

# The Human Instrument

When judged by its size, our vocal system fails to impress as a musical instrument. How then can singers produce all those remarkable sounds? **By Ingo R. Titze**

## KEY CONCEPTS

- Although the human vocal system is small, it manages to create sounds as varied and beautiful as those produced by a variety of musical instruments.
- All instruments have a sound source, a resonator that reinforces the basic sound and a radiator that transmits the sound to listeners.
- A human's sound source is the vibrating vocal folds of the larynx; the resonator is the sound-boosting airway above the larynx; and the radiator is the opening at the mouth.
- The human voice can create such an impressive array of sounds because it relies on nonlinear effects, in which small inputs yield surprisingly large outputs.

—The Editors

The human vocal system would not receive much acclaim if instrument makers placed it in a lineup of traditional orchestral instruments. Arranged by size, for example, the voice box (larynx)—and the airway it sits in—would be grouped with the piccolo, among the smallest of mechanical music makers. And yet experienced singers compete well with all man-made instruments, one on one and even paired with full orchestras. Recent investigations of how our singing voice generates a remarkable range of sounds have revealed surprising complexity in the behavior of the vocal system's elements and in the ways they interact.

For more than half a century, scientists explained the voice's ability to create song by invoking a so-called linear theory of speech acoustics, whereby the source of sound and the resonator of sound (or amplifier) work independently [see "The Acoustics of the Singing Voice," by Johan Sundberg; *SCIENTIFIC AMERICAN*, March 1977]. Researchers have now learned, however, that nonlinear interactions—those in which source and resonator feed off each other—play an unexpectedly crucial role in generating human sound. Such insights now make it possible

to describe how great singers produce those amazing sounds.

## Music-Making Keys

Structural and operational shortcomings in the human vocal apparatus are apparent in all its parts. To make music, an instrument needs three basic components: a sound source that vibrates in the air to generate a frequency that we perceive as pitch, together with higher frequencies that define the timbre (sound color); one or more resonators that reinforce the fundamental frequency by increasing its vibration strength; and a radiating surface or orifice that transfers the sounds to free air space and, eventually, to a listener's ear.

In the case of, say, a trumpet, a player's lips vibrate as lung-pumped air rushes between them into a cup-shaped mouthpiece to create a fundamental frequency and several higher frequencies that are called overtones. The instrument's metal tubes serve as the resonators, and the expanding aperture of the horn radiates the sound. Trumpeters alter fundamental frequency by modifying the lip tension and by pressing the valves to change the effective length of the tubes. Or take a violin: the strings vibrate to create



AARON GOODMAN

**AMAZINGLY FLEXIBLE**, the human voice creates sounds as rich and complex as those of conventional musical instruments—but with much smaller equipment.

© 2007 SCIENTIFIC AMERICAN, INC.

itches, the central air cavity and wooden top supply resonance, and the *f* holes in the top plate help to send the sound into the surrounding air.

A singer, on the other hand, relies on vibrating vocal folds, blowing air across them to generate the sound frequencies. Vocal folds are two small bundles of specialized tissue, sometimes called “vocal cords,” that protrude pouchlike from the walls of the larynx. They generate a fundamental frequency by rapidly oscillating as they contact each other, separate and come in contact again. The glottis (the space between the folds) opens and closes. The laryngeal vestibule, an airway passage just above the larynx, acts like the mouthpiece of the trumpet to couple the sound to the remaining part of the resonator known as the vocal tract. The lips radiate the sound outward like the bell of the trumpet.

Instrument manufacturers examining the vocal folds, which are collectively the size of your thumbnail, would not find their potential for making orchestral musical sounds impressive. Beyond their small size, one immediate objection would be that they would seem too soft and spongy to sustain vibration and create a variety of pitches.

