Speech Science Primer

Physiology, Acoustics, and Perception of Speech

Third Edition

1994

Gloria J. Borden, Ph.D.
Emerita, Department of Speech, Temple University, Philadelphia, Pennsylvania, and
Haskins Laboratories, New Haven, Connecticut

Katherine S. Harris, Ph.D.
Department of Speech and Hearing Sciences, The Graduate School,
City University of New York, New York, New York, and
Haskins Laboratories, New Haven, Connecticut

Lawrence J. Raphael, Ph.D.
Department of Speech and Theatre, Herbert H. Lehman College,
City University of New York, Bronx, New York, and
Haskins Laboratories, New Haven, Connecticut

LIPPINCOTT WILLIAMS & WILKINS
A **Wolters Kluwer** Company
Philadelphia • Baltimore • New York • London
Buenos Aires • Hong Kong • Sydney • Tokyo

is less a problem of inability to change lung volumes sufficiently to produce appropriate pressure than a problem in affording proper resistance to the airstream at the vocal folds or in the upper vocal tract. Sometimes, a speaker with a motor dysfunction will contract all of the internal intercostal muscles concurrently with the external intercostal muscles as he inhales, a seemingly contradictory maneuver, even though the external intercostals win the battle (the speaker actually inhales). Often, one can observe in less efficient breathers a tendency to expend a great deal of muscular energy lifting the sternum and upper rib cage, a process sometimes called clavicular breathing, when with the same energy applied to different muscles, a speaker could lift the lower ribs and produce a greater thoracic expansion. People with normal speech, however, seem to vary a great deal in whether they move the abdomen-diaphragm region more or the upper chest, and may engage in contradictory respiratory maneuvers, as people with speech disorders do.

PHONATION

Conversion of Air Pressure into Sound

The power supply for speech is the expired air from the lungs, but it falls to the actions of the upper airways to convert this air supply into audible vibrations for speech. As mentioned before, speakers use two methods of transforming the air into sounds for speech. The first method involves using the air pressure to set the elastic vocal folds in the larynx into vibration, producing a periodic sound wave (one with a repeated pattern). The second method involves allowing air to pass through the larynx into the upper vocal tract (the passages between the vocal folds and the outside air), where various modifications of the airstream result in noises: bursts, hisses, or combinations of these aperiodic sound waves (with no repeated pattern of vibration). The first method is termed *phonation*, and it is this vocal mode and its variation that we now consider.

Myoelastic Aerodynamic Theory of Phonation

The vocal folds are shelf-like elastic protuberances of tendon, muscles, and mucous membrane that lie behind the "Adams's apple" or thyroid cartilage and run in an anterior-posterior direction. Their tension and elasticity can be varied; they can be made thicker or thinner, shorter or longer; they can be opened wide, closed together, or put into intermediate positions, and they can be elevated or depressed in their vertical relationship to the cavities above. In running speech, all these adjustments occur at very rapid rates. These dynamic variations of vocal fold position and shape are the result of an evolutionary change from a simple sphincter or valve-like mechanism in lower forms of life to the human larynx, in which the muscles controlling the folds are divided into several groups with specific functions, allowing a wide range of adjustments.

When the vocal folds are together and vibrating, they are in the phonatory mode. Before considering the laryngeal structures and their functions in phonation, you can gain an immediate understanding of vocal fold physiology by producing that vibration of the lips known in the United States as the "Bronx cheer" and in Great Britain as a "raspberry." Insure your privacy and then put your lips together in such a way that air pressure from behind sets them into an audible vibration. The sound is that of air escaping in rapid bursts, not the sound of the lips moving. It is apparent that it is the air pressure producing the lip movements, not the lip muscles, and yet the lips have to be put together and with just the right amount of pressure for it to work. Try it with lips slightly apart, or tightly pursed and you will meet with failure. Although the "Bronx cheer" is quite easily observed, there was not general agreement until relatively recently that the vocal folds worked in somewhat the same way.

In the middle of the 18th century, it was generally thought that vocal folds vibrated like strings, thus directly producing vibrations in air. Even as late as 1950, Husson, with the neurochronaxic theory, proposed that the vocal folds vibrated as a consequence of individual nerve impulses, at the fundamental frequency rate, to the vocalis muscle rather than as a consequence of the action of expired air on the vocal folds. The currently accepted theory of

phonation,
posed by b
19th centur
a series of p
aerodynami
word is aero
vated by the
than by nerv
the ways in
their elastici
frequency of

The nu
and close pe
fold vibratio
bration dire
quency (fun
that is produ
average fund
mately 125 H
with a f_0 ove
Hz. Size of t
of fundamen
brating mass
frequency. T
folds than wo
approximatin
vocal folds ar
and 17 mm.
particular ler
son can incre
preciably by l
thus decreasi
vocal folds m
Singers are tr
(each octave i
voice can go a
soprano highe
thus importar
cles act to bri
can vibrate, a
ness and tensi
quency.

The esse
ory, however,
vibratory cycl
from the lung
vibration. The
ing each vibr
elasticity and

Adams's apple"
n anterior-pos-
and elasticity
thicker or thin-
e opened wide,
rmediate posi-
or depressed in
cavities above.
stments occur
mic variations
are the result
om a simple
nism in lower
x, in which the
re divided into
tions, allowing

gether and vi-
y mode. Before
ures and their
gain an imme-
physiology by
lips known in
cheer" and in
sure your pri-
ther in such a
ind sets them
und is that of
the sound of
at it is the air
nents, not the
ve to be put
t amount of
h lips slightly
vill meet with
eer" is quite
eneral agree-
hat the vocal
me way.
entury, it was
vibrated like
vibrations in
with the neu-
at the vocal
of individual
tal frequency
han as a con-
d air on the
ed theory of

---

phonation, however, is essentially that proposed by both Helmholtz and Müller in the 19th century and amplified by van den Berg in a series of papers in the 1950s, the *myoelastic aerodynamic theory* of phonation. The key word is aerodynamic. The vocal folds are activated by the airstream from the lungs rather than by nerve impulses. "Myoelastic" refers to the ways in which the muscles (myo-) change their elasticity and tension to effect changes in frequency of vibration.

The number of times the vocal folds open and close per second is the frequency of vocal fold vibration. The frequency of vocal fold vibration directly determines the lowest frequency (fundamental frequency) of the sound that is produced. Men have voices that have an average fundamental frequency (f_0) of approximately 125 Hz. Women are more apt to phonate with a f_0 over 200 Hz, and children, over 300 Hz. Size of the vocal folds is one determinant of fundamental frequency. The larger the vibrating mass of the vocal folds, the lower the frequency. Typically men have larger vocal folds than women. They have vocal fold lengths approximating 17 to 24 mm, while women's vocal folds are more apt to range between 13 and 17 mm. Given a pair of vocal folds of a particular length and weight, however, a person can increase the frequency of vibration appreciably by lengthening and tensing the folds, thus decreasing the effective mass. Usually, the vocal folds may be stretched by 3 or 4 mm. Singers are trained to have a range of 2 octaves (each octave is a doubling of frequency). A bass voice can go as low as about 80 Hz and a lyric soprano higher than 1 kHz. Muscular action is thus important in the control of voicing. Muscles act to bring the folds together so that they can vibrate, and muscles regulate their thickness and tension to alter the fundamental frequency.

The essential point of the myoelastic theory, however, is that the determinants of the vibratory cycle are aerodynamic. Air pressure from the lungs opens the glottis during each vibration. The folds come together again during each vibration because of their inherent elasticity and the sudden pressure drop between the folds (Bernoulli principle) as the air streams through the open glottis.

More recent theorizing about vocal fold vibration has led to the development of the Cover-Body Theory (also called the Two-Mass Model) of vocal fold vibration. This theory attempts to account for the wide range of frequencies, amplitudes, and vocal qualities that a human can produce while phonating in terms of the unique structure of the vocal folds. That structure, described in great detail by Hirano and others, consists of a relatively pliable "cover," formed partly of mucous membrane, which overlies a relatively stiff "body," composed mainly of muscle fibers. The vibration of the cover ranges from being very similar to being very dissimilar to that of the body. The degree of similarity or difference, which determines many of the acoustic characteristics of phonation, depends on the activity of the intrinsic laryngeal muscles.

Whatever the theory, an explanation of the details of phonation requires anatomical knowledge of the larynx. Only the anatomy essential to a basic understanding of vocal fold function for speech will be presented in this text.

Framework of the Larynx

In addition to its use for speech, the larynx is used to control the flow of air into and out of the lungs, providing oxygen to the body and eliminating carbon dioxide, to prevent food, water, or other substances from entering the lungs, to aid in swallowing, and to enable a build-up of pressure within the thorax for such functions as coughing, vomiting, defecating, and lifting heavy objects.

The larynx is suspended from the *hyoid bone* and sits on top of the *trachea* (Fig. 4.31). The trachea, formed by a series of horseshoe-shaped cartilages with the open part at the back, can be located at the base of the neck, while the hyoid bone is floating under the jaw and can best be felt by tilting the head back slightly. A small horseshoe-shaped bone, it can be distinguished from the cartilages by its rigidity. The laryngeal framework lies anterior to

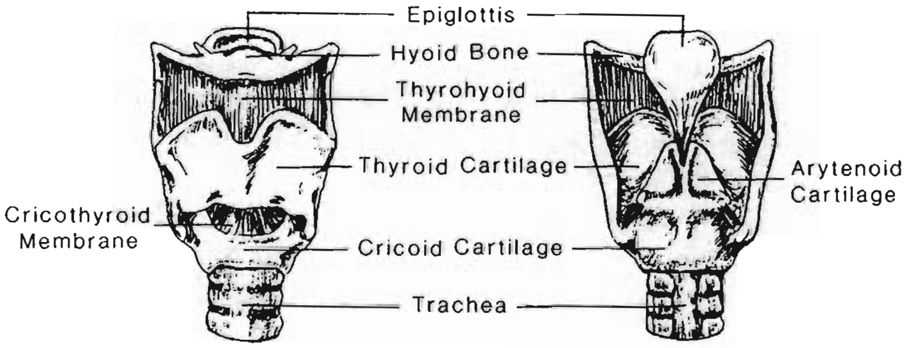


Figure 4.31. Anterior and posterior views of the larynx.

