# The Acoustics of the Singing Voice

*The voice organ is an instrument consisting of a power supply (the lungs), an oscillator (the vocal folds) and a resonator (the larynx, pharynx and mouth). Singers adjust the resonator in special ways*

by Johan Sundberg

Clearly there is something quite unusual about the voice of a first-class opera singer. Quite apart from the music, the intrinsic quality of such a voice can have a forceful impact on the listener. Moreover, a well-trained singer produces sounds that can be heard distinctly in a large opera house even over a high level of sound from the orchestra, and can do so week after week, year after year. If a second-rate singer or a completely untrained one tried to be heard over an orchestra, the result would be a scream and the singer's voice would soon fail. Is it only training that makes the difference? Or is the instrument that produces an excellent singer's voice itself different from other people's?

Let us begin with a description of that instrument. The voice organ includes the lungs, the larynx, the pharynx, the nose and the mouth. The main voice function of the lungs is to produce an excess of air pressure, thereby generating an airstream. The air passes through the glottis, a space at the base of the larynx between the two vocal folds (which are often called the vocal cords but are actually elastic infoldings of the mucous membrane lining the larynx). The front end of each vocal fold is attached to the thyroid cartilage, or Adam's apple. The back end of each is attached to one of the two small arytenoid cartilages, which are mobile, moving to separate the folds (for breathing), to bring them together and to stretch them. The vocal folds have a function apart from that of producing sound: they protect the lungs from any small objects entrained in the inspired airstream. Just above the vocal folds are the two "false" vocal folds, which are engaged when someone holds his breath with an overpressure of air in the lungs. The vocal folds are at the bottom of the tube-shaped larynx, which fits into the pharynx, the wider cavity that leads from the mouth to the esophagus. The roof of the pharynx is the velum, or soft palate, which in turn is the door to the nasal cavity. When the velum is in its raised position (which is to say during the sounding of all vowels except the nasalized ones), the passage to the nose is closed and air moves out through the mouth.

The larynx, the pharynx and the mouth together constitute the vocal tract, a resonant chamber something like the tube of a horn or the body of a violin. The shape of the tract is determined by the positions of the articulators: the lips, the jaw, the tongue and the larynx. Movements of the lips, jaw and tongue constrict or dilate the vocal tract at certain sites; protruding the lips or lowering the larynx increases the length of the tract.

Now consider the voice organ as a generator of voiced sounds. Functionally the organ has three major units: a power supply (the lungs), an oscillator (the vocal folds) and a resonator (the vocal tract). With the glottis closed and an airstream issuing from the lungs, the excess pressure below the glottis forces the vocal folds apart; the air passing between the folds generates a Bernoulli force that, along with the mechanical properties of the folds, almost immediately closes the glottis. The pressure differential builds up again, forcing the vocal folds apart again. The cycle of opening and closing, in which the vocal folds act somewhat like the vibrating lips of a brass-instrument player, feeds a train of air pulses into the vocal tract. The frequency of the vibration is determined by the air pressure in the lungs and by the vocal folds' mechanical properties, which are regulated by a large number of laryngeal muscles. In general the higher the lung pressure is and the thinner and more stretched the vocal folds are, the higher is the frequency at which the folds vibrate and emit air pulses. The train of pulses produces a rapidly oscillating air pressure in the vocal tract; in other words, a sound. Its pitch is a manifestation of the vibratory frequency. Most singers need to develop full control over a pitch range of two octaves or more, whereas for ordinary speech less than one octave suffices.