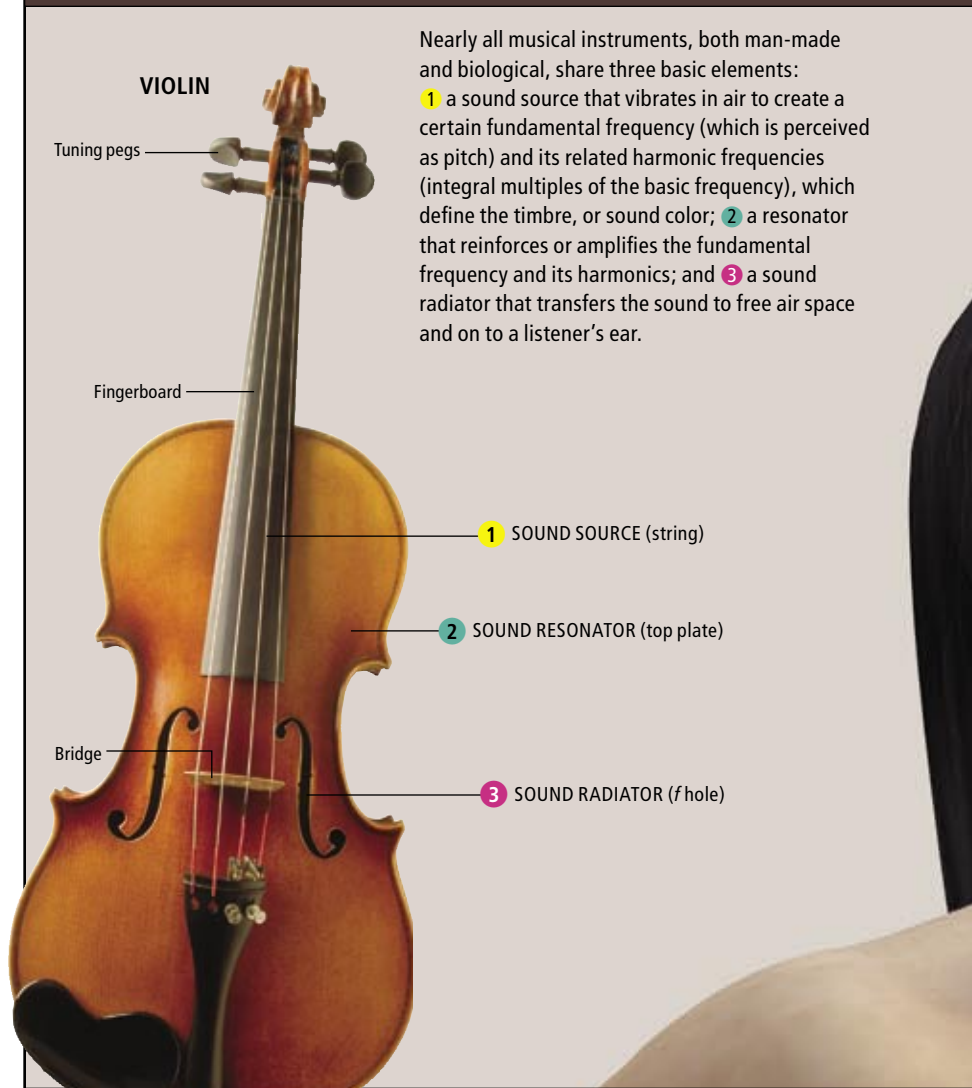
Nature, biology’s instrument maker, might respond that although the folds are certainly undersized, the airways can produce enough resonance to reinforce the sound of the larynx substantially. But here, too, the musical instrument maker would probably still fail to be persuaded: the typical air tube extends just 15 to 20 centimeters above the larynx and 12 to 15 centimeters below it, no more than the length of a piccolo. The rest of the body contributes little or nothing. Wind instruments that approximate the pitches created by the human voice (trombones, trumpets, bassoons) typically contain much longer tubes; the bell and valves of a trumpet, for instance, uncoil to about two meters and those of a trombone to about three meters.

### Source Design

To understand how nature the instrument maker has developed the vocal folds that perform beyond expectations, first consider some standard requirements for sound sources. For a reed or string to sustain its vibration, it needs to be made of an appropriately elastic material so it can snap back when deformed. Elasticity is measured by its stiffness (or, conversely, flexibility) or its tension: a reed has a bending stiffness; a string vibrates under tension. Generally, a sound source’s stiffness or tension determines the sound

## [THE BASICS]

# HOW INSTRUMENTS MAKE MUSIC



Nearly all musical instruments, both man-made and biological, share three basic elements: **1** a sound source that vibrates in air to create a certain fundamental frequency (which is perceived as pitch) and its related harmonic frequencies (integral multiples of the basic frequency), which define the timbre, or sound color; **2** a resonator that reinforces or amplifies the fundamental frequency and its harmonics; and **3** a sound radiator that transfers the sound to free air space and on to a listener’s ear.