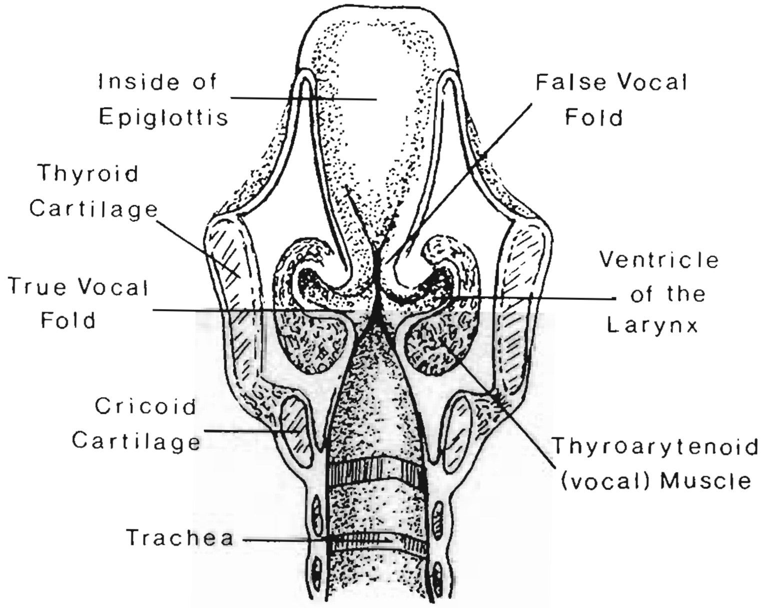


Figure 4.32. Frontal section of the larynx. Notice the constrictions formed by the ventricular folds and the "true vocal folds" below.

the lower pharynx, which leads to the esophagus and on to the stomach. Therefore, food and liquids must pass over the entrance to the lungs to gain access to the entrance to the stomach, a seemingly inefficient arrangement, which is the price paid for the evolutionary adaptation of the larynx as a sound source for speech. During swallowing, a leaf-shaped cartilage, the *epiglottis*, covers the entrance to the larynx. In other animals, the larynx is positioned high in the throat and can be coupled to the nasal airways, in which case food and liquids pass from the mouth around the sides of the larynx and straight into the esophagus, with no danger of entering the windpipe.

The larynx is a tube composed of cartilages connected by ligaments and connecting membranes and covered by mucous membrane. The enclosed area forms an hourglass space (Fig. 4.32) with a vestibule above two sets of folds, the *ventricular folds* or "false vocal folds," and the "true vocal folds" used for phonation. The ventricular folds form a second constriction, just above the true vocal folds. The vertical space between the two sets of folds is called the *laryngeal ventricle*, and the horizon-

---

tal space bet
the *glottis*.
widens agai
work.

The ca
laryngeal sp
regulate its
and *aryteno*
so named be
(Greek, krik
sidered to b
the top rin
because of t
its posterio
tracheal rin
narrow fron
arch, and t
the signet-l
teriorly (Fi

Altho
to the cric
with three
folds: the t
cartilages.
pyramidal i

For arti
aryter

C

Figure 4.3

Arytenoid
Cartilage

ɔid
le

olds and the "true

› the esophagus,
e windpipe.
nposed of carti-
and connecting
mucous mem-
ıs an hourglass
e above two sets
or "false vocal
used for phon-
n a second con-
ocal folds. The
sets of folds is
ıd the horizon-

tal space between the true vocal folds is called the *glottis*. Below the vocal folds the space widens again within the cartilaginous framework.

The cartilages that serve to maintain the laryngeal space and to support the muscles that regulate its changes are the *thyroid, cricoid,* and *arytenoid cartilages*. The cricoid cartilage, so named because it is shaped like a signet ring (Greek, krikos = ring, oid = like), can be considered to be an overgrown tracheal ring. It is the top ring of the trachea, and is distinctive because of the large plate (lamina) that forms its posterior surface, in contrast to the other tracheal rings, which are open at the back. The narrow front and sides of the cartilage form the arch, and the broad lamina at the back forms the signet-like part of the ring, which faces posteriorly (Fig. 4.33).

Although the vocal folds are not attached to the cricoid cartilage, the cricoid articulates with three cartilages that do support the vocal folds: the thyroid cartilage and two arytenoid cartilages. The arytenoid cartilages are roughly pyramidal in shape and articulate with the cricoid cartilage at oval depressions on their inferior surfaces that correspond with convex facets on the top sides of the cricoid lamina. When the arytenoids are in place, a small projection at the base of each cartilage (the vocal process) points anteriorly and is the point of attachment for the vocal ligament with its associated folds. The vocal ligament and the *thyroarytenoid muscle* lying along it are stretched between the vocal process of the arytenoid cartilages in the back and the deep angle of the thyroid cartilage in the front (Fig. 4.34).

These structures (the vocal process of the arytenoid cartilage, the vocal ligament, the thyroarytenoid muscle, and the mucous membrane lining the inner surface on both sides) form the true vocal folds. The cartilage projects into the posterior one-third of the folds, the cartilaginous portion, leaving the anterior two-thirds as the membranous portion. As might be expected, the cartilaginous part often differs from the membranous part in its pattern of vibration.

The larger extension at the base of each arytenoid cartilage is called the muscular pro-

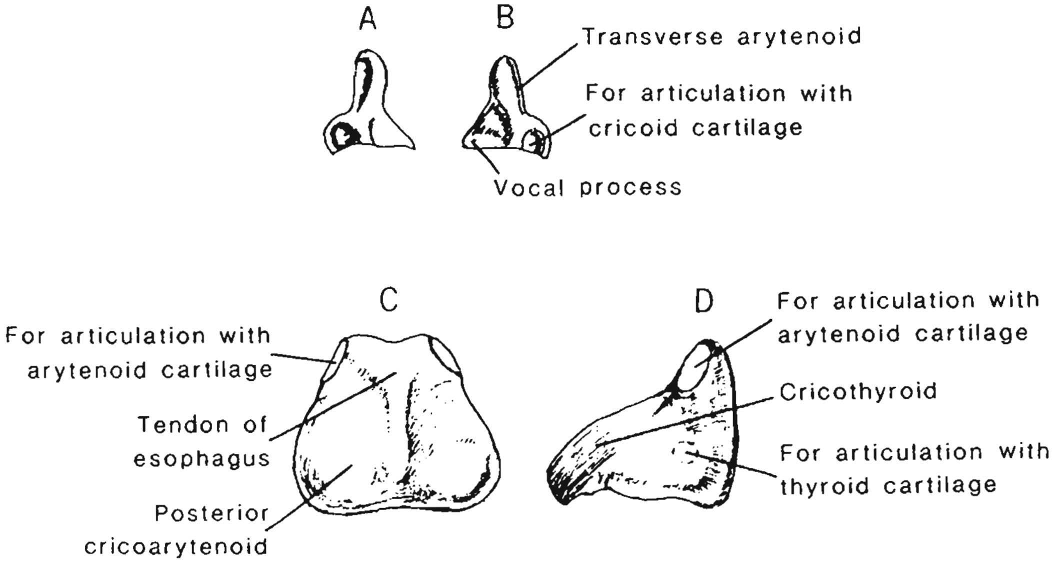


Figure 4.33. The arytenoid and cricoid cartilages. (*A*) The left arytenoid cartilage, medial aspect. (*B*) The right arytenoid cartilage, medial aspect. (*C*) The cricoid cartilage, posterior aspect. (*D*) The cricoid cartilage, left lateral aspect.

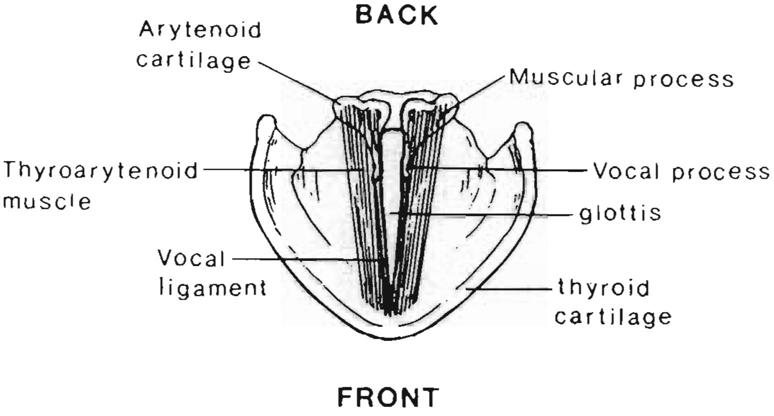


Figure 4.34. The larynx from a superior view, showing the relationships among the thyroid, cricoid, and arytenoid cartilages, and the thyroarytenoid muscle.

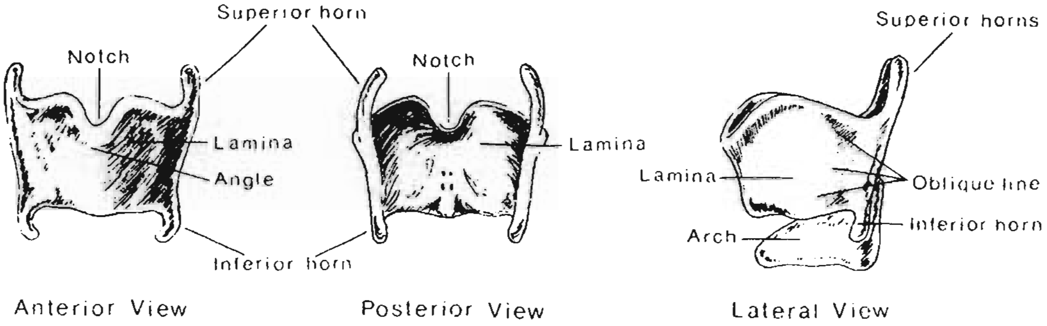


Figure 4.35. Anterior, posterior, and lateral views of the thyroid cartilage. Lateral view includes the cricoid cartilage.

cess because three muscles important for the positioning of the vocal folds are attached to it. The muscular process extends posteriorly and somewhat laterally.

