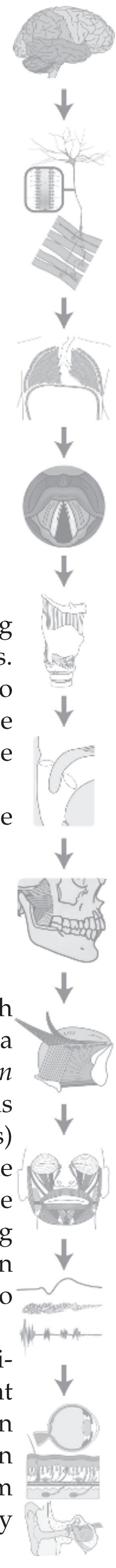
Chapter 1

The Speech System and Basic Anatomy



Sound is movement. You can see or feel an object even if it – and everything around it – is perfectly still, but you can only hear an object when it moves. When things move, those movements can be heard, seen, or felt. In order to understand the sounds of speech (the central goal of phonetics as a whole), we must understand how the different parts of the human body move to produce those sounds. This is the central goal of articulatory phonetics.

This chapter describes the roadmap we follow in this book, as well as some of the background basics you'll need to know.

1.1 The Speech Chain

Many accounts have described the process of producing and perceiving speech in terms of a *feed-forward* system, leading in one direction from a speaker to a listener. This kind of system can be represented as a *feed-forward speech chain* (e.g., Denes and Pinson, 1993), as shown in Figure 1.1. A feed-forward system is one in which a movement plan (in this case, a sequence of speech movements) is constructed and carried out without any monitoring or correction along the way. It's easy to identify a feed-forward system, as it is one in which all the arrows would go in one direction in a diagram (as in Figure 1.1), never leading you back to their starting point. Because the arrows never meet up, they form an incomplete loop like the letter "U," which is why a feed-forward system is also sometimes referred to as an "open-loop" system.

In a simple feed-forward auditory *speech chain* model, a speaker's communicative intentions are converted into a set of *motor commands* (brain signals sent to muscles) that produce vocal tract movements – *articulations* – resulting in acoustic speech output. A listener can then pick up this acoustic information through hearing, or *audition*, in which air pressure changes are conveyed from the ear via sensory nerves to the brain, where they are perceived and finally interpreted as conveying meaning.

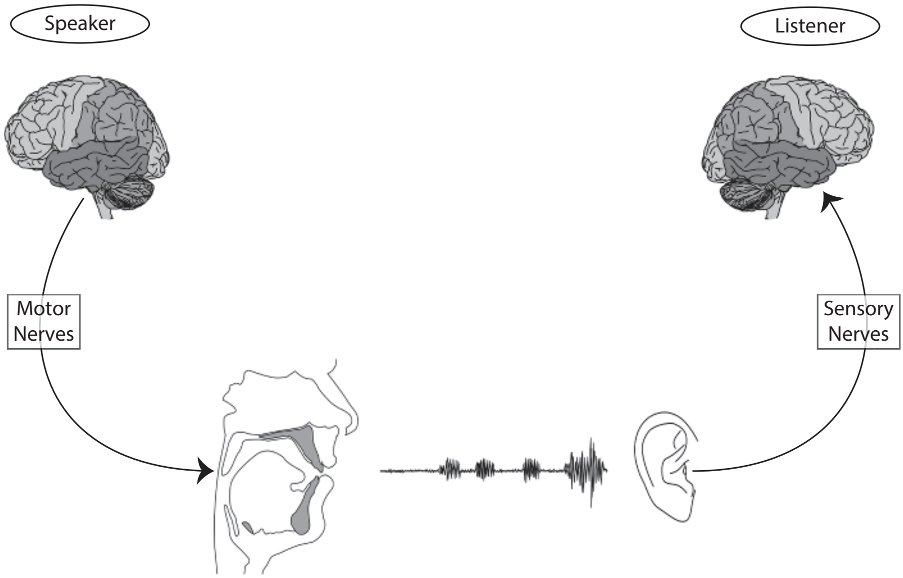


Figure 1.1 Feed-forward, auditory-only speech chain (image by W. Murphey and A. Yeung).

While feed-forward control and audition are certainly important in speech, we know that speech is not always unidirectional, and it involves more than just sound. Rather, when we speak, we are also constantly using sensory information to monitor and adjust what we're doing as we move along the chain. This constant flow of sensory information is called *feedback*. The arrows in a feedback system connect back to form a complete loop, like the letter "O," so that a feedback system is also sometimes called a "closed-loop" system (sensory feedback will be discussed in more detail in Chapter 12).

While both feed-forward and feedback control is needed for speech to work, speech movements are ultimately produced for communication. This means that our loops of information need to include feedback we get from our environment, including our interlocutors, indicating whether they understood what we said or whether they responded in the way we had expected. This kind of external (ecological) feedback is much less consistent than the more direct internal (sensorimotor) feedback we get from the feeling or sound of our own movements, but it is no less important for training the system to make movements (see Ting *et al.*, 2024, Gick and Mayer, in preparation).

It is also important to bear in mind that the sensory information generated and exchanged during speech is *multimodal*. That is, we use not just our sense of hearing when we perceive speech, but all of our sense modalities – even some you may not have heard of before. Thus, while information in the speech chain as a whole generally flows forward, each link in the chain functions through a constant flow of internal and external feedback information. As you can see from the schematic shown in Figure 1.2, this version of the "speech chain" is much more complex.