## LINEAR VS. NONLINEAR

**Voice scientists used to explain the performance of the human vocal system in terms of linear effects, ones in which the outputs of a function are proportional to the inputs (and so can be represented as a line, for example). More recently, researchers have concluded that the human vocal system behaves nonlinearly. In a nonlinear feedback system, small changes can result in disproportionately large effects.**

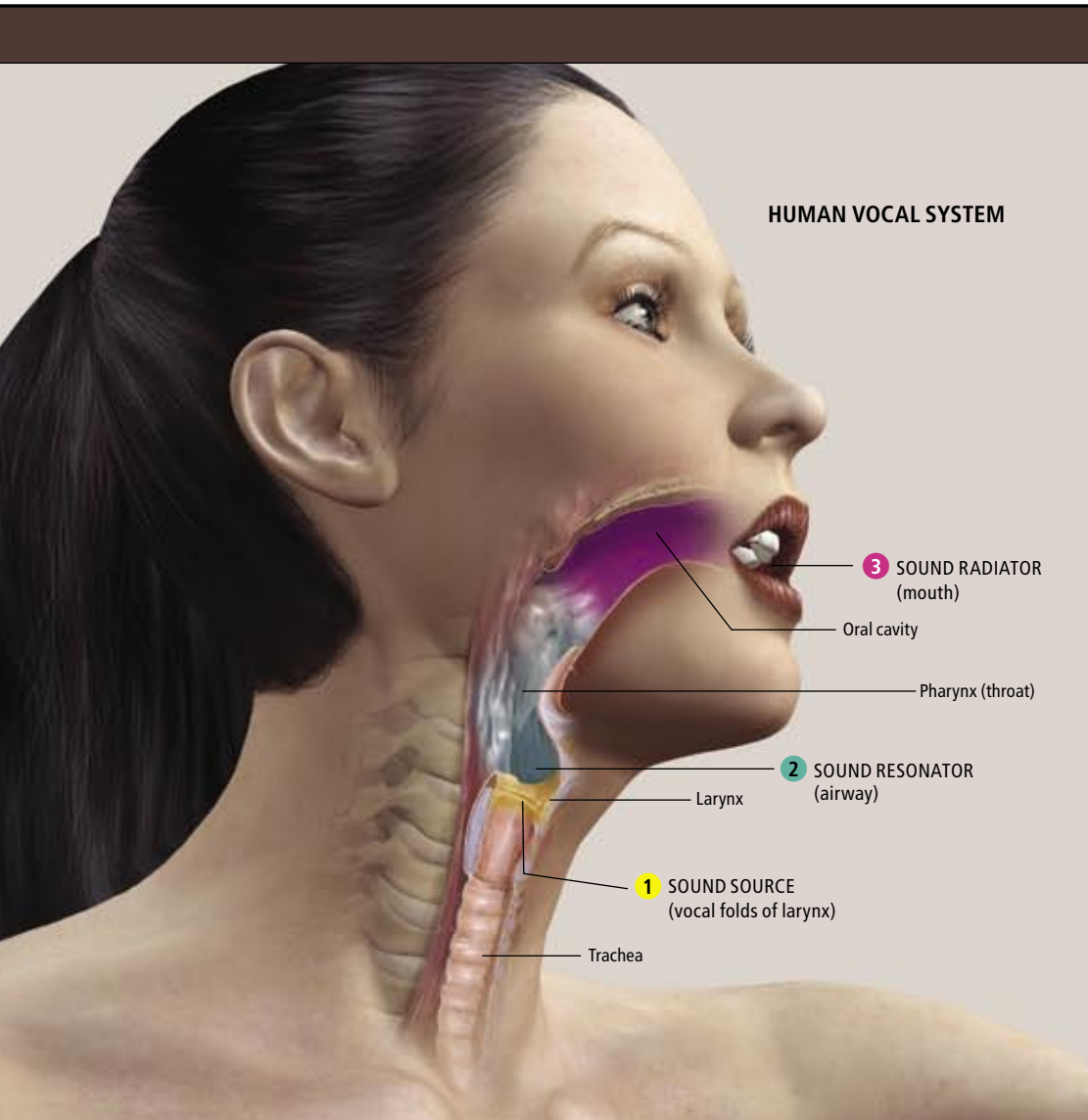
frequencies via a square root relation. Thus, to make a steel string of a given length double its frequency (raise the pitch by an octave), one must quadruple the string tension. This rather stringent requirement may limit the range of frequencies that can reasonably be obtained by altering a source’s stiffness or tension.