The largest cartilage, the thyroid cartilage, so named because it is like a shield (Greek, thyreos = large shield), is positioned anterior to the arytenoid cartilages, which its sides enclose, and superior to the cricoid cartilage, the top laminal crest of which it also encloses. It forms an angle in the front that is more acute in men (~90°) than in women (~120°), hence the common term "Adam's apple" instead of "Eve's apple." There is a notch (Fig. 4.35) where the laminae separate above the angle that can usually be located by feeling along the midline of your neck with your index finger. The plates are widely separated in the back and extend into two superior horns that project toward the cornua (horns) of the hyoid bone above. The two smaller inferior horns articulate with the cricoid cartilage below by fitting into a round facet on each side of the cricoid lamina. The cartilages of the larynx can move in relation to one another to a limited degree. The thyroid and cricoid cartilages can rock back and forth upon each other. We will describe these motions later in reference to pitch change. The arytenoids can rotate and rock on the cricoid cartilage and can slide a bit toward one another. The muscles attached to the muscular process of the arytenoids control these movements, as we shall see in the following discussion of vocal fold adjustments.

Vocal Fold Adjustments during Speech

The vocal folds at rest are apart (*abducted*), creating a V-shaped glottal space with its apex behind the thyroid cartilage and its widest separation at the back, where the folds attach to the vocal process of the arytenoid cartilages. During running speech, the vocal folds are separated for voiceless speech sounds, such as the consonants /s/ or /t/, are brought together (*adducted*) for voiced sounds, such as the vowels and diphthongs /u/, /i/, and /aɪ/ in the words "two" /tu/, "tea" /ti/, and "tie" /taɪ/. They are less firmly brought together for voiced consonants such as /z/ and /v/ where phonation is needed in addition to large air pressures in the oral cavity (Fig. 4.36).

Voiceless Consonants. The simplest adjustment of the vocal folds for speech is the one made for voiceless consonants. The folds abduct in order to allow the passage of sufficient air from the lungs to create noises in the oral cavity. As speech proceeds, voiceless consonants are interspersed in the speech stream singly or in clusters, demanding rapid glottal opening to interrupt phonation. The job is done by a pair of large triangular muscles attached by tendons to the top of the muscular process of each arytenoid cartilage; the muscle fibers fan out as they course back and down to attach to the dorsal plates of the cricoid cartilage (Fig. 4.37). Named for their position and attachments, the *posterior cricoarytenoid muscles* (PCA) upon contraction rotate the arytenoid cartilages by pulling the muscular processes down and medially, thereby moving the vocal processes apart. Innervation to these and almost all of the other intrinsic muscles of the larynx is supplied by the *recurrent nerve*, a branch of the Xth cranial nerve, the vagus nerve.

Voiced Speech Sounds. Widely separated vocal folds cannot be set into vibration, so in order to produce the voiced sounds of speech, the normally separated folds must be adducted or nearly so. To approximate the vocal folds, the arytenoid cartilages must be brought closer together with their vocal processes rocked inward toward one another. A strong band of muscular fibers runs horizontally across the posterior surfaces of the arytenoid

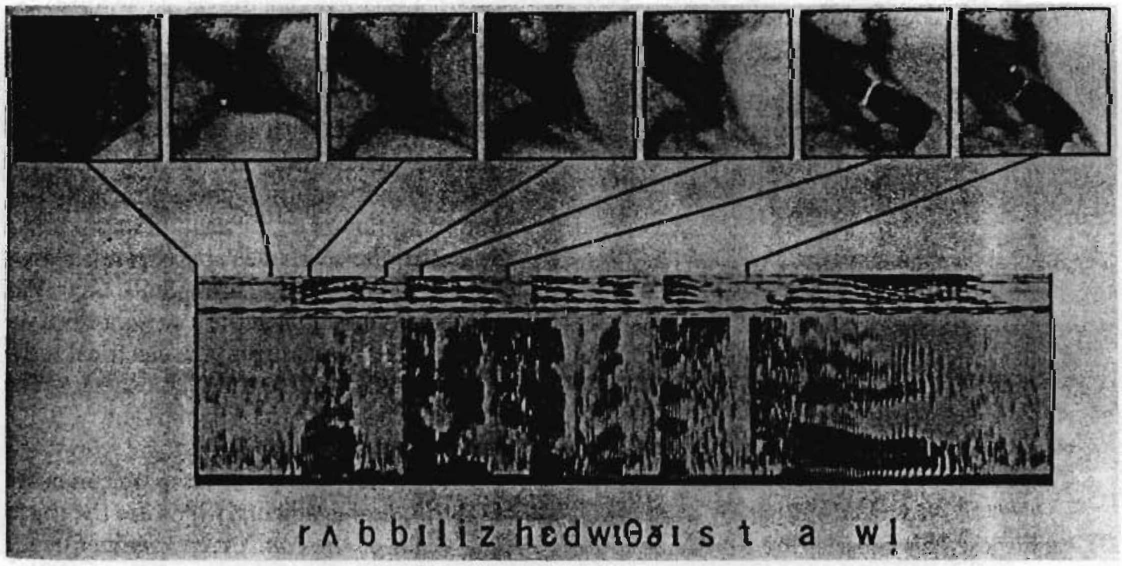


Figure 4.36. The glottis, viewed from above with a fiber bundle, at various times in the production of a sentence. The posterior part of the glottal chink is at the *lower right* in each view. Note the open glottis in the first frame for inspiration, the relatively closed glottis in the third frame for a vowel, and the relatively open glottis in the sixth frame for a voiced consonant. (The structure apparently connecting the folds in the *rightmost views* is a bead of mucus.) (Reprinted with permission from M. Sawashima *et al.*: © 1970.)

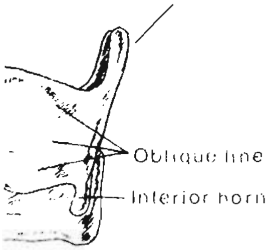
ocess

l process

s

ie

the thyroid, cricoid, and



ral View

l view includes the cricoid

separated in the back and
rior horns that project to-
orns) of the hyoid bone
ler inferior horns articu-
cartilage below by fitting
n each side of the cricoid
es of the larynx can move
other to a limited degree.
icoid cartilages can rock
ı each other. We will de-
later in reference to pitch
ds can rotate and rock on
and can slide a bit toward
scles attached to the mus-
arytenoids control these
hall see in the following
old adjustments.

Vocal Fold Adjustments during Speech

The vocal folds at rest are apart (*abducted*), creating a V-shaped glottal space with its apex behind the thyroid cartilage and its widest separation at the back, where the folds attach to the vocal process of the arytenoid cartilages. During running speech, the vocal folds are separated for voiceless speech sounds, such as the consonants /s/ or /t/, are brought together (*adducted*) for voiced sounds, such as the vowels and diphthongs /u/, /i/, and /aɪ/ in the words "two" /tu/, "tea" /ti/, and "tie" /taɪ/. They are less firmly brought together for voiced consonants such as /z/ and /v/ where phonation is needed in addition to large air pressures in the oral cavity (Fig. 4.36).

Voiceless Consonants. The simplest adjustment of the vocal folds for speech is the one made for voiceless consonants. The folds abduct in order to allow the passage of sufficient air from the lungs to create noises in the oral cavity. As speech proceeds, voiceless consonants are interspersed in the speech stream singly or in clusters, demanding rapid glottal opening to interrupt phonation. The job is done by a pair of large triangular muscles attached by tendons to the top of the muscular process of each arytenoid cartilage; the muscle fibers fan out as they course back and down to attach to the dorsal plates of the cricoid cartilage (Fig. 4.37). Named for their position and attachments, the *posterior cricoarytenoid muscles* (PCA) upon contraction rotate the arytenoid cartilages by pulling the muscular processes down and medially, thereby moving the vocal processes apart. Innervation to these and almost all of the other intrinsic muscles of the larynx is supplied by the *recurrent nerve*, a branch of the Xth cranial nerve, the vagus nerve.

Voiced Speech Sounds. Widely separated vocal folds cannot be set into vibration, so in order to produce the voiced sounds of speech, the normally separated folds must be adducted or nearly so. To approximate the vocal folds, the arytenoid cartilages must be brought closer together with their vocal processes rocked inward toward one another. A strong band of muscular fibers runs horizontally across the posterior surfaces of the arytenoid

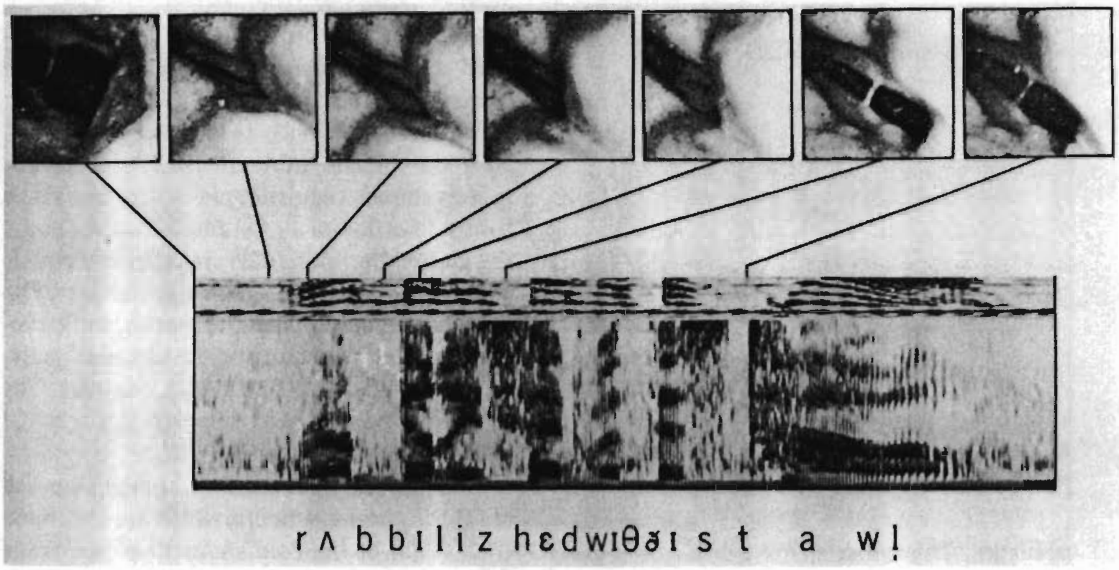


Figure 4.36. The glottis, viewed from above with a fiber bundle, at various times in the production of a sentence. The posterior part of the glottal chink is at the *lower right* in each view. Note the open glottis in the first frame for inspiration, the relatively closed glottis in the third frame for a vowel, and the relatively open glottis in the sixth frame for a voiced consonant. (The structure apparently connecting the folds in the *rightmost views* is a bead of mucus.) (Reprinted with permission from M. Sawashima *et al.*: [illegible] 1970.)