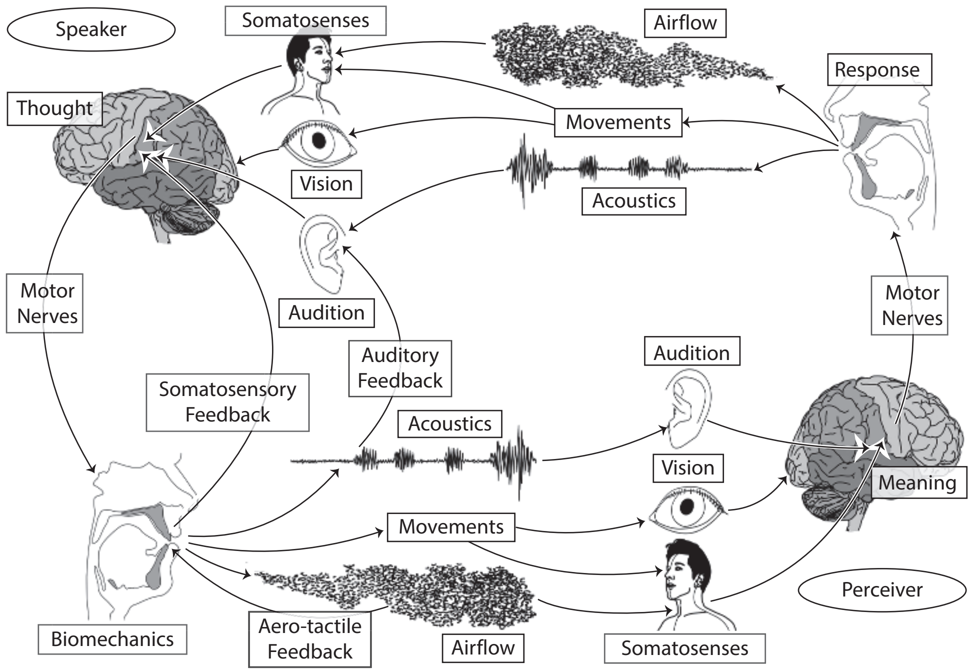


Figure 1.2 Schematic illustration of a multimodal speech chain with internal (sensorimotor) and external (ecological) feedback loops (image by W. Murphey, A. Yeung and D. Derrick).

Sweet Sensations

Speech communication depends on information from many different sensory modalities. Because of this, speech is often talked about in terms of multimodal or *multisensory* perception.

The sensory information people most commonly associate with speech is sound, which is normally conveyed through the *auditory* modality. Auditory information (i.e., sound) does indeed provide a useful medium for sharing information: it can travel a long distance through and around objects, it can be conveyed in the dark, it can provide additional information about the physical characteristics of the speaker, and so on. Auditory feedback provides a speaker with a constant flow of information about their own speech movements. While sound is indeed an important modality for speech, it is by no means the only one we use.

Speech can also be perceived through the *visual* modality, by seeing movements of the face and body. While the visual modality is important for perceiving others' speech, it is of less use for providing speech feedback from one's own articulators, as one normally has relatively little direct visual evidence of one's own speech (though there are exceptions,

(*Continued*)

such as speaking in front of a mirror or seeing steam from one's breath on a cold day).

Speech is also perceived through *somatosensory* modalities. The term *somatosenses* (from Greek, *soma-* "body" + Latin, *sens* "feeling"), sometimes referred to as the "body senses," is a broad cover term for a dizzying range of different sense modalities, some of which have barely been studied, and many of which are important for speech (see, e.g., Keough *et al*., 2019). The *tactile* (touch) senses are a subset of the somatosenses, often described as relevant in perceiving one's own or others' speech. For example, the *vibro-tactile*, *aero-tactile,* and *pressure* senses (relating to vibrations, airflow, and movements, respectively) each provide speakers with different information about their own and sometimes others' speech movements. Other somatosenses can only be detected inside one's own body, such as *proprioception* (also known as *kinesthetic sense*), which conveys information about the position and movements of one's own body parts. Taken together, the tactile senses plus proprioception are sometimes referred to under the term *haptic* (Greek, "grasp").

Chapter 12 will explore sensory information and its use in controlling, monitoring, and correcting speech articulations in more detail.

1.1.1 A speech production roadmap

Because this textbook is about articulatory phonetics, we'll focus mainly on the internal loop of the speech chain, up to where speech sounds leave the mouth. This part of the chain has sometimes been called the *speech production chain* (see Figure 1.3). This book will use a roadmap that follows along the path, starting with the brain and moving in turn through the processes involved in producing different speech sounds. Chapter 12 will "close the loop" with the sensory systems that guide and correct speech articulations.

This roadmap follows the order in which we often think about speech movements being produced in a forward model: our brain initiates a plan, sometimes called a *movement plan* or *motor plan*, which generates motor commands, with nerve impulses distributed to relevant regions of the body. These nerve impulses reach muscles, causing them to contract. Muscle movements expand and contract our lungs, initiating the movement of air. This air progresses through the glottis and into the upper vocal tract, which we shape with additional muscle movements. By changing the configuration of structures in the larynx and upper vocal tract, we can initiate vibration or turbulence, obstruct or release airflow, change frequencies or resonances, and so on, all of which enable us to produce different speech sounds. The sound, air, vibrations, and movements we produce through these actions can then be sensed and perceived by ourselves (sensorimotor feedback) or by other people, our interlocutors.

While it may seem surprising at first for a book about speech articulation to end with a chapter about sensory systems (Chapter 12), it is important to bear

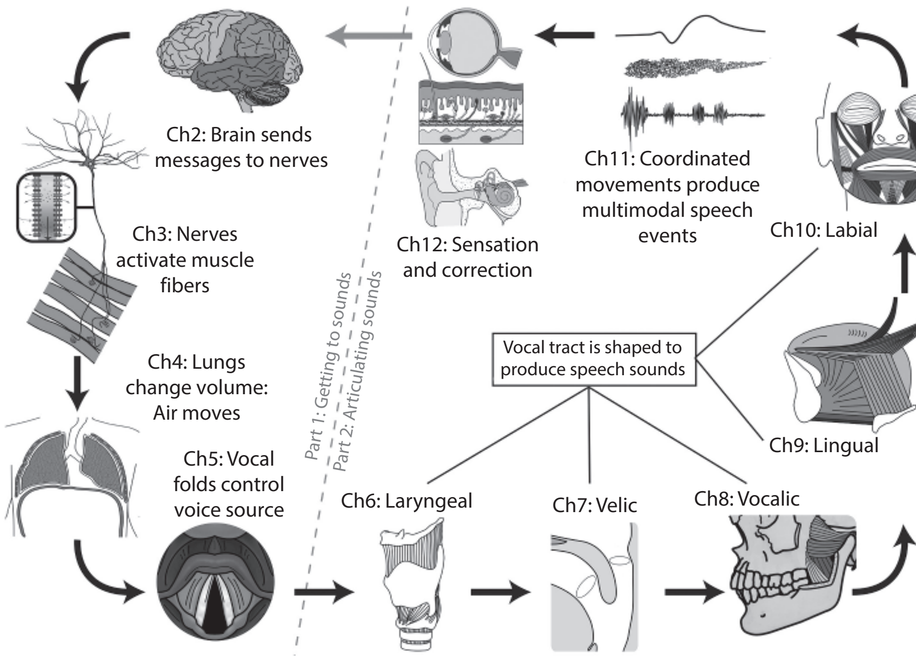


Figure 1.3 An articulatory phonetics roadmap, following the speech production chain. The first half (left) takes you through Part I of the book, and the second half (right) covers Part II (image by D. Derrick and W. Murphey).