Fortunately, a player can also change the frequency of a sound source’s vibration by effectively lengthening or shrinking the oscillating element. Within a vibrating string, for instance, frequencies are inversely proportional to the length of the vibrating segment. By pinning the string on one end with the finger, a player selects different frequencies. If a string’s vibrating length is cut in half without changing the tension, for example, the vibration frequency doubles. To produce a wider range of frequencies, a single

[THE AUTHOR]



Ingo R. Titze, a world leader in the scientific study of the human voice, has published more than 500 articles on the topic. He currently serves as the University of Iowa Foundation Distinguished Professor in that institution's department of speech pathology and audiology, and he directs the National Center for Voice and Speech ([www.ncvs.org](http://www.ncvs.org)) at the Denver Center for the Performing Arts. Titze, who received his Ph.D. in physics from Brigham Young University in 1972, teaches singing and sings in multiple styles, including opera, Broadway and pop.



HUMAN VOCAL SYSTEM

3 SOUND RADIATOR (mouth)

Oral cavity

Pharynx (throat)

2 SOUND RESONATOR (airway)

Larynx

1 SOUND SOURCE (vocal folds of larynx)

Trachea



**SHRIEK ROCKER** Steven Tyler is celebrated for his ability to scream tunefully. Aerosmith's lead singer creates that extreme sound by using a large amount of nonlinear effects in his vocalizations.

musical instrument often uses multiple strings.

String instruments, then, have three distinct mechanisms for changing frequency: altering the length of a string, modifying its tension, or skipping to another string. Players of stringed instruments typically set the tensions by turning pegs around which the strings are wrapped; the strings retain this same tension between end points. Players almost never can manipulate both the length and tension simultaneously.

**The Little Source That Could**

In playing the human vocal folds, in contrast, singers must do what no other string instrument can do: vary the length and tension of the vibrating material simultaneously to change frequency. Rather than pinning down a vocal fold with a finger to shorten its effective length, we use mus-

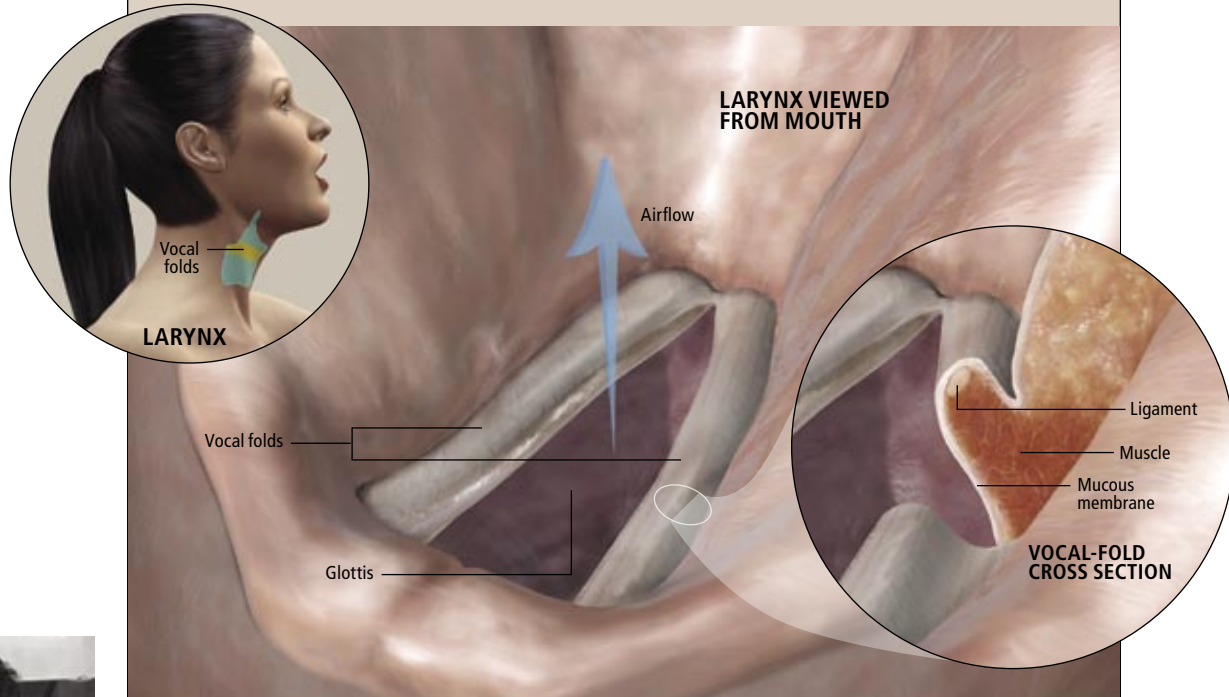
cles to shift its end points. But should we lengthen or shorten the vocal folds to raise frequency? An argument could be made for either adjustment. Longer vocal folds would vibrate at lower frequency, but tenser folds would vibrate at higher frequency.