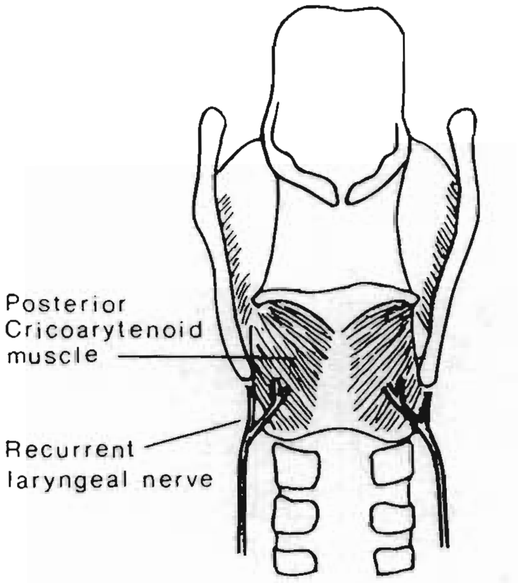


Figure 4.37. Posterior view of the posterior cricoarytenoid muscle. Innervation is via the recurrent branch of the vagus (Xth cranial) nerve.

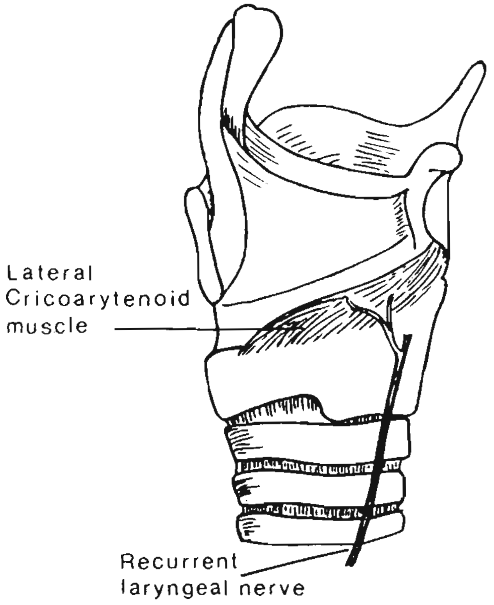


Figure 4.39. Lateral view of the lateral cricoarytenoid muscle with the left side of the thyroid cartilage removed.

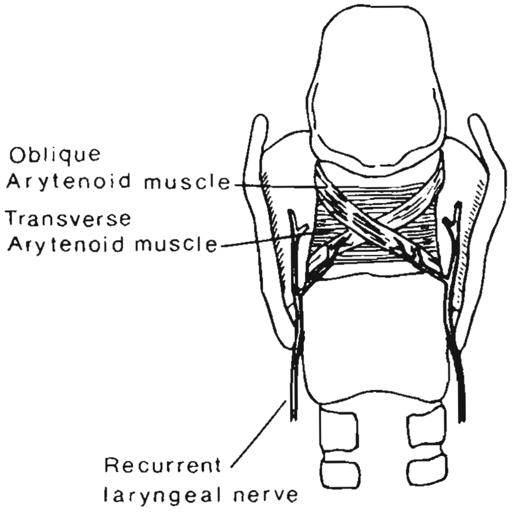


Figure 4.38. Posterior view of the transverse and oblique arytenoid muscles. Together, these muscles are referred to as the interarytenoid muscle.

cartilages. This muscle, the *transverse arytenoid muscle,* is overlaid by some muscular fibers in the shape of an X called the *oblique arytenoid muscles* (Fig. 4.38). Together, termed the *interarytenoid muscle* (*IA*), they adduct the arytenoid cartilages and thereby, the vocal folds. The interarytenoid muscle is thought to be the primary adductor of the vocal folds. The *lateral cricoarytenoid muscles* (*LCA*) also aid in adduction of the vocal folds, by rocking the muscular process of the arytenoids forward and down, thereby pressing the vocal processes together (Fig. 4.39). For stronger adduction of the folds, as in vowel production, both the IA and LCA are usually used. For speech sounds requiring phonation and also a continuous flow of air for a sound source above the glottis, the vocal folds are often less closely completely adducted. Hirose and Gay (Fig. 4.40) have differentiated the functions of laryngeal muscles by measuring the electrical activity generated as these muscles contract. The recording method (electromyography or EMG) is explained in Chapter 7.

The vocal folds themselves are composed of (1) the vocal ligaments, which are the thickened edges of the *conus elasticus* membrane rising from the cricoid cartilage, (2) the muscles that are attached to the ligaments, the internal part of the *thyroarytenoid muscles* commonly called the *vocalis muscles,* and (3) the mucous membrane that covers them. The vocal

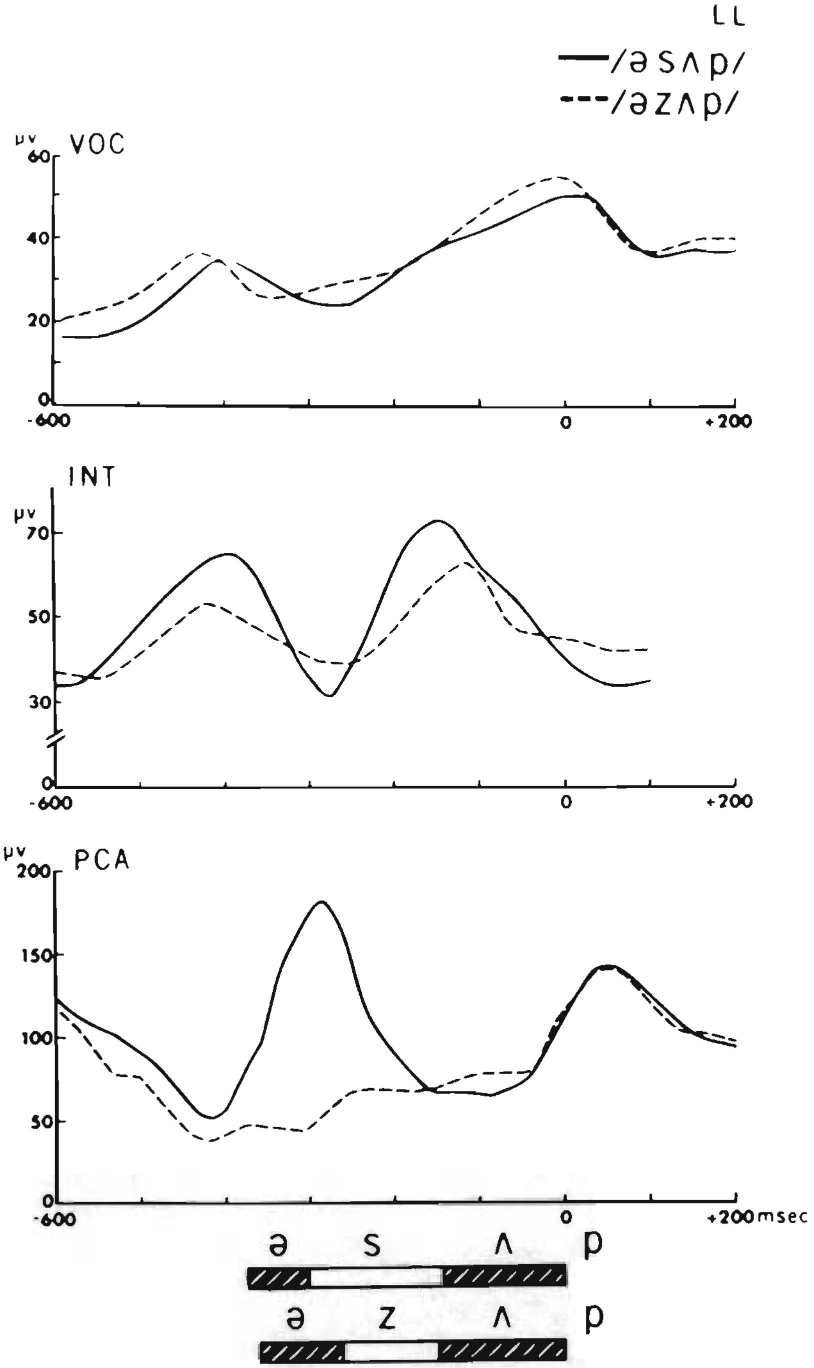


Figure 4.40. Superimposed EMG curves for voiced (z; *dashed line*) and voiceless (s; *solid line*) fricatives in nonsense syllables. Although vocalis muscle (*VOC*) activity is similar for both, for /s/, interarytenoid (*INT*) activity is reduced during the time that posterior cricoarytenoid activity (*PCA*) is greatly increased. (Reprinted with permission from H. Hirose and T. Gay: *Phonetica, 25,* S. Karger AG, Basel, © 1972.)

ligaments and the vocalis muscles, which form the body of the vocal folds, emerge from the projection of the arytenoid cartilage known as the vocal process. The posterior portions of the vocal folds are thus stiffer, because of the presence of the arytenoid cartilage, and become increasingly flexible in their anterior portions. When relaxed, the vocal folds are relatively thick, and open and close in an undulating manner, the mucous membrane moving somewhat independently like flabby skin on a waving arm. The lateral fibers of the thyroarytenoid muscle extend to the muscular process of the arytenoid cartilage, some of them wrapping around the arytenoid where they commingle with the interarytenoid muscles. More research is needed to differentiate the role of the medial (vocalis) and lateral fibers of the thyroarytenoid muscles in phonation, but in general they are thought to tense the folds.

