Cambridge University Press
978-1-108-41834-8 — Phonetics
Ratree Wayland
Excerpt
More Information

1 Speech Articulation

Manner and Place

Learning Objectives

By the end of this chapter, you will be able to:

- Identify systems in the body used for speech production, such as
  - The respiratory system
  - The phonation system
  - The articulatory system: the vocal tract
- Describe speech: consonants
- Describe speech: consonants with multiple articulations
- Describe speech: vowels

Introduction

Phoneticians and phonologists are linguists specializing in the study of speech sounds. Phoneticians study their physical properties, whereas phonologists are interested in the structure, organization, and patterns – the more abstract aspects of speech sounds. The division of labor between them is similar to that of an architect (a phonologist) and a builder (a phonetician). The architect is interested in the overall look of the building and its layout, such as the locations of rooms, doors, and windows, as well as the relationship between spaces within the building, whereas a builder focuses on the details of how to construct each part of the building, including the types and sizes of the materials to be used. The analogy is not perfect, but I hope you get the point.

Physical properties of speech sounds can be described in terms of how they are made in the human vocal tract, and this falls under the domain of **articulatory phonetics**. **Acoustic phonetics** focuses on how speech sound propagates or travels, and how its acoustic properties can be measured. Questions on how speech signals are registered and processed by the human auditory system, and which acoustic properties are used to differentiate classes of speech sounds, belong to the domain of auditory phonetics or **speech perception**.

Phoneticians ask questions such as:

- How are speech sounds made?
- What are the speech sounds attested in the world's languages?
- How can they be written down?
- What are their acoustic properties and how can they be measured?
- How are they processed in the ears and perceived or interpreted by the brain?

Phonologists, on the other hand, ask big-picture questions such as:

- How do languages organize speech sounds to differentiate words?
- How are they organized into larger units such as syllables, words, phrases, etc.?
- What are the permissible and impermissible sequences?
- What kinds of modifications occur in permissible and, particularly, in impermissible sequences?

Additionally, such questions as:

- How do children learn how to perceive speech sounds?
- How are speech sounds from a different language learned?

are also of interest, not only to phonologists and phoneticians, but also to cognitive psychologists, among others.

In this chapter, we will take a first step into the study of speech and learn how speech sounds are made by our vocal tract. To do so, it's necessary that we first identify the systems in the body and their roles in the speech production process. Then we will discuss how and where the articulators move to create two main types of speech sounds, namely consonants and vowels.

Systems in Speech Production

In physical terms, speaking is simply the act of making air moving through the mouth and the nose audible to communicate thoughts. It is not a random act, but one that is planned and controlled by the speaker's linguistic knowledge of his or her language and neurally coded in the brain.

As shown in Figure 1.1, speech production begins with a speaker's desire to communicate a thought to a listener. The thought is converted into linguistic representations and coded as neural impulses in the brain. These neural impulses are converted into muscle movements in the vocal tract, generating an outbound acoustic signal. The acoustic signal is picked up by the listener's auditory system and decoded by the brain for meaning. In linguistics, meaning refers to the information (i.e., the concept, the referent, the action, etc.) that the speaker intends to convey to the listener.

Speech production involves three systems in the body: **the respiratory system**, **the phonation system**, and the **articulation system** (Figure 1.2).