The physics formula describing the frequency of a string fixed at both ends under tension says that to get the maximum increase in frequency, one should increase the tension (actually the tensile stress, or tension per cross-sectional area) while decreasing the length. Such a response requires an unusual material, because most materials can increase tension (stress) only when they are elongated. Think of a rubber band; pull on it, and it tightens up. Thus, length and tension are in competition for changing frequency.

COURTESY OF INGO R. TITZE (Titze); GEORGE PIMENTEL (WireImage/Getty Images (Tyler); ADAM QUESTELL (Illustration)

## HOW OUR VOCAL FOLDS WORK

Unlike violin strings, the human sound source—the vocal folds (or vocal cords) in the larynx—has a complex, three-part structure that allows us to generate several octaves of frequencies. At the core of each fold is a stringlike ligament (*cross section*). Internal to the ligament are contractile muscles. And covering it all is a highly flexible mucous membrane. Each component brings a special capability to the whole. The ligament's tensile stress rises rapidly with elongation (by muscles that move the cartilages attached to the folds), which helps to produce higher frequencies. The vocal-fold muscle can increase tensile stress as it contracts. In doing so, it generates an even larger frequency range. The soft, pliable surface of the outer membrane, which oscillates like a flag in the wind as air from the lungs passes over it, exchanges vibrational energy with the airstream to create sound waves.



**BROADWAY STAR** Ethel Merman belted out songs with precise enunciation and pitch so the audience could hear her even without amplification. The female “belt” voice gains resonance from vocal-tract inertive reactance [see box on opposite page], which boosts the second harmonic frequency (double the fundamental).

Nature has addressed these problems by constructing the vocal folds out of a three-part material that displays properties not found in standard instrument strings. One component is a ligament that looks somewhat stringlike, which is why the folds came to be called “cords” popularly. Scientists have shown in biomechanical tests that the stress in this ligament rises nonlinearly when it is stretched just a little; it can be virtually limp when short but impressively tense when elongated. Stretching its length from 1.0 to just 1.6 centimeters, for example, can raise its internal stress by a factor of 30, which would yield a frequency change ratio of more than 5 to 1 (recall the square root relation mentioned earlier). But the fact that the length increases by 60 percent lowers the vibration rate, bringing the true frequency ratio back to around 3 to 1, about one and a half octaves in musical terms. Most of us speak and sing in this frequency range, but some singers can produce as much as four to five octaves, which is still considered extraordinary by scientists.

### Complex Cords

Biology has also found a second way to expand the pitch range of the vocal folds, including a material that can increase in tension as it *shortens*, namely, muscle tissue. The internal contraction of muscle fibers can raise the stress between a vocal fold's end points, even when the fold itself shortens. About 90 percent of the volume of the vocal fold is muscle tissue. In essence, nature has solved the pitch problem largely by growing a group of strings side by side as a laminate, with some layers having contractile properties and others not. But how can this complex tissue be kept in vibration when it cannot be bowed or repeatedly plucked inside the larynx? The only source of energy available to deform the folds and thereby induce vibration—the way that wind passing across a flag makes it flap—is air flowing from the lungs. A muscle and a ligament alone would be too stiff to develop such vibrations as air passes over their surfaces. For the needed air-driven oscil-

lation to occur, a soft, pliable surface tissue is required, one that can respond to the airstream by generating waves akin to those the wind forms at the surface of the ocean [see “The Human Voice,” by Robert T. Sataloff; *SCIENTIFIC AMERICAN*, December 1992].

And indeed, the folds have a third layer, a mucous membrane that stretches over the muscle-ligament combination to provide this energy-transfer function. This mucosa, which consists of a very thin skin (epithelium) with a fluidlike substance underneath, is easily deformed and can support a so-called surface wave. My colleagues and I have shown mathematically that this airstream-driven wave sustains vibration. The buckling, ribbonlike motion often makes the tissue look like it is folding bottom to top, which is how the name “vocal fold” arose.

### Playing the Vocal Folds

How can this triple-decker system be played over several octaves so as to produce a single frequency? Only with much experience and dexterity. Chaotic effects always lurk in the background during vocalizations as multiple

natural (freely vibrating) frequencies compete in these tissues for dominance. This competition may result in unexpected pitch jumps or roughness in the sound [see “The Throat Singers of Tuva,” by Levin Edgerton; *SCIENTIFIC AMERICAN*, September 1999].