The muscle activity needed to adduct and tense the vocal folds simply readies them for vibration but does not cause the vibration itself. For the "Bronx cheer," you had to put your lips together, and that required muscular effort, but the sound itself was produced by aerodynamic forces acting upon the elastic bodies of your lips. The two aerodynamic forces that produce vibration of the vocal folds are the *subglottal air pressure* (P_s) applied to the lower part of the folds, forcing them apart, and the negative pressure that occurs as air passes between the folds (the *Bernoulli effect*). These positive and negative pressures set the vocal folds into vibration because of the elasticity of the folds.

Subglottal Air Pressure

Consider first the subglottal air pressure that parts the vocal folds. During each opening, a tiny puff of air escapes, cut off sharply by the abruptly closing glottis. Since the folds vibrate rapidly (usually over 100 times per second), one puff follows another in rapid succession. This chain of air puffs sets up a pressure wave at the glottis that is audible. A necessary condition for phonation is that the air pressure below the folds must exceed the pressure above the folds. If the pressure above the folds builds up so that the pressure drop across the glottis necessary for phonation is lost, then phonation ceases. You can test this by trying to prolong a voiced stop consonant such as a [b]. You can phonate during the [b] closure only for a short time because the labial and velopharyngeal closures for the [b] cause the supraglottal air pressure to increase until it equals the subglottal air pressure. Since there is no longer more pressure below the folds than above them, phonation is not possible. For speech at conversational level, a subglottal air pressure in the range of 7–10 cm of H_2O (centimeters of water pressure) is sufficient to produce phonation at approximately 60 dB intensity.

The effect of subglottal air pressure sufficient to separate a pair of vocal folds can be seen in Figure 4.41, schematic diagrams made from a movie of a vibrating larynx. The vocal folds open at the bottom first and then the opening proceeds up to the top of the folds. As the top part of the vocal folds opens, the bottom part can be seen to be closing. Thus, there is a vertical phase difference, creating a wave-like motion of the folds, the normal movement during vibration for chest voice. If the speaker speaks or sings in a high falsetto voice, however, the vertical phase difference is lost and each of the taut folds moves as a unit. The closing phase of each cycle is the result of the tendency of the elastic folds to move into their rest positions and of the second aerodynamic phenomenon important to voicing, the pressure drop ascribed to the Bernoulli effect.

The Bernoulli Effect

Daniel Bernoulli, an 18th century Swiss mathematician and physician, developed the kinetic theory of gases and liquids, part of which is known as the Bernoulli effect or principle. The Bernoulli effect is based upon the observation that when a gas or liquid current runs through a constricted passage, the velocity (speed in a given direction) increases. Simply stated, the Bernoulli principle is that such an increase in velocity results in a drop in the

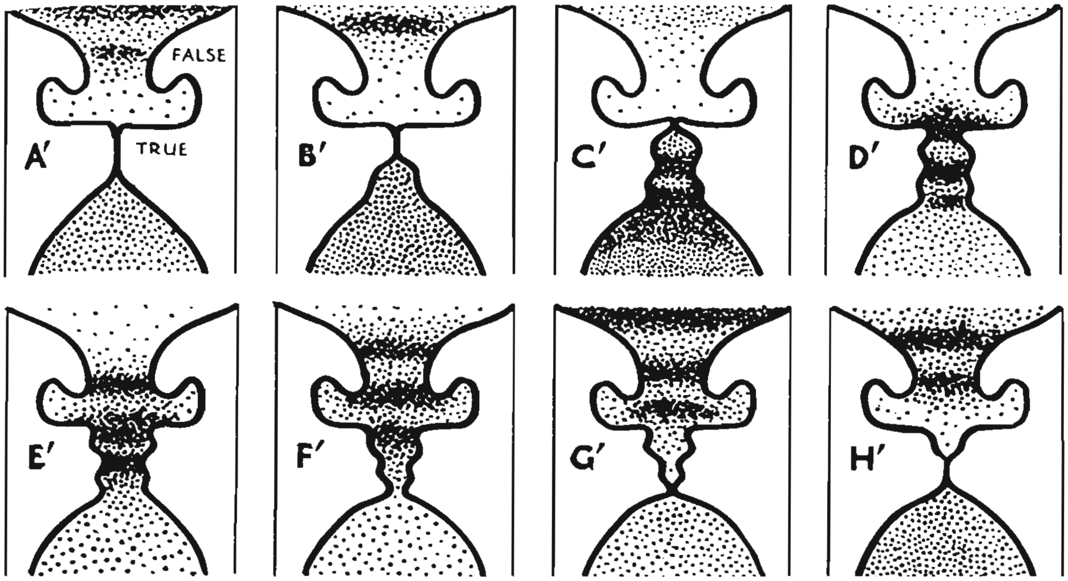


Figure 4.41. Schematic cross sections of the vocal folds during vibration. It can be seen that the folds open and close from *bottom* to *top.* (From *Singing: the Mechanism and the Technic,* 4th Ed., by William Vennard, © 1967 by Carl Fischer, Inc. All rights reserved. Reprinted by permission of the publisher.)

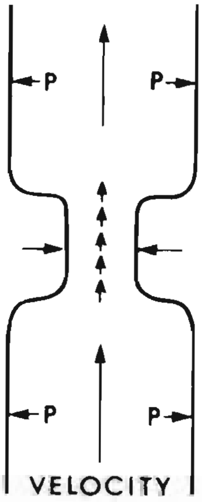


Figure 4.42. Schematic diagram of flow through a constricted passage. In the constriction, velocity is greater, but pressure on the inner sides of the constriction is correspondingly reduced.

pressure exerted by the molecules of moving gas or liquid, the pressure drop being perpendicular to the direction of the flow. Figure 4.42 illustrates the increase in velocity within a narrow portion of a passage and the resulting decrease in pressure against the lateral walls.

The conventional airplane wing is designed to take advantage of the Bernoulli effect to elevate the aircraft. The wing is streamlined on the top surface (Fig. 4.43), permitting a higher velocity of air current than that passing underneath. The higher velocity results in a drop in pressure against the top surface, which creates a difference between the pressures under and over the wings, thereby elevating the plane. You can elevate a piece of paper, using the same principle, by holding one end of it under your lips and blowing air across the top (Fig. 4.44).

We experience the Bernoulli phenomenon constantly. When a draft of air flows through a narrow corridor, the doors opening into rooms off the hall slam shut, because the pressure on the hall side of the doors is lower than that on the room side. If you have ever been in a lightweight car cruising alongside a heavy truck on a highway and felt your car being sucked alarmingly close to the truck, it is because the faster airstream created between your car and the truck has lowered the pressure against the truck side of your car relative to the other side.

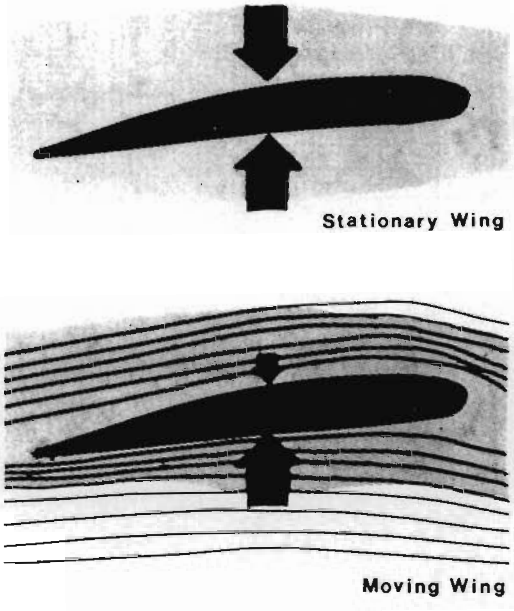


Figure 4.43. Aerodynamic forces on an airplane wing. (See text for discussion.)

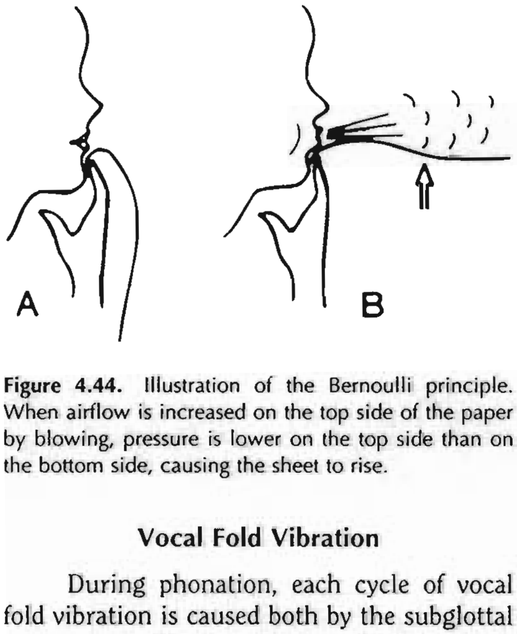


Figure 4.44. Illustration of the Bernoulli principle. When airflow is increased on the top side of the paper by blowing, pressure is lower on the top side than on the bottom side, causing the sheet to rise.