in mind that no part of the speech communication system functions in isolation. As shown in Figure 1.2, speech production and perception are richly intertwined. Sensory information is essential in understanding how speech movements are controlled and corrected through sensorimotor feedback. For questions specifically related to acoustic signals and auditory information as inputs to speech perception, we refer the reader to Keith Johnson's *Acoustic and Auditory Phonetics* book (Johnson, 2011).

1.2 The Building Blocks of Articulatory Phonetics

Because the field of articulatory phonetics is concerned with the physical movements we make when we speak, part of learning about articulatory phonetics involves learning a good deal of anatomy. Figure 1.4 shows an overview of some of the important anatomical structures of speech production. The speech production chain begins with the brain and other parts of the nervous system and continues with the respiratory system, composed of the ribcage, lungs, trachea, and all the supporting muscles. Above the trachea is the larynx with its tiny but important cartilages and muscles and the pharynx, which is divided into laryngeal, oral, and nasal sections. The upper vocal tract includes the nasal and oral passages, including the flexible muscular structures of the mouth, such as the tongue and soft palate. The oral passage opens to the teeth and lips as well as the face, an important part of the body both for speech articulation and for the visual and tactile communication of speech.

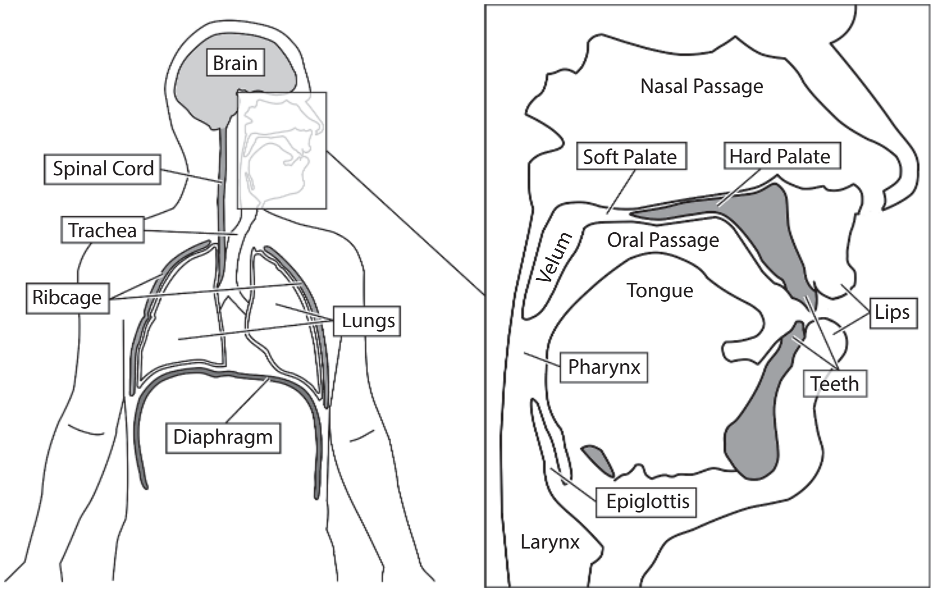


Figure 1.4 Anatomy overview: upper body (left), vocal tract (right) (image by D. Derrick).

Scientists use many terms to describe anatomical structures, and anatomy diagrams often represent anatomical information along two-dimensional (2D) slices or planes (see Figure 1.5). A *midsagittal* plane divides a body or part of the body down the middle into two halves: *dextrad* (Latin, "rightward") and *sinistrad* (Latin, "leftward"). The *sagittal* plane was so named by the ancient Romans, who apparently thought the suture running along the midline of the skull resembled an arrow (Latin *sagitta*, "arrow"). The two axes of the sagittal plane are (1) vertical (running superior–inferior or up–down) and (2) *anteroposterior* (running anterior–posterior or front–back). The midsagittal plane, shown, for example, in Figures 1.4 (right) and 1.5 (right), is by far the most commonly used cross-section in articulatory phonetics. Structures near this midline plane are called *medial* or *mesial*, while structures farther away from the midline are called *lateral*.

Coronal slices divide a body or parts of a body into *anterior* (front) and *posterior* (back) parts. The two axes of the *coronal* plane are (1) vertical and (2) lateral, or side-to-side.

The *transverse* plane is horizontal, dividing a body or body structure into *superior* (top) and *inferior* (bottom) parts.

The direction of the top of the spine is *cranial* or *cephalad* (Greek, "toward the head"), while the direction of the tail is *caudal* (Latin, "tail"). Also, *ventral* (Latin, "belly") refers to the belly, while *dorsal* (Latin, "back") refers to the back, relative to the spine. For creatures that have a horizontal spine, the ventral direction points inferiorly to the belly area underneath, while dorsal points superiorly to the back area on top (as a shark's dorsal fin). For creatures like humans with a vertical spine, on the other hand, ventral is equivalent to anterior, while dorsal is equivalent to posterior for most of the body (the brain being the notable exception). To enable upright posture, the human