For low pitches and moderate-to-loud sound volumes, the singer activates the vocal fold muscle and sets all the layers into vibration. The vocal folds are short, and the muscle stress largely determines the pitch. In this case, the mucosa and the ligament are both relaxed and serve mainly to propagate the desired surface waves for self-sustained oscillation. For the sound volume to be reduced at these pitches, the muscle does not vibrate and is used only to adjust the vocal-fold length. It is the combined elasticity of the mucosa and ligament that determines the frequency. To create high pitches, the singer elongates the vocal fold; ligament stress alone then dictates the frequency while the mucosa carries the surface wave.

It is not hard to imagine the complicated control system and innervation of the laryngeal muscles needed to finely regulate these tensions

### DID YOU KNOW?

Humans tend to think of the entire body as the human instrument, which would make it comparable in size to a double bass. But most of the human body contributes nothing to the sound—neither the chest, the back, the belly, the buttocks nor the legs. All the sound comes from the voice box (larynx) and the air passages.

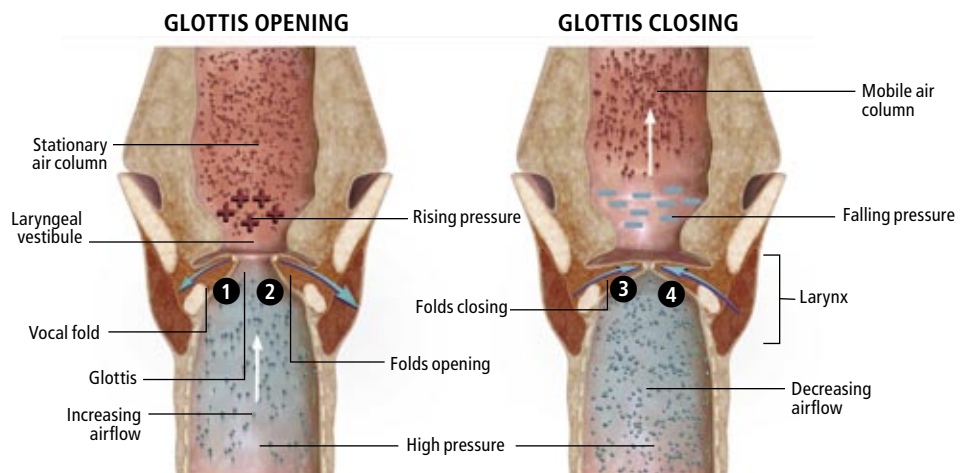
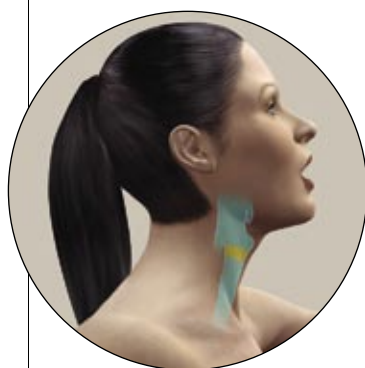
#### [HUMAN SOUND RESONATOR]

## HOW THE VOCAL AIRWAYS AMPLIFY SOUND

Singers use a nonlinear energy-feedback process in the laryngeal vestibule (the airway above the larynx) to resonate, or amplify, sounds created by the vocal folds. This process, called inertive reactance, occurs when singers create special conditions in the vestibule to provide an extra, precisely timed “kick” to each cyclic opening and closing of the folds that reinforces their vibration to create stronger sound waves.

The kick comes in when the motion of the air column in the vestibule lags with respect to the movement of the vocal folds. When the vocal folds start to separate at the beginning of an oscillation (1), lung-driven air flows into the glottal space between the folds and

presses against the motionless air column in the vestibule. The inertia of the stationary air column raises the air pressure in the glottis, which drives the folds even farther apart (2). Then the lungs begin to accelerate the air mass upward. As the air column moves, the elastic recoil of the folds begins to send them back together to close the glottis, cutting off the airflow from the lungs (3). Those responses leave a partial vacuum in the glottis that acts to slam the folds together strongly (4). Thus, like a well-timed push on a kid’s swing, the inertive reactance—the push-pull action—of the air in the laryngeal vestibule augments each swing of the vocal folds, creating resonance.