Vocal Fold Vibration

During phonation, each cycle of vocal fold vibration is caused both by the subglottal air pressure that has built up sufficiently to separate the folds and to the Bernoulli effect which, as the air rushes through the glottis at an increased velocity, accounts for a sudden drop in pressure against the inner sides of each fold and sucks them together again. The whole process is made possible by the fact that the folds themselves are elastic. Their elasticity not only permits them to be blown open for each cycle, but the elastic recoil force (the force that restores any elastic body to its resting place) works along with the Bernoulli effect to close the folds for each cycle of vibration.

The vocal folds move in a fairly periodic way. During sustained vowels, for example, the folds open and close in a certain pattern of movement that repeats itself. This action produces a barrage of air bursts that sets up an audible pressure wave (sound) at the glottis. The pressure wave of sound is also periodic; the pattern repeats itself. Like all sound sources that vibrate in a complex periodic fashion, the vocal folds generate a harmonic series (see Chapter 3), consisting of a fundamental frequency and many whole-number multiples of that fundamental frequency. The fundamental frequency is the number of glottal openings per second.

The human voice is a low frequency sound compared to most of the sounds of the world, including the other sounds that humans make above the larynx. Since it contains many harmonics, the voice is also a complex sound. We never hear the unmodified sound of vocal fold vibration, however, because by the time it has reached the lips of the speaker, it has been changed by the vocal tract. If we were to lower a microphone down to the vocal folds, we would record a sound that has a spectrum resembling Figure 4.45. The lowest frequency, the frequency of the vibration itself, sets up a 2nd harmonic (2 times the f_0), a 3rd harmonic (3 times the f_0), and so forth. Notice that it is characteristic of the human voice that the higher harmonics have less intensity than the lower harmonics, so that although the voice contains many high frequency components, the emphasis is on the low frequencies. The intensity falls off at about 12 dB per octave (each doubling of the frequency).

Low pitched and high pitched voices sound different in part because the spacing of their harmonics is different. Figure 4.45 shows the difference. Notice the closer spacing of the harmonics in the adult male's voice, which has

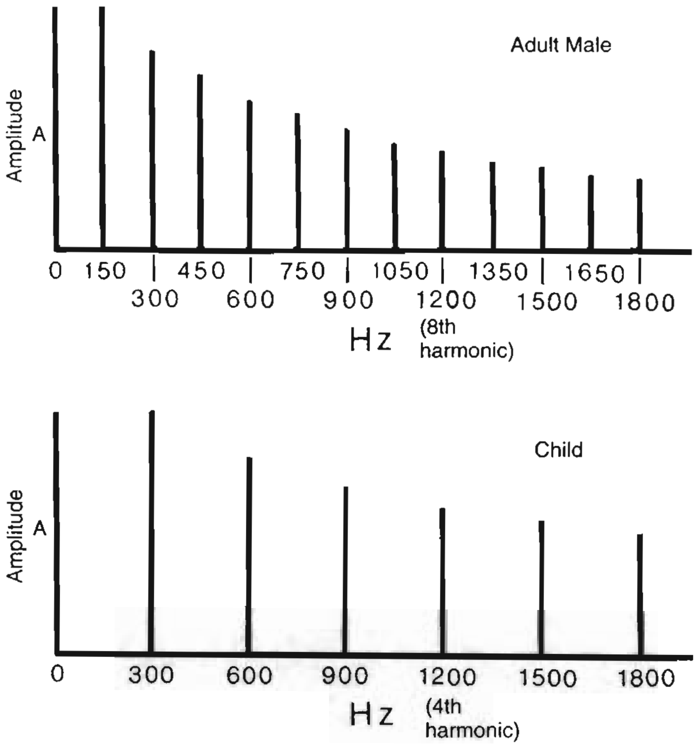


Figure 4.45. Schematic spectra of sounds resulting from vocal fold vibration. The spectra represent two different frequencies of phonation, thus the harmonic spacing is different.

a lower fundamental frequency. A child with a f_0 of 300 Hz would have a 2nd harmonic at 600 Hz, a 3rd at 900 Hz, and a 4th at 1200 Hz. In contrast, a man with a f_0 of 150 Hz would have a 2nd harmonic at 300 Hz, and the 8th harmonic in the adult voice would correspond with the child's 4th harmonic. In the same way, a single person adjusting the frequency of his voice also changes the harmonic spacing. Notice, in the figure, that the shape and slope of the spectrum remain similar for man and child.

Fundamental Frequency

The human voice is composed of many frequencies; it is a complex tone. The human listener perceives the lowest frequency, the fundamental frequency, as the speaker's pitch. The fundamental frequency is constantly changing, as we know when we listen for the *intonation* patterns of sentences. "Are you sure?" has a rising intonation pattern, while "I'm sure" has a falling intonation pattern. The speaker produces these different patterns by altering the fundamental frequency of vocal fold vibration.

According to the myoelastic aerodynamic theory of phonation, frequency of vocal fold vibration is determined by the elasticity, tension, and mass of the vocal folds. More massive folds (longer and thicker) vibrate at naturally lower frequencies than shorter and thinner folds. Vocal folds vibrate faster when they are tense than when they are slack. The primary way to make a given set of vocal folds more tense is to stretch them.

You may have noted that longer folds contribute to increased mass and lower f_0 in one condition and to increased tension and higher f_0 in another condition. This is because a longer pair of vocal folds (compared to other speakers) will be more massive and produce a lower frequency voice; men's voices are lower than children's voices. Yet a lengthening of

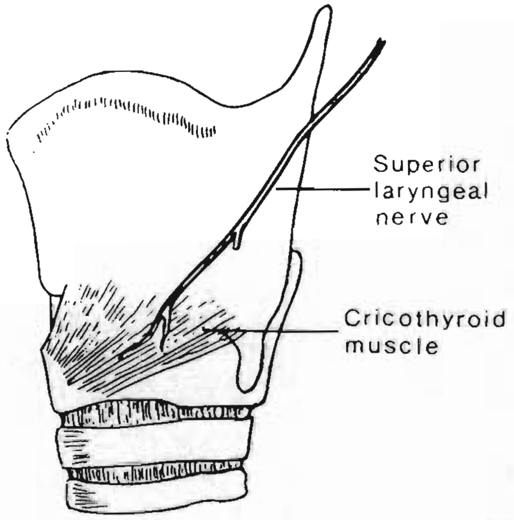


Figure 4.46. Lateral view of the cricothyroid muscle, innervated by the superior laryngeal branch of the vagus (Xth cranial) nerve.

vocal folds (within the same speaker) will stretch and thin the effective vibrating portion of the vocal folds, adding tension and thereby producing a higher fundamental frequency. The pair of muscles responsible for stretching the vocal folds and thereby controlling f_0 change are the *cricothyroid muscles*.

Since the vocal folds lie between the thyroid cartilage and the two arytenoid cartilages, the way to stretch the folds would be to increase the distance between these cartilages. The cricothyroid muscles can do just that. Since they are attached to the side of the cricoid ring and rise (part straight up and the other part at an oblique angle) to the thyroid cartilage, their contraction pulls the two cartilages toward one another by lifting the anterior arch of the cricoid cartilage toward the thyroid cartilage. The closing of the space between the cricoid arch and the front of the thyroid has been likened to the closing of the visor on a suit of armour. Figure 4.46 shows the location of one of the cricothyroid muscles on the left side of the larynx. The effect that their contraction has in elevating the front of the cricoid cartilage, is to tip the posterior plate of the cricoid backward. The arytenoid cartilages ride on the cricoid cartilage and the vocal folds are stretched. Van den Berg refers to this effect of cricothyroid muscle action as longitudinal tension. The innervation of the cricothyroid muscle is from the superior laryngeal nerve (vagus, Xth cranial nerve) unlike all the other intrinsic muscles of the larynx, which are innervated by the recurrent nerve (another branch of the vagus nerve).

The addition of longitudinal tension to the vocal folds increases the fundamental frequency at which they vibrate, at least for much of the frequency range used in speech. For extreme frequencies, other mechanisms are thought to be instrumental in pitch control. At high frequencies, such as for falsetto voice, the cricothyroid is used to further increase tension although no further lengthening is possible. The vocal folds are pulled extremely taut and forego their usual wave-like motion. The vocal ligaments vibrate more like strings.

At extremely low frequencies, the strap muscles of the neck (particularly the *sternohyoid muscle*, Fig. 4.47) assume more responsibility for lowering f_0. You may have noticed the larynx move up slightly in the vertical plane for high frequencies or, more noticeably, move down in the neck for low frequencies. The muscles above the hyoid bone (*suprahyoid mus-*

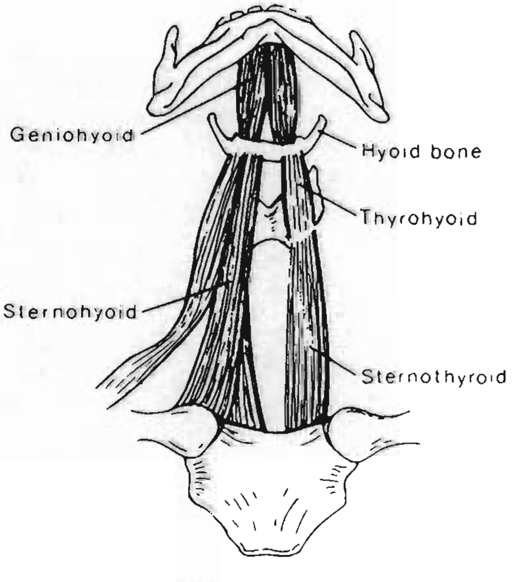


Figure 4.47. The strap muscles of the neck, anterior view from below.

cles) elevate the larynx. These movements are thought by some to add vertical tension to the membranes that serve as a lining for the larynx, and for the trachea below. Increased vertical tension in the conus elasticus during laryngeal elevation and decreased vertical tension in the case of laryngeal lowering would affect the vocal folds. The conus elasticus membrane emerges from the cricoid cartilage and rises in a medial direction to the vocal folds where its thickened border becomes the vocal ligament.