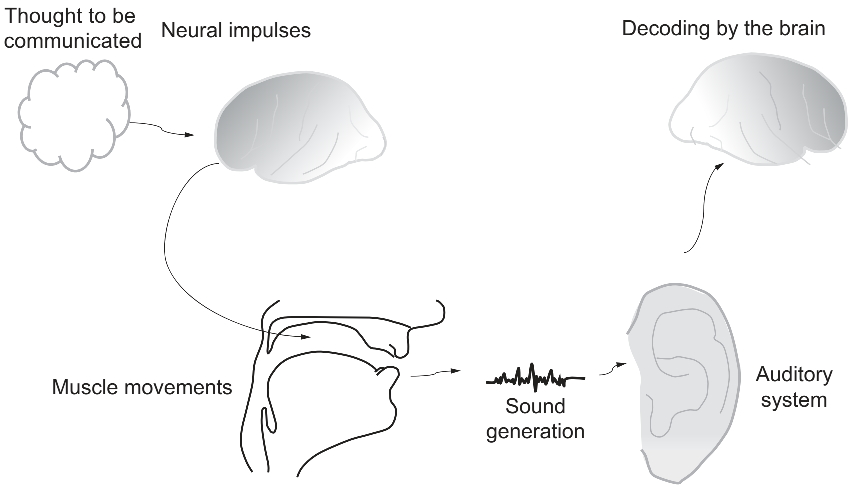

Figure 1.1 Speech production process

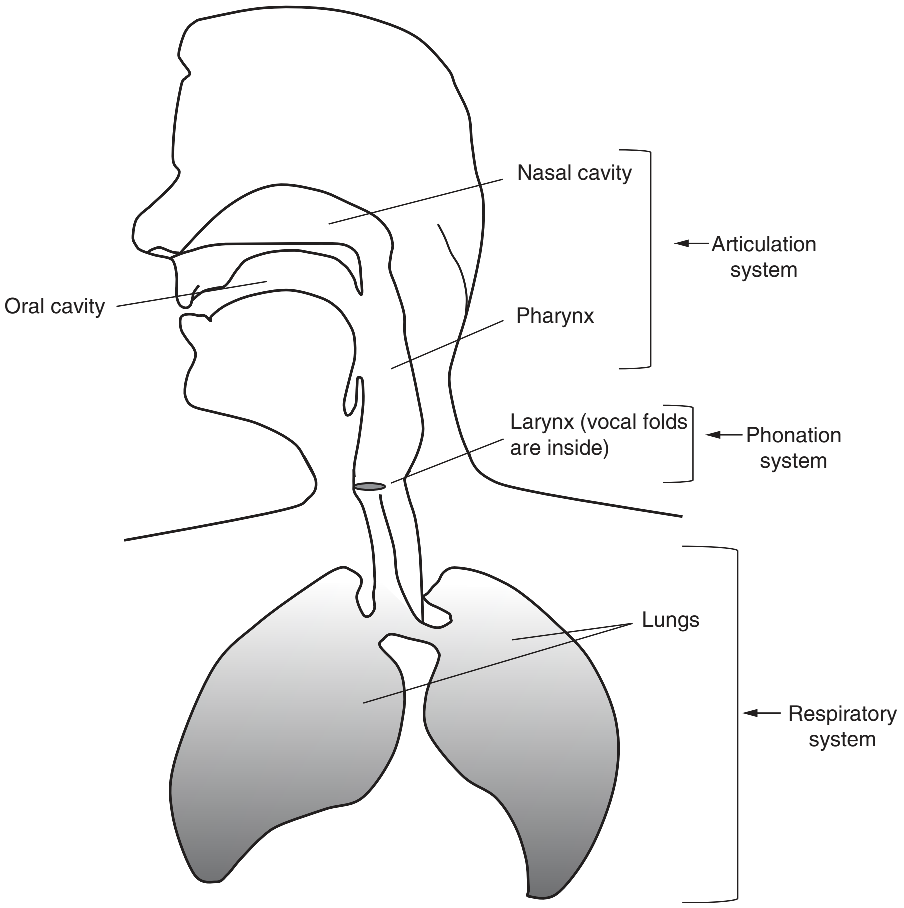

Figure 1.2 Speech production systems

These systems are primarily designed for biological functions such as breathing and eating, and are only secondarily adapted for speech. The respiratory system provides the air supply needed to produce speech. Most speech sounds are powered by pushing air from the lungs upward and outward, that is, the pulmonic airstream. However, many sounds are initiated by moving air trapped inside the vocal tract using the larynx and the tongue. These other airstream mechanisms are discussed more fully in Chapter 2.

The phonation system modifies the airflow as it passes through the larynx. Different adjustments of the vocal folds, two flaps of layered tissue inside the larynx, produce different phonation types, a topic more fully treated in Chapter 2. Finally, the articulation system, which includes the pharyngeal cavity, the mouth, and the nose, performs the final shaping of the airflow before it exits the body. The where and the how of the airflow being shaped are referred to as the place and the manner of speech articulation, respectively. We discuss both later in the chapter. Let's first take a look at each of the three systems in a little more detail.