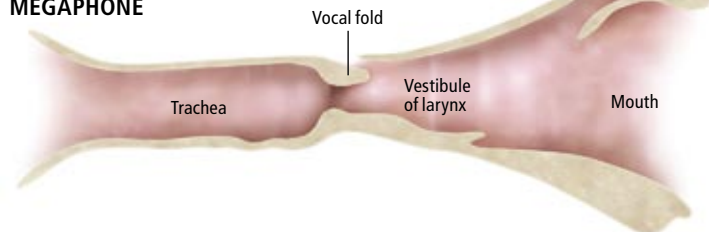
ADAM QUESTEIL

# BIG MOUTHS AND SMALL MOUTHS

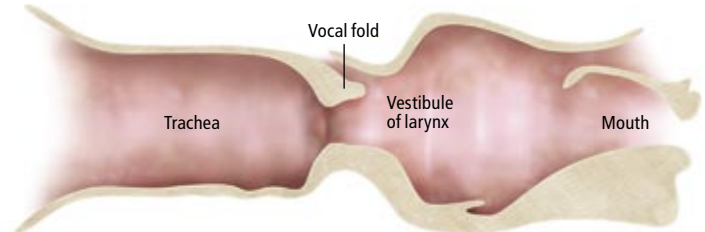
Acting as a tube resonator, the vocal tract adopts certain shapes to better project certain pitches and resonant overtones. To produce powerful high notes, belters often open their mouths as wide as possible. This so-called megaphone configuration resembles a trumpet (*top*), with the vocal folds and vestibule serving as the “lips and mouthpiece” and the mouth as the “horn.” Other singing styles are better produced when the vocal tract takes on an inverted megaphone shape—that is, with the mouth narrowed (*bottom*).



### MEGAPHONE



### INVERTED MEGAPHONE



to produce a desired frequency and volume level. Laryngeal muscles outside the vocal folds precisely coordinate length changes in the vocal fold. During these complicated manipulations, voice quality may suddenly change, a phenomenon known as registration. It is caused to a large extent by overusing or underusing the vocal-fold muscle to regulate tension. Singers use registration artistically to present two contrasting sounds to the listener, as in yodeling. If a singer involuntarily or accidentally changes register, however, it can cause embarrassment, because such a slip suggests a lack of control of the singer's instrument.

## Resonating Airways

In musical instruments, the resonator for the most part determines the instrument size, but singers have to make do with a pint-size resonator. The human resonator, however, performs effectively despite its overt limitations.

In a musical instrument, boards, plates, kettles, horns or tubes typically act to reinforce and amplify the frequencies that the sound source produces. In the violin, for instance, the strings pass over a bridge support that connects to the top plate, which has been carefully fashioned to vibrate sympathetically at many of the same natural frequencies that the strings can produce, thereby boosting them. The air mass between the top and bottom plates can also oscillate at the strings' natural frequencies. In

many brass and woodwind instruments, the horn (with its valves) is designed to match many of the source frequencies at whatever pitch is played.

Because physical law dictates that all steady (continuous) sounds are composed of source frequencies that are harmonically spaced—meaning that all source frequencies are integer multiples (2:1, 3:1, 4:1, ...) of the fundamental—the resonator must often be quite large to accommodate these wide frequency spacings. This physical law dictates that trumpet horns are 1.2 to two meters long, trombone horns stretch three to nine meters, and French horn tubing uncoils from 3.7 to 5.2 meters.

Nature is stingy with the size of the singer's resonator. The total size of the human airway above the vocal folds is only about 17 centimeters long. The lowest frequency that can be resonated is around 500 hertz (cycles per second)—and half that when certain vowels are sung, such as /u/ or /i/ (as in “pool” or “feel”). Because the vocal tract is a resonant tube that is nearly closed at one end, its resonant frequencies include only the odd-integer multiples (1, 3, 5, ...) of the lowest resonance frequency. Therefore, this short tube can resonate simultaneously only the odd harmonics of a 500-Hz source frequency (500 Hz, 1,500 Hz, 3,500 Hz, ...). And because the vocal tract cannot change tube length with valves or slides (other than a few centimeters by protruding a lip or lowering the



**DAME Joan Sutherland knew instinctively that some vowels cannot be used when singing certain pitches. The Australian soprano changed some of the vowels in her opera lyrics (even to the point of mispronouncing the words) to best match them with the desired pitches.**

larynx), our resonator seems as if it should be hopelessly restricted in what it can do.

## Resonating a Short Tube

Here again recent studies indicate that nonlinear effects come to the rescue. This time it is a nonlinear interaction among the system's elements. Rather than reinforcing each harmonic with a specific tube resonance (as occurs, for example, in organ pipes of different sizes, each of which resonates certain harmonics), our short vocal tract reinforces a cluster of harmonics simultaneously by using an energy-feedback process. The vocal tract can store acoustic energy in one part of the vibration cycle and feed it back to the source at another, more advantageous time. In effect, the vocal tract gives a "kick" to each cycle of oscillation of the vocal folds so as to increase the amplitude of vibratory motions. In analogy to pushing someone on a playground swing, this cyclic kick resembles a carefully timed push to boost the amplitude (travel distance) of the swing's oscillations.