A further source of tension in the vocal folds is the internal tension possible with the contraction of the thyroarytenoid muscles themselves, especially the vibrating portions known as the vocalis muscles. The vocalis muscles are antagonistic to the cricothyroid muscles, since they shorten rather than lengthen the folds, but both sets of muscles can increase tension in the folds to raise f_0. Perhaps the vocalis muscles serve to tune the vocal folds to make the cricothyroid lengthening more effective. Research is needed to clarify the interaction of muscular and non-muscular contributions to frequency change. Atkinson suggests that the relative contributions of various muscular and non-muscular forces may vary in different parts of an individual's pitch range.

It seems that fundamental frequency is primarily affected by applying more or less longitudinal tension to the vocal folds via the cricothyroid muscles. Secondary affects result from (1) applying more or less vertical tension to the folds by elevating or depressing the larynx, and from (2) varying subglottal pressure.

Voice Quality

Much of what distinguishes one voice from another results from the effects of the resonating cavities and structures above the larynx, but part of what is called voice quality or timbre is due to the way in which the vocal folds themselves vibrate. One obvious difference among voices is fundamental frequency, which listeners perceive as pitch. Other differences have to do with how closely the folds are approximated or with irregularities along the edges of the folds. If one or both of the folds are paralyzed, compensations must be made to set up a vibration if possible. Sometimes one vocal fold can be trained to move more than halfway to meet the paralyzed one. If part or all of the larynx has been surgically removed because of cancer, the speaker must learn to vibrate other tissues and muscle masses such as scar tissue or the cricopharyngeus muscle. Some alaryngeal speakers (people whose larynges have been removed) have to resort to an artificial sound source which they hold to the outside of the neck. This produces a "voice" with an artificial, mechanical quality.

Quality differences depend upon various modes of vocal fold vibration. A breathy voice, popular with some movie stars and celebrities in the 1950s, is achieved by failing to adduct the vocal folds sufficiently for full voicing, particularly in the cartilaginous portions. They are close enough to be vibrated, but the sound of continuously released air accompanies the sound wave set up by the air pressure volleys. A hoarse voice is caused by irregularities in the folds. When the vocal folds are irritated or swollen, as they may be during a cold with laryngitis, the voice becomes hoarse. Hoarseness can also be indicative of vocal abuse, either from focusing too much tension in the larynx, causing *contact ulcers*, lesions produced by the arytenoid cartilages banging against one another, or from overusing the voice as happens commonly to women and occasionally to men who develop nodules along the edges of the vocal folds. *Vocal fry* or *creaky* voice (Ladefoged's term) consists of extremely low frequency phonation. About the only practical value of vocal fry is that (because of its low frequency) it provides an audible example of the individual puffs of air that are released in each vibratory cycle of the vocal folds.

One characteristic of voice quality is related to the way in which some speakers initiate the vibrations. Efficient use of the voice requires phonation to be initiated from lightly adducted or approximated folds. Some speakers initiate phonation with what is called a *glottal attack* (or sometimes harsh glottal attack) for which the vocal folds are tightly adducted prior to vibration. This causes their eventual separa-

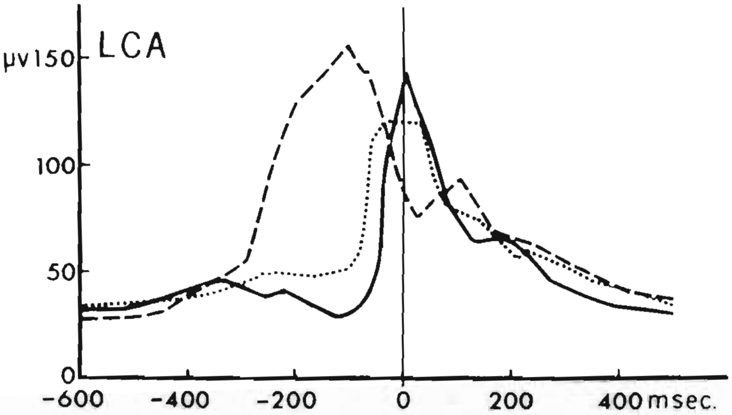


Figure 4.48. Contrasting muscle activity patterns in the lateral cricoarytenoid muscle (*LCA*) for various forms of vocal attack. The onset of the vowel is marked as *O*. Activity is earliest for glottal attack (*dashed line*) later for normal voiced attack (*dotted line*) and latest for voiceless aspirate attack (*solid line*). The measure of muscle activity is obtained by smoothing muscle interference patterns. (See Fig. 4.13 and the discussion of Fig. 7.24). (Reprinted with permission from H. Hirose and T. Gay: *Folia Phoniatrica (Basel). 25*, S. Karger AG, Basel, 1973.)

tion to release the air burst of a stop consonant, similar to those of /p/ or /t/ but produced at the glottis rather than at the lips or the alveolar ridge. Hirose and Gay have shown that glottal attack is accompanied by an early onset of activity in the lateral cricoarytenoid muscles (Fig. 4.48) which compress the center of the vocal folds. This early compression provides the time needed for the pressure build-up required for the stop-like release burst mentioned above.

The symbol for this glottal stop is [ʔ]. Thus, we transcribe a vowel or diphthong initiated by a glottal stop as [ʔaɪ] or [ʔa], rather than with the simple symbols usually associated with them. Harry Belafonte, a popular singer in the United States, made repeated visits to the hospital during the height of his career to have vocal nodules removed from his vocal folds, nodules caused by glottal attack. This is an affliction suffered by many singers, teachers, and others who make heavy use of their voices at high intensity levels.

Relationship between Frequency and Intensity

We have seen that by increasing the subglottal air pressure, keeping other things constant, we can increase vocal intensity. However, if subglottal pressure is increased without muscular adjustments of the vocal folds, the fundamental frequency as well as the intensity will increase. If someone is phonating a steady tone and is (gently) punched in the stomach, the tone not only gets louder but increases in pitch. The pitch rise may be caused by reflexive tensing of the vocal folds or be due to the fact that the increased subglottal air pressure causes the vocal fold closure to occur more quickly because of the Bernoulli effect. In contrast, when one is speaking, at the end of a breath, the f_0 drops naturally along with the intensity by about 2–7 Hz per cm of H_2O decrease. A singer or a speaker can reverse this affinity, however. If a singer wants to increase intensity but maintain the f_0, he must lower the resistance to the airflow at the vocal folds, either by relaxing the cricothyroid a bit or by lowering the internal tension by relaxing the thyroarytenoid muscle. Similarly, when asking "Are you sure?," in order to signal the question with a rising fundamental frequency, the speaker must work against the natural fall in frequency at the end of a breath group by increasing cricothyroid activity, stretching the folds, and at the same time, increasing internal intercostal muscle activity to give added stress to the word "sure."

Increased vocal intensity is due to greater resistance (afforded by the vocal folds) against

Table 4.3. Summary Chart of Events during Phonation*

Peripheral nerves	Muscles	Movements	Air pressure	Air movement
Xth cranial n. (vagus) Recurrent branch	PCA	Open vocal folds before thorax enlargement →	→	Air enters via larynx to lungs
→	IA →	Adduction of vocal folds		
→	LCA →	Medial compression of vocal folds	P_s builds Pressure drops across glottis,	
→	Voc. →	Intrinsic tension	$P_{sub} > P_{supra}$	
Xth cranial n. (vagus) Ext. branch of sup. laryngeal n. →	CT →	Longitudinal tension	Resistance offered to P_s by vocal fold tension	
		Vocal folds blown open →	Subglottal air pressure overcomes vocal fold resistance	
		→	→	Released puff of air
		Folds sucked together ←	Negative pressure *vs.* lateral edges of folds as velocity of air increases (Bernoulli effect)	
			→	Airstream cut off
		Vocal folds part ←	P_{sub} builds again	
		→	→	Another puff of air released

* The abbreviations used in the table are: PCA, posterior cricoarytenoid muscle; IA, interarytenoid muscle; LCA, lateral cricoarytenoid muscle; Voc., vocalis muscle; CT, cricothyroid muscle.

the increased airflow. The vocal folds are blown wider apart, releasing a larger puff of air that sets up a sound pressure wave of greater amplitude. The vocal folds not only move farther apart for each vibratory cycle of increased intensity, but they stay adducted for a larger part of each cycle. In Figure 4.49, the changes in the vocal fold movement are schematized and presented with the resulting change in waveform.

Summary

We have seen that phonation is a dynamic process, varying as it does during running speech in intensity, frequency, and quality. The output is a rapidly varying acoustic stream made up of segments of silence, periodic sounds, and noises. The alternation between sequences of phonated and unphonated segments is particularly complex. As a result, normal speakers frequently simplify such sequences to eliminate the alternation. For example, we say "cats" [kæts] with a voiceless (unphonated) [s], but after a voiced stop, it is easier to continue phonating and change the [s] to a [z] as in "dogs" [dɔgz]. Indeed, such alterations are so general in English that they are usually considered to be rule-governed: The word-final morpheme {-s}, whether a plural, a verb ending, or a possessive marker, is usually pronounced [s] after most voiceless sounds and [z] after voiced sounds. Contrast the final [s]'s in "mops," "loafs," "Kit's" and "pecks" with the final [z]'s in "mobs," "loaves," "kid's" and "pegs." For

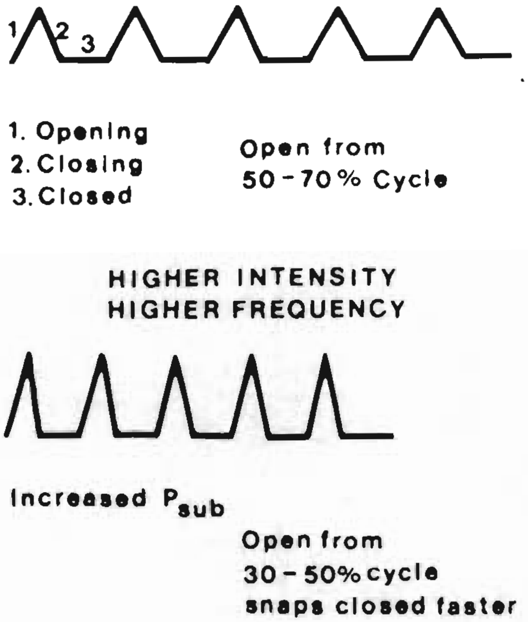


Figure 4.49. Schematic of the movement of the vocal folds during phonation. At higher subglottal pressure, the folds remain closed for a greater proportion of the vibratory cycle, and close more rapidly. Frequency tends to increase, as well as intensity.

those with certain speech disorders, it is particularly difficult to coordinate shifts back and forth between phonated and unphonated sound production. A person who stutters, for example, attempting to say his name, "Sam," might prolong the first sound, [sːæm], or repeat it [s-s-s-s-sæm], but he is not really stuttering the [s]. Rather, he produces the [s] very well, but fails to make a smooth transition to the voiced [æ].