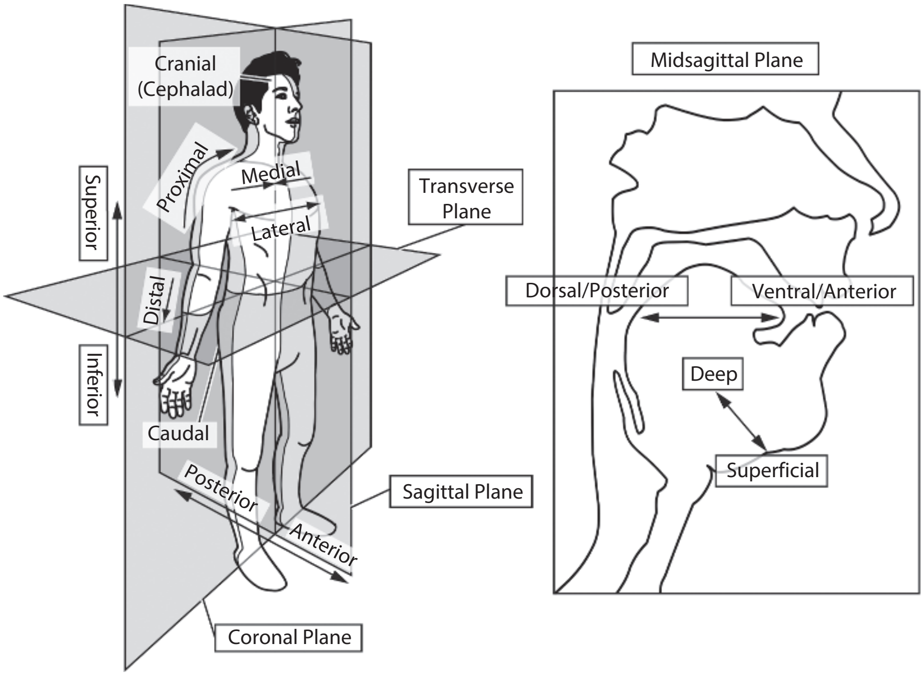


Figure 1.5 Anatomical planes and spatial relationships: full body (left), vocal tract (right) (image by D. Derrick).

brain's frontal lobe is positioned above the eyes, giving us a high, elevated forehead (and confusing terms for brain anatomy, which we'll revisit in Chapter 2).

There are also terms that refer to locations relative to a center point rather than planes. Areas closer to the center of the trunk are called *proximal*, while areas away from the center of the trunk, such as hands and feet, are *distal*. Finally, structures in the body can also be described in terms of depth relative to the skin surface, with *superficial* structures being closer to the surface and *deep* structures being farther away from the surface.

1.2.1 Materials in the body

Anatomical structures are made up of many different materials. The nervous system, not surprisingly, is made up of different kinds of nerves, as discussed in detail in Chapters 2 and 3 and as relevant throughout the book. Chapters 4–11 focus more on bones, cartilages, and muscles, while Chapter 12 discusses sensory systems.

The "hard parts" of the body consist of bones and cartilages. Bony, or *osseous*, material is the hardest. The skull, ribs, and vertebrae are all composed of *bone*. These bones form the support structure of the speech apparatus. *Cartilaginous* or *chondral* (Greek, "cartilage, grain") material is composed of semi-flexible material called *cartilage*, which makes up the stiff but flexible parts you can feel in your ears and nose. The ribcage and larynx contain many cartilages that are important for speech. Bones and cartilages are also the "hard parts" in the sense that you need to memorize their names, whereas most muscles are just named according to which hard parts they are attached to. For these reasons, we usually learn about the hard parts first and then proceed to the muscles.

Muscles are made up of long strings of cells that have the specialized ability to contract (we'll look in more detail at how muscles work in Chapter 3). The word "muscle" comes from the Latin *musculus*, meaning "little mouse" (when you look at your biceps muscle, it's not hard to imagine it's a little mouse moving under your skin!). In this textbook, we'll study only *striated muscles*, also known as skeletal muscles. Striated muscles are often named by combining their *origin*, which is the larger, unmoving structure (usually a bone) to which they attach, and their *insertion*, which is usually the part that moves most when a muscle contracts. Many muscle names take the form "origin-insertion." For example, the muscle that originates from the *thyroid* cartilage in the neck and inserts into the smaller *hyoid* bone in the base of the tongue is called – you guessed it! – the "thyrohyoid" muscle (this is a real muscle you'll learn about later!). As you can see, if you know the origin and insertion of a muscle, then in many cases, you can guess its name.

Muscles seldom act alone. Most of the time, they interact in agonist–antagonist pairs. The *agonist* produces the main movement of an articulation, while the *antagonist* pulls in the opposite direction, which helps to control the stronger pull of the agonist. Other muscles may also act as *synergists*. Activating a synergist lends stability to the system by preventing other unwanted motion.

1.3 Understanding Movement: Control and Biomechanics

Speech articulation is movement. As such, it's helpful to have some terminology for talking about how movement systems work. A key term that has already come up in this chapter is "control." When we're talking about movement systems, the word *control* refers to all of the processes by which the brain and body work together to plan, initiate, regulate, and adapt movements toward achieving goals (such as making speech sounds). We've already introduced several terms and concepts in this chapter that are part of movement control, including motor plans, motor commands, feed-forward, sensorimotor feedback, and correction. While movement control is usually thought of as happening mostly in the brain, the nervous system is distributed and structured throughout the body, making the brain and body indivisible partners in the movement relationship.

All intentional movements in the body are initiated by activating muscles. When muscles activate, they change in shape and stiffness, creating forces that act on surrounding structures. For our purposes in this book, this means we'll often be talking (nontechnically) about *biomechanics*, or how mechanical forces affect body structures. The field of biomechanics is concerned with any aspect of how physics can affect bodies, including how muscles contract (discussed mainly

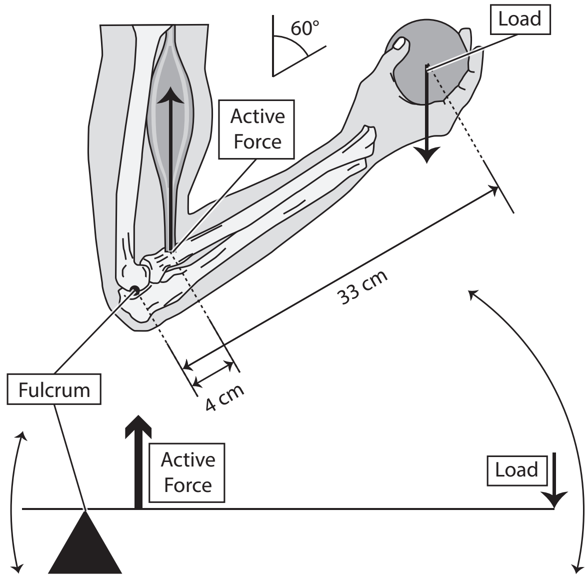


Figure 1.6 Mechanical advantage, type 3 lever (Adapted from Ling *et al.*, 2016, CC: 4.0).

in Chapter 3) or how the size, shape, position, or function of a body structure can be affected by mechanical forces (discussed mainly in Chapters 4–11).