The Respiratory System

The respiratory system supplies the air needed to initiate speech sounds (see Figure 1.3). It consists of parts of the body that allow us to breathe, including the lungs, the diaphragm, the muscles of the rib cage, and the abdominal muscles. To initiate speech, air has to be drawn in and forced out of the lungs. However, since the lungs do not have muscles of their own, they must rely on muscles of

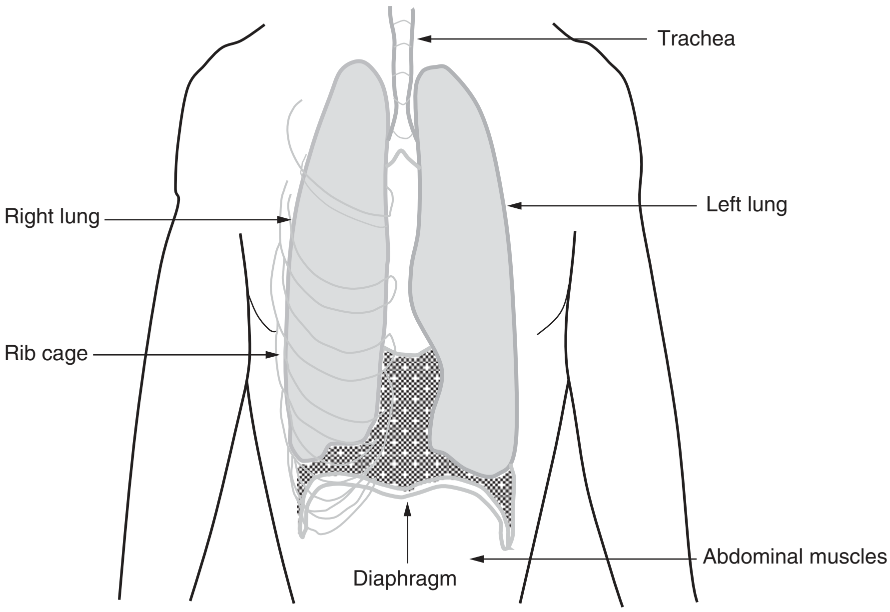

Figure 1.3 The respiratory system

the ribcage and the abdomen to expand or to contract. The diaphragm is a large dome-shaped muscle separating the chest cavity from the stomach. The dome flattens when the diaphragm contracts to allow the lungs to expand during inhalation. Air is squeezed out as the diaphragm and the muscles of both the rib cage and the abdomen contract during exhalation. Air flows upward from the lungs and is then modified by the phonation system.

The Phonation System

The phonation system comprises the larynx and its internal structure. Formed by two major cartilages, the thyroid and the cricoid, the larynx (commonly known as the Adam's apple) sits on a ring of connecting cartilage known as the trachea, or the windpipe (Figure 1.4a). Inside the larynx are the vocal folds (Figure 1.4b, and History Box). They are made up of layers of tissue attached to a pair of **arytenoid** cartilages at the back end, and to the **thyroid** cartilage at the front end. Movements of the arytenoids either bring the vocal folds together (adduct) and close off the airflow from the lungs, or move them apart (abduct) to allow the upward flow of air without obstruction. The spacing between the vocal folds is the **glottis**.

History Box: Vocal Cords

In 1741, Antoine Ferrein, a French anatomist, hypothesized that the ligaments of the larynx are similar to the cords of a violin, and therefore, erroneously called the vocal folds 'cordes vocales' or the vocal cords, the term that is still commonly used today.

Voicing occurs when the air from the lungs pushes the closed vocal folds apart, causing them to vibrate. Sounds produced with vocal fold vibration are **voiced**. In contrast, **voiceless** sounds are those made without vibration of the vocal folds.