The ideal timing of the kick comes when the movement of the air column in the tube is delayed with respect to the movement of the vocal folds. Scientists say that the air column then has inertive reactance (slow or sluggish response to an applied pressure). Inertive reactance helps to sustain the flow-induced oscillation of the vocal folds in a profound way [see box on page 99].

When the vocal folds begin moving apart at the inception of a vibratory cycle, air from the lungs starts to flow into the glottal space between them and begins pushing on the stationary air column located just above in the laryngeal vestibule. Air pressure in and above the glottis rises as the air column accelerates upward to allow new air to fill in behind it. This pressure increase pushes the folds even farther apart. When elastic recoil springs the folds back from the walls to close the glottis, the flow of air through the glottis subsides. Because of inertia, though, the air column continues to move up, leaving a partial vacuum in and above the glottis that acts to slam the folds more strongly together. Thus, like a well-timed push on a kid's swing, the inertive reactance of the air in the vocal tract augments each swing of the vocal folds with a push-pull action.

Still, the vocal tract does not automatically behave in this inertive way for all vocal shapes. A singer's task is to adjust the shape of the vocal tract (by carefully selecting favorable "singing" vowels) so that inertive reactance is experienced over most of the pitch range—no easy task.



**ITALIAN TENOR Luciano Pavarotti, famed for the brilliance and beauty of his tone, produced the rich resonance in his voice by fine-tuning nonlinear inertive reactance in his throat.**

## MORE TO EXPLORE

### The Physics of Small-Amplitude Oscillation of the Vocal Folds.

I. R. Titze in *Journal of the Acoustical Society of America*, Vol. 83, No. 4, pages 1536–1552; 1988.

### Acoustic Systems in Biology.

Neville H. Fletcher. Oxford University Press, 1992.

### Vocal Tract Area Functions from Magnetic Resonance Imaging.

B. Story, I. Titze and E. Hoffman in *Journal of the Acoustical Society of America*, Vol. 100, No. 1, pages 537–554; 1996.

### Principles of Voice Production.

Reprint. I. R. Titze. National Center for Voice and Speech, 2000. [www.ncvs.org](http://www.ncvs.org)

### The Physics of Musical Instruments.

Second edition (corrected fifth printing). N. H. Fletcher and T. D. Rossing. Springer, 2005.

## Megaphone Mouth

Different singing styles rely on different vocal tract shapes to make optimal use of inertive reactance. In producing an /æ/ vowel (as in "mad"), the vocal tract approximates a megaphone shape. A small cross section at the glottis is paired with a large opening at the mouth [see box on opposite page]. Singers can find inertive reactance as high as 800 or 900 Hz for males and 20 percent higher for females. At least two harmonic source frequencies can achieve inertive reactance for fairly high pitches, and several more can for low pitches. This fact means that one strategy for obtaining powerful high notes is for the singer to open the mouth as wide as possible, as in belting or calling. When the vocal tract adopts this megaphone configuration, it approximates the shape of an amputated trumpet (with no coiling tube or valves, but with a bell, or horn).

An alternative approach to reinforcing vocal-fold vibration with inertive reactance is to adopt the so-called inverted-megaphone shape, in which the laryngeal vestibule, the "mouthpiece," is kept narrow, the pharynx (the part of the throat situated immediately behind the mouth and nasal cavity) is expanded as widely as possible and the mouth is narrowed somewhat. This configuration is approximated to verbalize the /u/ vowel (as in "took"). The inverted-megaphone technique is ideal for female classical singers who wish to sing in the middle of their pitch range and male classical singers who wish to sing in the high part of their pitch range. Classical training involves finding more regions of the singing range where the vocal tract provides inertive reactance for the source frequencies, at all pitches and for many different vowels. The training also involves getting a "ring" into the voice, which is accomplished by the combination of the narrow vestibule and the wide pharynx. Singing teachers use terms such as "covering" the voice or "turning it over" to describe the process of choosing just the right vowel for the given pitch so that most of the source frequencies experience inertive reactance.

Singing styles are based on what human biology can offer to produce an acoustically efficient instrument. Researchers who study the elements of the human vocal system, and the unexpected ways in which it functions, are garnering an ever greater understanding of how accomplished singers ply their art. Both scientists and singers will benefit substantially from continued close cooperation and study. ■