We'll find principles of biomechanics at work every time movement happens in a body. For example, depending on whether a muscle is attached closer to or farther from a joint, it can have a higher or lower *mechanical advantage* (see Figure 1.6). A muscle attached farther from a joint has a higher mechanical advantage, giving the muscle greater strength but less speed and a smaller range of motion. A muscle attached closer to a joint has a lower mechanical advantage, reducing power but increasing speed and range of motion. Other principles, such as *volume conservation,* apply at every level, from cells (e.g., Gelbart *et al.*, 2012; discussed in Chapter 3) to tongues (discussed in Chapter 8).

One area where biomechanics has been important in understanding movement systems relates to the *degrees of freedom* (DoF) problem. When a body part is able to move along a single path (such as sliding along or around a single axis), it is said to have a single degree of freedom. For example, a sliding door has one DoF (side-to-side), as does a hinged door (rotation around the hinge). The more DoFs an object has, the more complex movements it can perform. But as they say, freedom isn't free: managing more DoFs means a higher workload for the brain.

The Soviet pioneer of movement science, Nikolai Bernstein (1967) famously pointed out that a key challenge for any movement system is to reduce DoFs, thereby reducing cognitive load on the system. In principle, for example, a flexible structure like a tongue could have unlimited DoFs, making it impossible to control (i.e., if there are infinitely many ways to bend the tongue to perform a task, how do you choose one?). Researchers continue to debate how the brain and body work together to manage so many DoFs.

Tongues, Tentacles, and Degrees of Freedom

Fortunately, nature provides us with many examples of strategies creatures use to reduce DoFs. Two examples of this are how babies reach for an object and how octopuses use their arms. If you've ever seen a baby reach for an object, you may have noticed large, sweeping movements hinging at the shoulder rather than the more complex, adult-like movements integrating and coordinating all the available DoFs in the human arm, wrist, and hand. This is because infants simplify their reaching tasks by locking joints, such as their wrists and elbows, thereby reducing DoFs (Spencer and Thelen, 2000; Berthier and Keen, 2006).

Similarly, an octopus could, in theory, move its arm in an unlimited number of ways to reach for food and get it into its mouth by curving or bending the arm at any point along its length. In reality, however, Sumbre *et al*. (2006) observe that octopuses tend to use highly consistent movements: they create three "pseudo" joints (proximal, medial, and distal) in each arm, forming an artificially articulated structure that is divided into segments that are strikingly similar to those of a human arm (proximal joint = shoulder; medial joint = elbow; distal joint = wrist). By using these three pseudo-joints, the octopus' motor task of putting food into its mouth is greatly simplified (Figure 1.7).

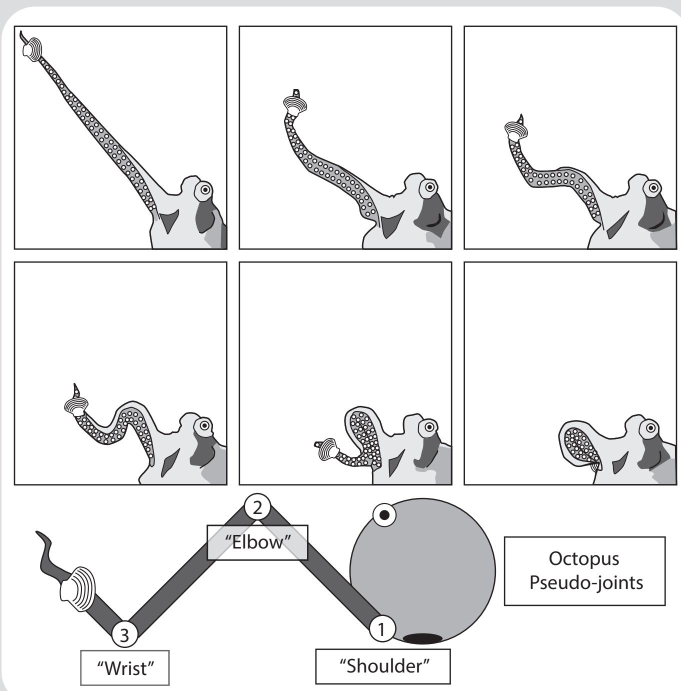


Figure 1.7 Octopus catching and eating a scallop (image by D. Derrick, using concepts from Figure 1, Sumbre *et al*., 2006).

The human tongue appears to present a similar DoF problem to octopus arms, as a tongue can in principle bend anywhere. Chapter 3 will look more closely at how complex body structures, such as human tongues, may lick the DoF problem.

One of the most powerful tools we have for studying biomechanics is computer simulation. This is especially true in the vocal tract, where the complexity of body structures exceeds researchers' capacity for intuitive understanding and where muscles are often so tiny and so intertwined that it can be all but impossible to understand what exactly a muscle does using conventional means.

Biomechanical simulation using open tools such as OpenSim (Delp *et al*., 2007) and ArtiSynth (e.g., Lloyd *et al*., 2012; Fels and Gick, 2016; www.artisynth.org) or the G-OBIM tongue model (Calka *et al*., 2025) can give us insights into the biomechanics of articulatory phonetics that would otherwise be unavailable.

Simulating the Speaking Vocal Tract

Biomechanical simulations of the vocal tract have allowed researchers to isolate variables, such as the effects of gravity, individual muscle activations, or changing tissue composition with age, in ways that real-world studies just can't get at. A few examples of the questions that vocal tract simulations have shed light on include whether gravity affects how astronauts speak (it does! Shamei *et al*., 2023); whether Neanderthal vocal tracts would have enabled them to speak (they would have! Dediu *et al*., 2021); and whether men's mumbling is because of the way their jaws are built (it is! Weirich *et al*., 2016).