An idealized cycle of vocal fold vibration is depicted in Figure 1.5. First, the vocal folds are adducted, closing off the airflow from the lungs (a). Air pressure builds up underneath the closed vocal folds (sub-glottal pressure). When the sub-glottal air pressure becomes greater than the air pressure above the vocal folds (supra-glottal pressure), the vocal folds are pushed apart, from the bottom layer (b) to the top layer (c, d), and set into vibration. According to the myoelastic-aerodynamic theory, the elasticity of the vocal folds and the aerodynamic mechanism known as the *Bernoulli principle* cause the vocal folds to close, again from the bottom layer (e) to the top layer (f, g), and the whole cycle repeats until the air in the lungs is exhausted (see also In Focus). However, current theories contend that the vocal folds would vibrate when there is an asymmetry between the aerodynamic forces created within the glottis and the opening and closing phases of the vocal folds, and that the Bernoulli effect plays only a secondary role.

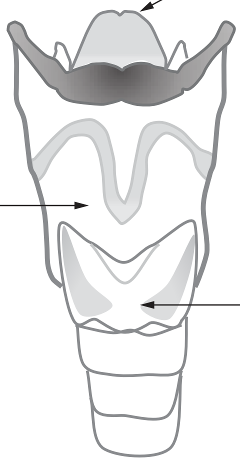

Figure 1.4a Front view of the larynx

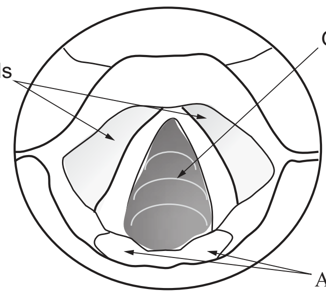

Figure 1.4b Top view of the larynx

Other modes of voicing are possible with different configurations of the glottis. This is discussed in more detail under phonation types (Chapter 2).

In Focus: Bernoulli Principle

The Bernoulli principle, named after the Swiss-Dutch mathematician Daniel Bernoulli, is invoked to explain why the vocal folds close after they are blown open by the force of sub-glottal pressure. Originally applied to fluid dynamics, the Bernoulli principle states that, for the flow of a fluid that has no viscosity, when the speed of the fluid increases, its pressure or its fluid potential energy decreases. The Bernoulli principle is also applicable to the flow of gas or air. In other words, as the speed of the flow of air increases, the

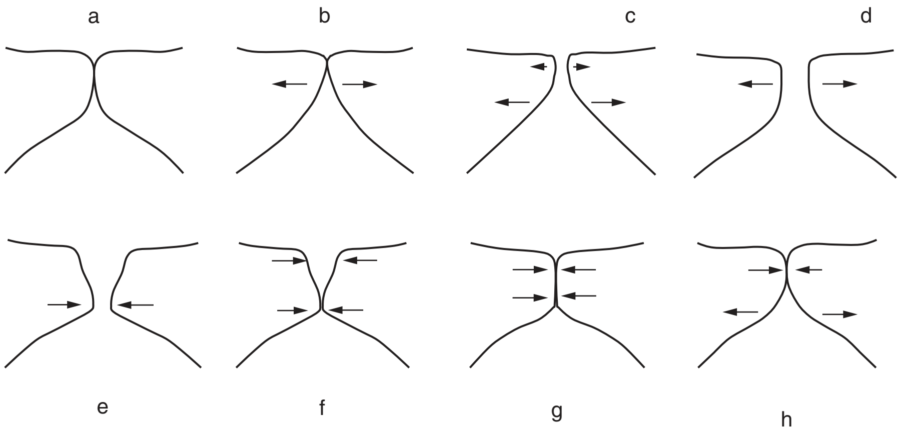

Figure 1.5 Voicing cycle
(Redrawn from Story, B. H. (2002). An overview of the physiology, physics and modeling of the sound source for vowels. Acoustical Science and Technology, 23(4), 195–206, with permission from the author, Dr. Brad Story.)

pressure of the flowing air decreases. In a vocal fold vibratory cycle, to maintain a steady rate of airflow below and above the glottis, the airflow rate at the glottis increases because of the narrowing of the passage, causing a drop in air pressure at the glottis. This drop in air pressure and the elastic property of the vocal folds allow the vocal folds to close. This pattern of the vocal folds closing and opening continues as long as the air supply lasts.