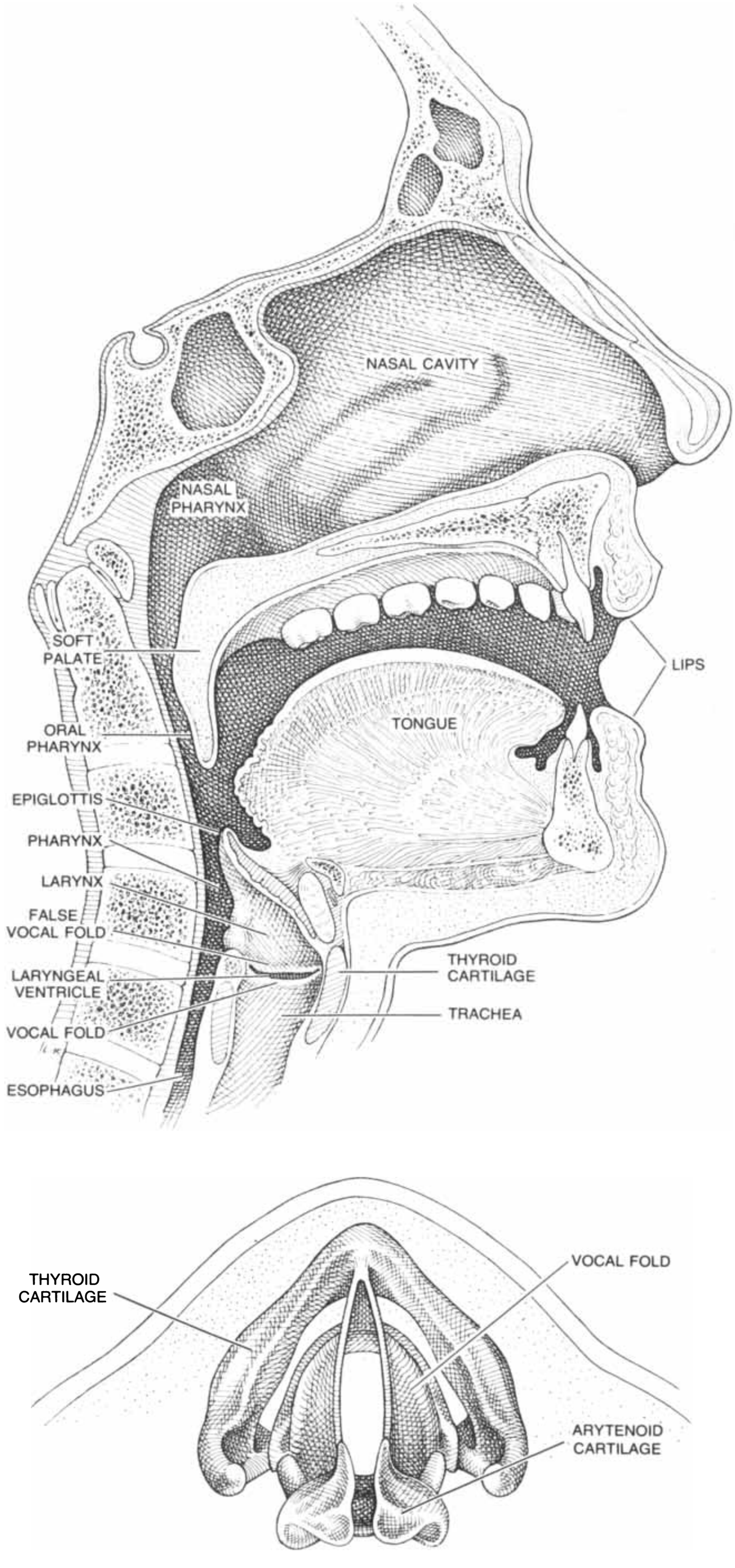
The sound generated by the airstream chopped by the vibrating vocal folds is called the voice source. It is in effect the raw material for speech or song. It is a complex tone composed of a fundamental frequency (determined by the vibratory frequency of the vocal folds) and a large number of higher harmonic partials, or overtones. The amplitude of the partials decreases uniformly with frequency at the rate of about 12 decibels per octave. The "source spectrum," or plot of amplitude against frequency, for a singer is not very different from that for a nonsinger, although the spectrum does tend to slope more steeply in soft speech than it does in soft singing.

The vocal tract is a resonator, and the transmission of sound through an acoustic resonator is highly dependent on frequency. Sounds of the resonance frequencies peculiar to each resonator are less attenuated than other sounds and are therefore radiated with a higher relative amplitude, or with a greater relative loudness, than other sounds; the larger the frequency distance between a sound and a resonance is, the more weakly the sound is radiated. The vocal tract has four or five important resonances called formants. The many voice-source partials fed into the vocal tract traverse it with varying success depending on their frequency; the closer a partial is to a formant frequency, the more its amplitude at the lip opening is increased. The presence of the formants disrupts the uniformly sloping envelope of the voice-source spectrum, imposing peaks at the formant frequencies. It is this perturbation of the voice-source envelope that produces distinguishable speech sounds: particular formant frequencies manifest themselves in the radiated spectrum as peaks in the envelope, and those peaks are characteristic of particular sounds.