Simulating the vocal tract has allowed researchers to see "under the hood" of speech. This has made simulation a valuable complement to the many tools that an articulatory phonetician uses to reveal the secrets of the speaking vocal tract.

1.3.1 The tools of articulatory phonetics

There is no single tool that can measure everything we need to know about articulatory phonetics. Rather, scientists use a wide array of tools to track and measure the movements of speech articulators. Some of these tools are shown in Figures 1.8a and 1.8b, including examples of data obtained using each tool. Each tool has important advantages and disadvantages, including speed, accuracy, participant comfort, availability and setup time, data storage and

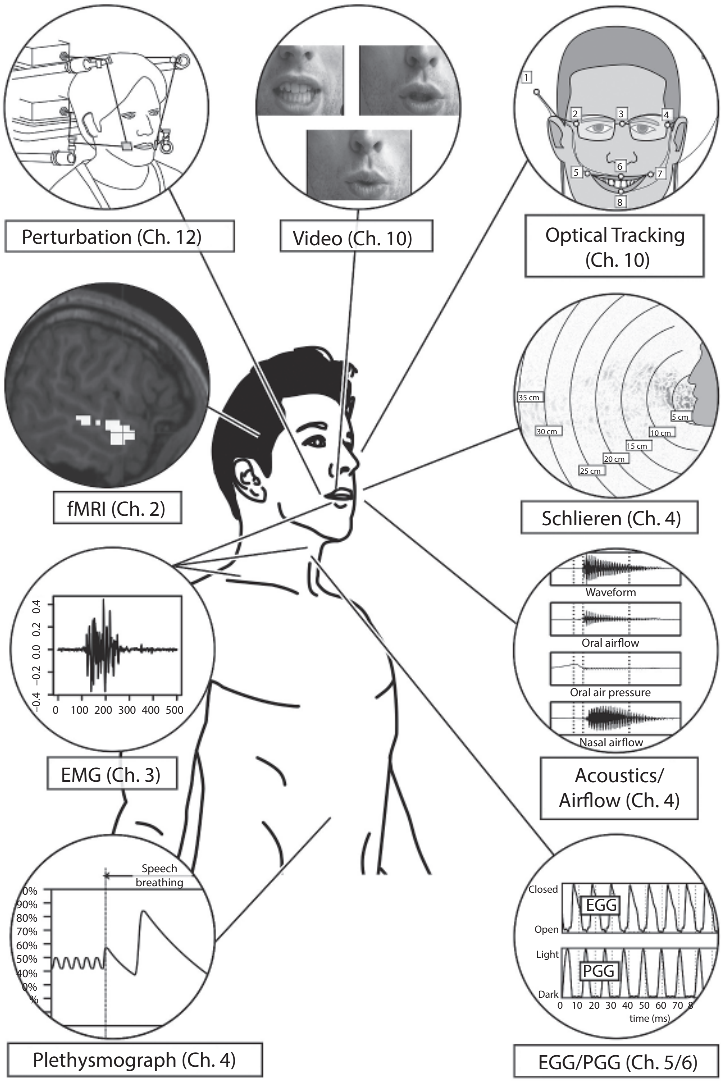


Figure 1.8a Measurement tools for articulatory phonetics (image by D. Derrick).

analysis, expense, and the type of data collected (e.g., electrical activity in nerves or muscles, airflow or pressure, position and movement of body structures, etc.). Issues relating to each tool are discussed in depth in their respective chapters.

Decades ago, when digital computer games first came out, the objects on the screen made choppy movements and looked blocky. Their movements were choppy because it took a long time for early computers to redraw images on the screen, indicating poor temporal resolution. *Temporal resolution* is a term for how often an event happens, such as how often a recording or sample is taken. Particularly when talking about digital signals, the term "temporal

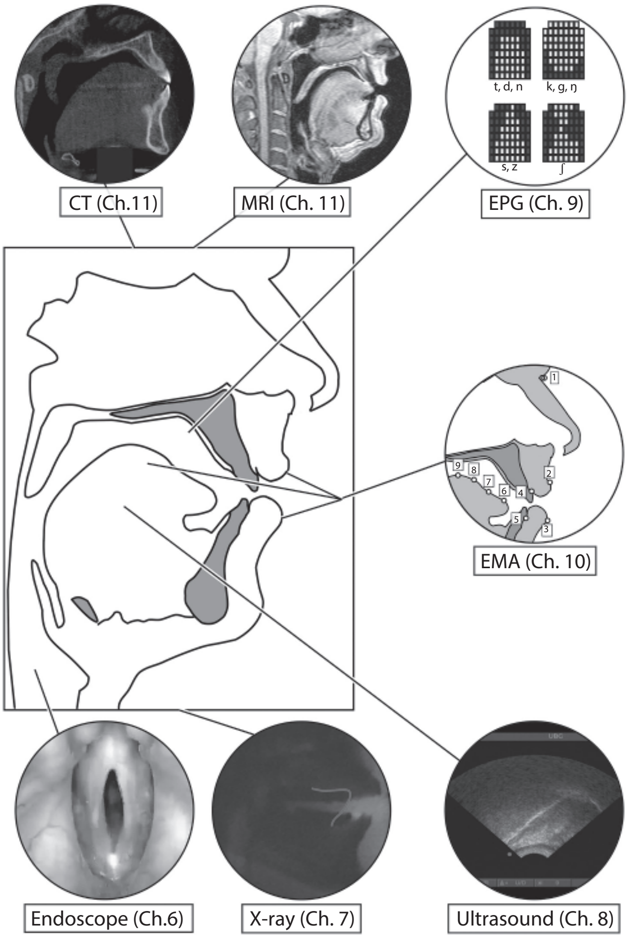


Figure 1.8b Measurement tools for articulatory phonetics (image by D. Derrick).

resolution" is often interchangeable with "sampling rate" and is often measured in samples per second, or *Hertz* (Hz). The "blocky" look was because only a few square pixels were used to represent an object, indicating low spatial resolution. From a measurement point of view, *spatial resolution* can be thought of as a term for how accurately you can identify or represent a specific location in space. Temporal and spatial resolution both draw on a computer's memory, as one involves recording and storing more detail in time, while the other requires recording and storing more detail in space. Temporal and spatial resolution are illustrated in Figure 1.9. Because of this, temporal and spatial resolution normally trade off, such that when one increases, the other often decreases.