Phonation must be coordinated with respiration. The motor commands to the larynx must be related to those of the respiratory system. To take a breath for speech, the glottis opens quickly before the thorax expands, and when the vocal folds adduct for voicing, the action is simultaneous with expiration. Table 4.3 summarizes the chain of events in vocal fold vibration from neural impulses to the results in air pressure and movements. For muscular adjustments the arrows move from left to right, while for aerodynamic forces the arrows move from right to left.

REFERENCES

General Works on Speech Production

Daniloff, R., Schuckers, G., and Feth, L., *The Physiology of Speech and Hearing*. Englewood Cliffs, NJ: Prentice-Hall Inc. 1980.

Dickson, D. R., and Maue-Dickson, W., *Anatomical and Physiological Bases of Speech*. Boston: Little, Brown and Co. 1982.

Kahane, D. R., and Folkins, J. F. *Atlas of Speech and Hearing Anatomy*. Columbus, OH: Charles E. Merrill, 1984.

Harris, K. S., Physiological Aspects of Articulatory Behavior. In *Current Trends in Linguistics*, Vol. 12, No. 4, T. A. Sebeok (Ed.) The Hague: Mouton, 1974, pp. 2281–2302.

Kent, R. D., Atal, B. S., and Miller, J. L. (Eds.), *Papers in Speech Communication: Speech Production*. Woodbury, NY: Acoustical Society of America, 1991.

Lieberman, P., *Speech Physiology and Acoustic Phonetics: An Introduction*. New York: Macmillan, 1977.

Minifie, F., Hixon, T. J., and Williams, F. (Eds.), *Normal Aspects of Speech, Hearing, and Language*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1972.

Perkell, J. S., *Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study*. Cambridge, MA: M. I. T. Press, 1969.

Van Riper, C., and Irwin, J. V., *Voice and Articulation*. Englewood Cliffs, NJ: Prentice-Hall Inc., 1958.

Zemlin, W. R., *Speech and Hearing Science: Anatomy and Physiology*, Ed. 2. Englewood Cliffs, NJ: Prentice-Hall Inc., 1981.

Neurophysiology References

Broca, P., Remarques sur la siege de la faculté du langage articule, suivies d'une observation d'aphémie (perte de la parole). *Bull. Soc. Anatom. Paris. VI: 36*, 1861, 330-357.

Eccles, J. C., *The Understanding of the Brain*. New York: McGraw-Hill, 1973.

Fromkin, V. A., Slips of the Tongue. *Sci. Am. 229*, 1973, 110–116.

Lassen, A. R., Ingvar, D. H., and Skinhoj, E., Brain Function and Blood Flow. *Sci. Am. 239*, 1978, 62–71.

MacKay, D. G., Spoonerisms: The Structure of Errors in the Serial Order of Speech. *Neuropsychologia 8*, 1970, 323–350.

Milner, B., Branch, C., and Rasmussen, T., Observations on Cerebral Dominance. In *Psychology Readings: Language*. R. C. Oldfield and J. C. Marshall (Eds.) Baltimore: Penguin Books, 1968, pp. 366–378 (Later figures given in present text from oral presentation by Milner at ASHA meeting, Las Vegas, 1974.)

Penfield, W., and Roberts, L., *Speech and Brain-Mechanisms*. Princeton, NJ: Princeton University Press, 1959.

Pribram, I
NJ: Pre
Wada, J., a
dium A
Domina
Neurost
Wernicke,
lau: Ma

Respiratio

Campbell,
Sci. 15
Draper, M
tory M
16–27.
Fenn, W.
1951, 7
Hixon, T.
Aspects
T. J. Hi
Prentic
Hixon, T.
the Ch
the Th
Speech
Hixon,
San D
135–19
297–35
Lofqvist,
tion. In
C. Sasa
York: T
Mead, J., I
erating
1968, 1
Netsell, R
the Int
1969, 6
Rahn, H.,
The Pr
Am. J.

Pribram, K. H., *Languages of the Brain*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1971.

Wada, J., and Rasmussen, T., Intracarotid Injection of Sodium Amytal for the Lateralization of Cerebral Speech Dominance: Experiments and Clinical Observations. *J. Neurosurg. 17*, 1960, 266–282.

Wernicke, C., *Der Aphasische Symptomencomplex*. Breslau: Max Cohn and Weigert, 1874.

Respiration References

Campbell, E., The Respiratory Muscles. *Ann. N. Y. Acad. Sci. 155*, 1968, 135–140.

Draper, M. H., Ladefoged, P., and Whitteridge, D., Respiratory Muscles in Speech. *J. Speech Hear. Res. 2*, 1959, 16–27. Reprinted in Kent *et al.*, 1991, 3–14.

Fenn, W. O., Mechanics of Respiration. *Am. J. Med. 10*, 1951, 77–91.

Hixon, T. J., Respiratory Function in Speech. In *Normal Aspects of Speech, Hearing, and Language*. F. D. Minifie, T. J. Hixon, and F. Williams (Eds.) Englewood Cliffs, NJ: Prentice-Hall, Inc., 1973.

Hixon, T. J., Mead, J., and Goldman, M. D., Dynamics of the Chest Wall during Speech Production: Function of the Thorax, Rib Cage, Diaphragm, and Abdomen. *J. Speech Hear. Res. 19*, 1976, 297–356. Reprinted in Hixon, T. J. *Respiratory Function in Speech and Song*. San Diego, CA: Singular Publishing Group, 1991, 135–197. Also reprinted in Kent *et al.*, 1991 (q.v.), 297–356.

Lofqvist, A., Aerodynamic Measurements for Vocal Function. In *Neurological Disorders of the Larynx*. A. Bletzer, C. Sasaki, S. Falin, M. Brill, and K. S. Harris. (Eds.), New York: Thieme, 1992.

Mead, J., Bouhuys, A., and Proctor, D. F., Mechanisms Generating Subglottic Pressure. *Ann. N. Y. Acad. Sci. 155*, 1968, 177–181.

Netsell, R., Subglottal and Intraoral Air Pressures during the Intervocalic Contrast of /t/ and /d/. *Phonetica 20*, 1969, 68–73.

Rahn, H., Otis, A. B., Chadwick, L. E., and Fenn, W. O., The Pressure-Volume Diagram of the Thorax and Lung. *Am. J. Physiol. 146*, 1946, 161–178.

Stetson, R., *Motor Phonetics*. Amsterdam: North-Holland, 1951.

van den Berg, J., Direct and Indirect Determination of the Mean Subglottic Pressure. *Folio Phoniatr. (Basel). 8*, 1956, 1–24.

Phonation References

Atkinson, J. E., Correlation Analysis of the Physiological Factors Controlling Fundamental Voice Frequency. *J. Acoust. Soc. Am. 63*, 1978, 211–222.

Faaborg-Andersen, K., Electromyographic Investigation of Intrinsic Laryngeal Muscles in Humans. *Acta. Physiol. Scand. 41*, Suppl. 140, 1957, 1–148.

Hirose, H., and Gay, T.. The Activity of the Intrinsic Laryngeal Muscles in Voicing Control. *Phonetica. 25*, 1972, 140–164.

Husson, R., *Étude des Phénomenes Physiologiques et Acoustiques Fondamentaux de la Voix Chantée*. Thesis, University of Paris, 1950.

Ishizaka, K., and Flanagan, J. L., Synthesis of Voiced Sounds from a Two-Mass Model of the Vocal Cords. *Bell System Technical Journal. 51*, 1233–1268. Reprinted in Kent *et al.*, 1991 (q.v.), 183–218.

Ludlow, C. L., and Hart, M. O. (Eds.), *Proceedings of the Conference on the Assessment of Vocal Pathology, ASHA Reports*, 1981.

Müller, J., *The Physiology of the Senses, Voice, and Muscular Motion with the Mental Faculties*. Translated by W. Baly. London: Walton and Maberly, 1848.

Negus, V. E., *The Comparative Anatomy and Physiology of the Larynx*. New York: Hafner Publishing Co., 1962. (A rewriting of V. E. Negus, *The Mechanism of the Larynx*. London: William Heinemann Medical Books, Ltd., 1928.)

Shipp, T., Vertical Laryngeal Position during Continuous and Discrete Vocal Frequency Change. *J. Speech Hear. Res. 18*, 1975, 707–718.

Sundberg, J. *The Science of the Singing Voice*. De Kalb, IL: Northern Illinois Press, 1987.

Titze, I. R., On the Mechanics of Vocal-Fold Vibration. *J. Acoust. Soc. Am. 60*, 1976, 1366–1380.

van den Berg, I., Myoelastic-Aerodynamic Theory of Voice Production. *J. Speech Hear. Res.* 1, 1958, 227–244.

Von Helmholtz, H., *Die Lehre der Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. Braunschweig: F. Vieweg und sohn, 1863.