The formant frequencies are determined by the shape of the vocal tract. If the vocal tract were a perfect cylinder closed at the glottis and open at the lips and 17.5 centimeters (about seven inches) long, which is about right for the average adult male, then the first four formants would be close to 500, 1,500, 2,500 and 3,500 hertz (cycles per second). Given a longer or shorter vocal tract, these basic frequencies are somewhat lower or higher. Each formant is associated with a standing wave, that is, with a static pattern of pressure oscillations whose amplitude is at a maximum at the glottal end and near a minimum at the lip opening [*see illustration on page 86*]. The lowest formant corresponds to a quarter of a wavelength, which is to say that a quarter of its wavelength fits within the vocal tract. Similarly, the second, third and fourth formants correspond respectively to three-quarters of a wavelength, one and a quarter wavelengths and one and three-quarters wavelengths.

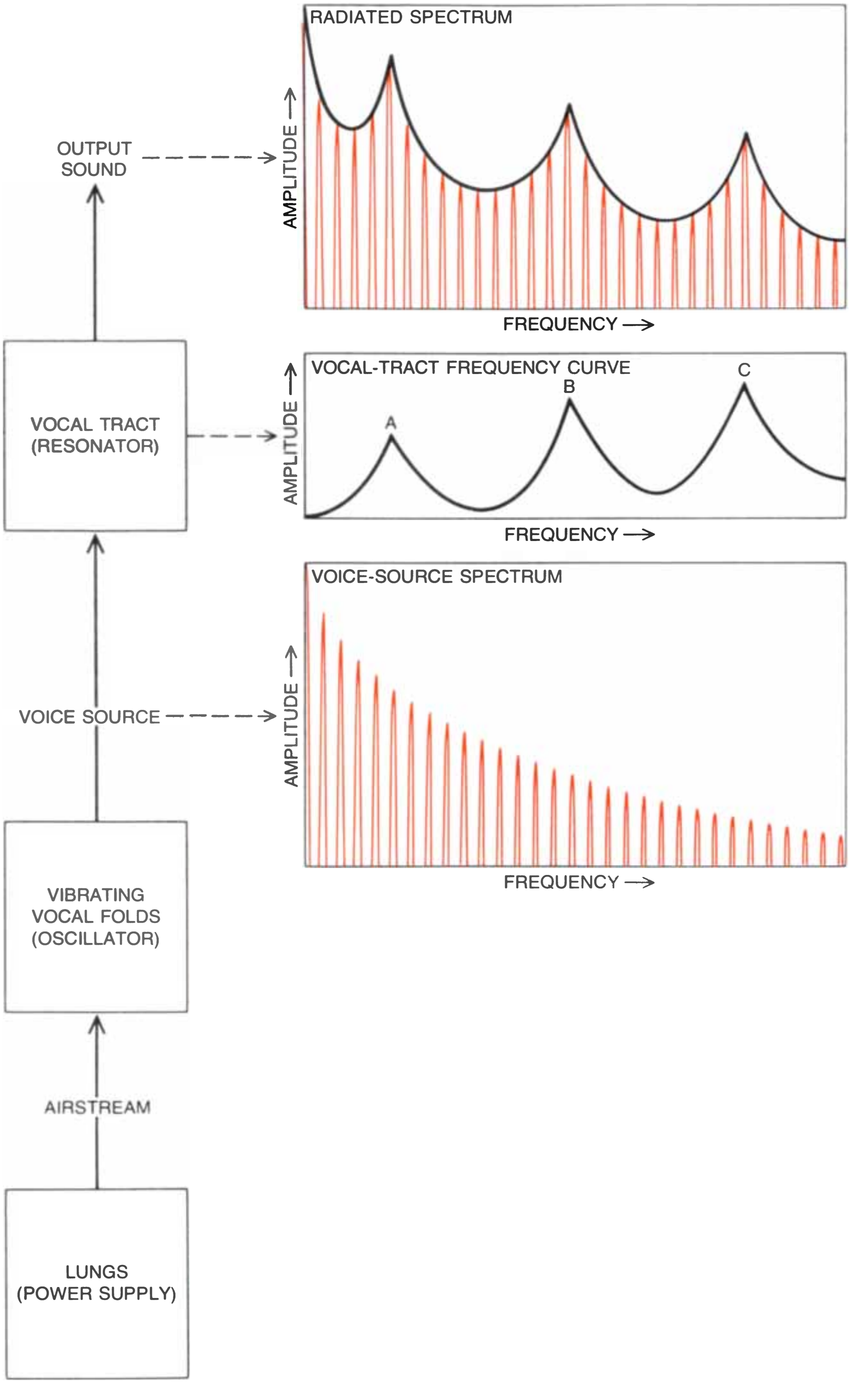
Any change in the cross section of the vocal tract shifts the individual formant frequencies, the direction of the shift depending on just where the change in area falls along the standing wave. For example, constriction of the vocal tract at a place where the standing wave of a formant exhibits minimum-amplitude pressure oscillations generally causes the formant to drop in frequency; expansion of the tract at those same places raises the frequency.