Many of the tools we'll look at in this book are ones that researchers use to measure anatomy relatively directly: imaging devices and point-tracking devices. *Imaging* devices take "pictures" of body structures, so that they tend to

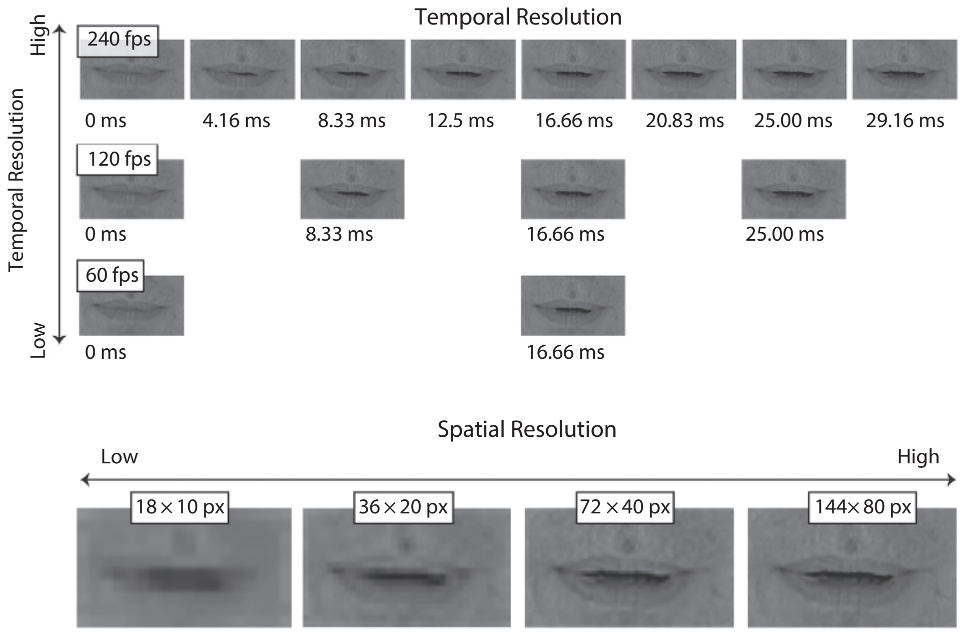


Figure 1.9 Temporal and spatial resolution (image by D. Derrick).

show more spatial detail, but often run more slowly than other kinds of devices. For example, using current ultrasound or X-ray technology, a 2D view of the tongue can be captured up to about 100 times per second, but a more common rate is about 30 times per second, as in standard movie-quality video (relatively low temporal resolution). These tools allow the measurement of slower articulator motion but show a picture of an entire section of the vocal tract (high spatial resolution).

Imaging tools we'll look at in this book include functional magnetic resonance imaging (fMRI), which images blood flow and oxygen in the brain (Chapter 2); positron emission tomography (PET), which images the brain's metabolic activity (Chapter 2); electroencephalography (EEG), which images the brain's electrical activity (Chapter 2); schlieren imaging, which images fluid (air) flow (Chapter 4); endoscopy, which films the vocal folds (Chapter 6); X-ray film, which was historically used to image the inside of the body, especially bone (Chapter 7); ultrasound, which images soft tissues in the body (Chapter 8); electropalatography (EPG), which images tongue contact with the palate (Chapter 9); video, for imaging face and body motion (Chapter 10); computed tomography (CT), which uses X-ray and computation to image cross-sections of the body (Chapter 11); and magnetic resonance imaging (MRI) and real-time magnetic resonance imaging (rtMRI), which images soft tissue in detail (Chapter 11). While most of these produce 2D images, some of them, such as CT, MRI, and some ultrasound machines, provide many slices of 2D information that can be combined to create 3D shapes.

Unlike imaging devices, *point-tracking* systems track the motion of a small number of fixed points or markers very accurately, giving them high spatial resolution for only those points, and they can often capture this information 1000 or more times per second, giving them high temporal resolution as well. Active, passive, and markerless face tracking are point-tracking tools that track points outside the body, whereas X-ray microbeam and electromagnetic articulometers (EMAs) are point-tracking systems that can track points both outside and inside the body (Chapter 10).

We'll also look at other measurement devices in this book. These typically have higher temporal resolution but may not directly convey any spatial information at all. These include electromyography (EMG), which measures electrical output from muscles (Chapter 3); audio, which records sound (Chapter 4); airflow meters, which measure airflow and pressure at a set location (Chapter 4); plethysmography, which tracks body volume (Chapter 4); electroglottography (EGG), which tracks electricity passing through the vocal folds (Chapter 5); and photoglottography (PGG), which tracks light passing through the vocal folds (Chapter 6). These measurement tools can capture information many thousands of times per second. For instance, a compact disc quality sound file contains audio information captured at 44,100 Hz.

Finally, Chapter 12 describes tools that allow us to perturb the sensory feedback loops we use to control, monitor, and correct our speech articulations.

Now that we have some basic background and an idea of where we're headed, the next chapter will start with the central nervous system, looking at the first steps of how the brain initiates movement.

Exercises

Sufficient jargon

Define the following terms: feed-forward, internal feedback loop, external feedback loop, sensorimotor, ecological, speech chain, articulations, audition, feedback, multimodal, multisensory, auditory, visual, somatosensory, somatosenses, tactile, vibrotactile, aero-tactile, pressure, proprioception, kinesthetic sense, haptic, speech production chain, movement plan, motor plan, midsagittal, dextrad, sinistrad, sagittal, anteroposterior, medial, mesial, lateral, anterior, posterior, coronal, transverse, superior, inferior, cranial, cephalad, caudal, ventral, dorsal, proximal, distal, superficial, deep, osseous, bone, cartilaginous, chondral, cartilage, muscles, striated muscles, origin, insertion, thyroid, hyoid, agonist, antagonist, synergists, control, biomechanics, mechanical advantage, volume conservation, degrees of freedom (DoF), temporal resolution, Hertz, spatial resolution, imaging, point tracking.