The Articulation System: The Vocal Tract

Airflow through the glottis is further modified inside the vocal tract, which consists of three main cavities: the pharyngeal cavity, the oral cavity, and the nasal cavity (Figure 1.6). The upper surface of the oral cavity contains relatively stationary or **passive articulators**, including the upper lip, the teeth, the alveolar ridge, the hard palate, the soft palate (also called the velum), and the uvula. The lower lip and the tongue are the main mobile or **active articulators** on the lower surface of the vocal tract. Different parts of the tongue are involved in speech production, and it is divided into different areas: tongue tip, blade, front, mid, body, back, and root. Active articulators move toward passive articulators to form varying degrees of constriction, which shape the airflow before it leaves the vocal tract as distinct speech sounds.

Describing Speech Sounds

Air movement can be made audible in a number of ways. It can be blocked, causing the pressure to build up and then suddenly be released. It can be forced to

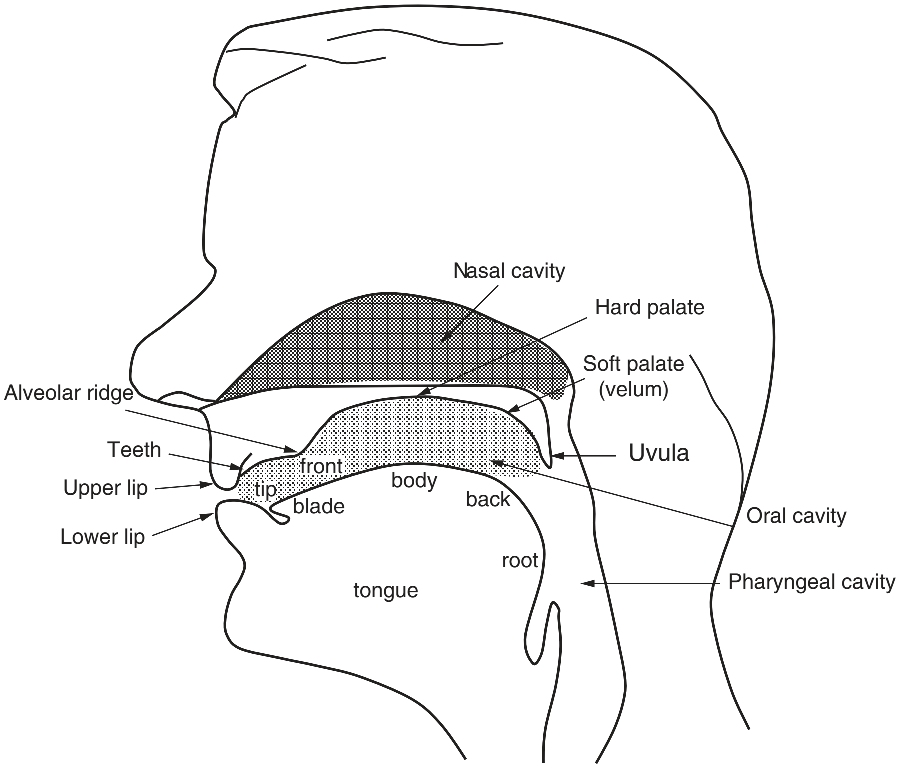

Figure 1.6 The articulatory system (vocal tract)

move fast through a small channel to generate friction or turbulence. Air can also be heard when its molecules vibrate or resonate inside a space such as the body of a guitar or, in the case of speech, the oral cavity. Languages make use of these methods, alone or combined, to make speech sounds audible.

Speech sounds are divided into two broad classes: vowels and consonants. Vibration of air molecules is the main source of energy for vowel sounds, whereas an abrupt release and turbulence alone or in combination with molecule vibration make consonants audible.

Consonants and vowels differ in how they are produced and where they occur in a syllable or a word. In comparison to vowel production, the vocal tract is more constricted, airflow is more obstructed, and less acoustic energy is generated during consonant production. In addition, while consonants can be voiced or voiceless, vowels are commonly voiced. Furthermore, producing a consonant takes a relatively shorter amount of time to complete than producing a vowel. Consequently, consonants are generally shorter and perceptually weaker than vowels.