The vocal tract is constricted and expanded in many rather complicated ways, and constricting it in one place affects the frequency of all formants in different ways. There are, however, three major tools for changing the shape of the tract in such a way that the frequency of a particular formant is shifted in a particular direction. These tools are the jaw, the body of the tongue and



**VOICE ORGAN** is composed of the lungs and the larynx, pharynx, mouth and nose, shown in longitudinal section (*top*). The larynx is a short tube at the base of which are twin infoldings of mucous membrane, the vocal folds. The larynx opens into the pharynx; the opening is protected during swallowing by the epiglottis. The larynx, pharynx and mouth (and in nasal sounds also the nose) constitute the vocal tract. It is a resonator whose shape, which determines vowel sounds, is modified by changes in the position of the articulators: the lips, the jaw, the tip and body of the tongue and the larynx. The vocal folds, seen from above in a transverse section (*bottom*), are opened for breathing and are closed for phonation by the pivoting arytenoid cartilages.

the tip of the tongue. The jaw opening, which can constrict the tract toward the glottal end and expand it toward the lip end, is decisive in particular for the frequency of the first formant, which rises as the jaw is opened wider. The second-formant frequency is particularly sensitive to the shape of the body of the tongue, the third-formant frequency to the position of the tip of the tongue. Moving the various articulatory organs in different ways changes the frequencies of the two lowest formants over a considerable range, which in adult males averages approximately from 250 to 700 hertz for the first formant and from 700 to 2,500 hertz for the second. Moving the articulatory organs is what we do when we speak and sing; in effect we chew the standing waves of our formants to change their frequencies. Each articulatory configuration corresponds to a set of formant frequencies, which in turn is associated with a particular vowel sound. More specifically, the formant frequencies enhance voice-source partials of certain frequencies and thus manifest themselves as the peaks characterizing the spectrum envelope of each vowel sound.



**VOICE ORGAN is composed functionally of a power supply, an oscillator and a resonator. The airstream from the lungs is periodically interrupted by the vibrating vocal folds. The resulting sound, the voice source, has a spectrum (*right*) containing a large number of harmonic partials, the amplitude of which decreases uniformly with frequency. The air column within the vocal tract has characteristic modes of vibration, or resonances, called formants (*A*, *B*, *C*). As the voice source moves through the vocal tract each partial is attenuated in proportion to its distance from formant nearest it in frequency. The formant frequencies thus appear as peaks in the spectrum of the sound radiated from the lips; the peaks establish particular vowel sounds.**

All the elements and functions of the voice organ that I have been describing are common to singers and nonsingers alike. Do singers bring still other faculties into play or manipulate the voice instrument in different ways? Let us begin by comparing normal male speech and operatic singing. Careful attention to a singer's voice reveals a number of modest but very characteristic deviations in vowel quality from those of ordinary speech. For example, the *ee* sound of a word such as "beat" is shifted toward the umlauted *ü* of the German "*für*"; the short *e* of "head" moves toward the vowel sound of "heard." The general impression is that the quality of the voice is "darker" in singing, somewhat as it is when a person yawns and speaks at the same time; voice teachers sometimes describe the effect as "covering."