Short-answer questions

1 Which anatomical plane(s) can intersect both eyes in the same plane?
2 When perceiving speech, one can use vision as well as audition. List 10 English consonants that normally provide clearly visible cues use the International Phonetic Alphabet [IPA] where possible; where the IPA character is not available on a standard keyboard, standard English digraphs and/or broad phonetic descriptions are acceptable; for example, some English vowels could be listed as follows: a, u, lax u, open o, schwa, caret.
3 Give brief examples of feedback that arise during the production of the sound /v/ from each of the (a) auditory, (b) vibrotactile, (c) aero-tactile, and (d) visual domains.
4 The *sternohyoid* muscle connects the *sternum* (a bone in the chest) and the *hyoid* (a bone in the neck).
   A Which of these bones is the origin?
   B Which is the insertion?
   C What do you think is the name of the muscle that originates from the sternum and inserts into the *thyroid* (a cartilage in the neck)?
   D If the *palatoglossus* muscle runs between the palate and the tongue, what muscle inserts into the tongue and originates from the hyoid bone?

Practical assignment: biomechanical simulation

Part of learning to do articulatory phonetics research is learning how to read articles and uncover research techniques from those articles. Many of those techniques – especially those relying on specialized computer software – may be outdated, difficult to implement, or no longer operational on modern computers. For this chapter, we provide a short assignment geared to help you think through some of these issues, with a focus on biomechanical simulation of speech.

1 Find and cite a paper not cited in this chapter that talks about biomechanical simulation of speech; follow the citation formatting (APA style) used in the references section at the end of each chapter of this book. The paper can refer to simulation tools mentioned in this chapter or to other tools not mentioned here. Make sure the paper uses a simulation tool that has a website or download link.
2 Read the website to determine whether the simulation tool is still regularly updated; report on when the last update took place. Provide the website URL and when it was last updated.
3 Write about three possible uses of the tool in speech research, either as described in the paper, described on the website, or that you can figure out on your own based on what you learned in this chapter. Keep the simulation tool in mind as you progress through the textbook and revisit this question once you have read this book from cover to cover.

Bonus: Attempt to download the tool and see if the instructions are good enough to allow you to test the simulation tool. Use the simulation tool to run any demonstrations that are available on the website. Report on whether it works or if it is broken, and if so, report on the outcome(s) or error(s).

References

Bernstein, N. A. (1967). *The Coordination and Regulation of Movements*. London, UK: Pergamon.

Berthier, N. E. and Keen, R. (2006). The development of reaching in infancy. *Experimental Brain Research*, *169*, 507–518.

Calka, M., Badin, P., Nazari, M. A., Rochette, M., Perrier, P., and Payan, Y. (2025). The G-OBIM tongue model: an accurate open-source biomechanical model of a male human tongue. *PLoS Computational Biology*, *21*(9), e1013378, 1–28.

Dediu, D., Moisik, S. R., Baetsen, W. A., Bosman, A. M., and Waters-Rist, A. L. (2021). The vocal tract as a time machine: inferences about past speech and language from the anatomy of the speech organs. *Philosophical Transactions of the Royal Society of London B, Biological Sciences*, *376*(1824), 20200192, 1–11.

Delp, S. L., Anderson, F. C., Arnold, A. S., Loan, P., Habib, A., John, C. T., Guendelman, E., and Thelen, D. G. (2007). OpenSim: open-source software to create and analyze dynamic simulations of movement. *IEEE Transactions on Biomedical Engineering*, *54*(11), 1940–1950.

Denes, P. B. and Pinson, E. N. (1993). *The Speech Chain: The Physics and Biology of Spoken Language* (2nd edition). New York: W. H. Freeman.

Fels, S. and Gick, B. (2016). Interdisciplinary approaches for advancing articulatory speech theory and synthesis. *Canadian Acoustics*, *44*(3), 21–22.

Gelbart, M. A., He, B., Martin, A. C., Thiberge, S. Y., Wieschaus, E. F., and Kaschube, M. (2012). Volume conservation principle involved in cell lengthening and nucleus movement during tissue morphogenesis. *Proceedings of the National Academy of Sciences*, *109*(47), 19298–19303.

Gick, B. and Mayer, C. (In preparation). How bodies talk: Speech as an embodied movement system.

Johnson, K. (2011). *Acoustic and Auditory Phonetics* (3rd edition). Chichester, UK: Wiley-Blackwell.

Keough, M., Derrick, D., Taylor, R. C., Schellenberg, M. H., and Gick, B. (2019). Learning effects in multimodal perception with real and simulated faces. In S. Calhoun, P. Escudero, M. Tabain, and P. Warren (eds), *Proceedings of the 19th International Congress of Phonetic Sciences*. Australia: Melbourne, 1189–1192.

Ling, S. L., Sanny, J., and Moebs, W. (2016). *University physics: Volume 1*. OpenStax https://openstax.org/books/university-physics-volume-1/pages/1-introduction.

Lloyd, J. E., Stavness, I., and Fels, S. (2012). ArtiSynth, A fast interactive biomechanical modeling toolkit combining multibody and finite element simulation. In Y. Payan (ed), *Soft Tissue Biomechanical Modeling for Computer Assisted Surgery*. Springer, 355–394.

Shamei, A., Sóskuthy, M., Stavness, I., and Gick, B. (2023). Postural adaptation to microgravity underlies fine motor impairment in astronauts' speech. *Scientific Reports*, *13*(8231), 1–8.

Spencer, J. P. and Thelen, E. (2000). Spatially-specific changes in infants' muscle coactivity as they learn to reach. *Infancy*, *1*, 275–302.

Sumbre, G., Fiorito, G., Flash, T., and Hochner, B. (2006). Octopuses use a human-like strategy to control precise point-to-point arm movements. *Current Biology*, *16*(8), 767–772.

Ting, L. H., Gick, B., Kesar, T. M., and Xu, J. (2024). Ethnokinesiology: towards a neuromechanical understanding of cultural differences in movement. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *7*, 20230485 https://doi.org/10.1098/rstb.2023.0485.

Weirich, M., Fuchs, S., Simpson, A., Winkler, R., and Perrier, P. (2016). Mumbling: macho or morphology? *Journal of Speech, Language, and Hearing Research*, *59*(6), S1587–S1595.