Besides their articulation, consonants and vowels also differ in number, patterning, and function. Consonants outnumber vowels and tend to occur at the beginning or at the end of a syllable. Vowels more frequently occur at the center of the

Table 1.1 *Differences between vowels and consonants*

Vowels	Consonants
More open vocal tract	Constricted vocal tract
Generally voiced	Can be voiced or voiceless
Higher acoustic energy (perceptually louder)	Less acoustic energy (perceptually softer)
Occur more often at the center of a syllable or a word	Likely occur at edges of a syllable or a word
Function as syllable nuclei	Only some consonants can function as syllable nuclei
Function as the pitch-bearing unit	Only some consonants can bear pitch information
Fewer in number	Larger in number

syllable and function as syllable nuclei. They also serve as the pitch-bearing unit. The differences between vowels and consonants are summarized in Table 1.1.

Consonants

As already mentioned above, consonants are produced with a complete or partial obstruction of airflow. Therefore, one way to describe them is by the amount of airflow obstruction or the size of the oral constriction, commonly referred to as manner of articulation.

Manner of Articulation

Major manners of articulation are plosives, fricatives, approximants, lateral approximants, trills, tap/flaps, and affricates. These consonants are described in the following section with examples drawn from English and other languages. English examples are given in both standard orthography (in italics) and transcribed using the International Phonetic Alphabet (IPA) enclosed in the square brackets []. Examples from other languages are written in the IPA. IPA transcriptions are enclosed in square brackets. Try to memorize each symbol and the sound it represents as you go. We will discuss the IPA chart in Chapter 4.

Plosives are produced with the highest degree of constriction in the vocal tract. The articulators make a complete closure and momentarily block off the airflow before releasing it abruptly. Both oral and nasal plosives occur. For oral plosives, the velum is raised (velic closure) to seal off the nasal cavity so that the air can only exit through the mouth. For nasal plosives, the velum is lowered, and the airflow exits through the nose (Figure 1.7). Plosive sounds such as [p] and [b] in English words ***peak*** and ***beak*** are oral while [m] and [n] in ***meat*** and ***neat*** are nasal consonants.

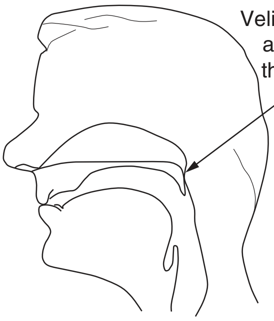

Figure 1.7a Oral plosives [p], [b]

Figure 1.7b Nasal plosive [m]

Figure 1.7 Oral plosives are produced with a raised velum (velic closure) to prevent air from escaping through the nose while nasal plosives are produced with the velum lowered allowing air to escape through the nasal cavity

Fricative consonants are produced by two articulators approximating each other, forming a narrow channel to obstruct the air passage. Turbulent and noisy hissing or frication occurs as high-pressured airflow is sped up and forced through the small slit or groove between the articulators. The first consonants [f, v, θ, ð, s, z, ʃ] in the English words ***f**ive*, ***v**ibe*, ***th**igh*, ***th**y*, ***s**ip*, ***z**ip*, ***sh**ip,* and the medial consonant [ʒ] in *vi**s**ion* are fricative consonants (see Figures 1.8a for labio-dental fricatives [f, v], and 1.8b for post-alveolar fricatives [ʃ, ʒ]).

(Central) approximant consonants are produced when an active articulator approaches a passive articulator without narrowing the vocal tract to the extent that a turbulent airstream is produced. The first consonant [j] in the English word ***y**arn* is an example of an approximant consonant produced with the front of the tongue raised toward the hard palate (Figure 1.9a). The first consonants [w, ɹ] in ***w**ide* and ***r**ide* are the other two English approximants. For [w], the back of the tongue is raised toward the velum (soft palate) with lip rounding (see also "Complex Consonants" below) (Figure 1.9b). For, [ɹ], there are many ways to