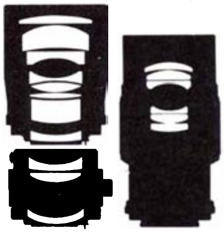
These shifts in vowel quality have been found to be associated with peculiarities of articulation. In "covered" singing the larynx is lowered, and X-ray pictures reveal that the change in the position of the larynx is accompanied by an expansion of the lowest part of the pharynx and of the laryngeal ventricle, the space between the true vocal folds and the false ones. It is interesting to note that voice teachers tend to agree that the pharynx should be widened in singing, and some of them mention the sensation of yawning. In other words, a low larynx position and an expanded pharynx are considered desirable in singing.

What we recognize as a darkened voice quality in singing is reflected very clearly in the spectrum of a sung vowel sound. A comparison of the spectra of the vowel in "who'd" as it is spoken and sung shows that the two lowest formant frequencies are somewhat lower in the sung version and that the spectral energy, or amplitude, is considerably higher between 2,500 and 3,000 hertz [*see top illustration on page 89*]. This spectral-envelope peak is typical of all voiced sounds sung by professional male singers. Indeed, its presence, regardless of the pitch, the particular vowel and the dynamic level, has come to be consid-

# New
# Vivitar Series 1
# 90mm f2.5
# macro lens

## Optical Specifications

| | |
|---|---|
| Elements/Groups: | |
| Main lens: | 8 elements, 7 groups. |
| Macro Adapter: | 3 elements, 3 groups. |
| Lens coating: | VMC Vivitar multicoating. |
| Angle of acceptance: | 27° |
| Aperture range: | f2.5 to f22 |
| Minimum focus distance from film plane: | 39.3cm (15.5 in.) without Adapter, 35.5cm (14 in.) with Adapter. |
| Maximum reproduction ratio: | 1:2 without Adapter, 1:1 with Adapter. |



*Vivitar Series 1 design* *conventional design*

The Vivitar Series 1 90mm f2.5 macro lens may well be the sharpest lens in 35mm photography. The lens utilizes the floating group concept to maintain optimum performance throughout its focusing range. Tests for resolution and contrast from infinity to life-size (1:1 reproduction) give the Series 1 90mm f2.5 lens some of the highest overall axial to corner ratings obtained for macro lenses.

A 90mm macro lens yielding good performance could have been produced using a normal double Gauss design. To achieve and maintain very high levels of performance from infinity to life-size, however, Vivitar Series 1 designers used a unique 8 element/7 group configuration to bring aberrations to an absolute minimum and to stabilize them throughout focusing distances from a reproduction ratio of 1:2 to infinity. The extremely stringent performance demands to eliminate aberrations in the lens also required the use of optical glass of a very high index of refraction and some uncommonly thick elements.

Using the concept of a null lens, borrowed from astronomical optics, the designers created a 3 element macro corrector-lens adapter that achieves a true flat-field image, high resolution and excellent contrast in the 1:2 to 1:1 reproduction range. The macro corrector-lens adapter is not a magnifying lens. Its sole function is to compensate aberrations produced when the lens is moved away from the film plane for life-size photography.

The selection of 90mm as the focal length of this lens provides two distinct benefits to photographers. It is an ideal focal length for portraiture and general purpose photography. When used with its macro adapter the lens allows life-size photography at a greater working distance from the subject than shorter focal length macro lenses. This greater working distance increases the photographer's options in illuminating macro subjects and lessens chances of disturbing live subjects.

As with all Vivitar Series 1 lenses, the mechanical configuration has been as carefully engineered and manufactured as the optics. The lens engravings give the photographer maximum information and legibility. The lens barrel styling is entirely functional, all controls being placed in the most appropriate positions for precise, comfortable operation.

## Mechanical Specifications

| | |
|---|---|
| Length at infinity: | 90mm (3.5 in.) without Adapter. 138mm (5.4 in.) with Adapter. |
| Maximum barrel diameter: | 70mm (2.8 in.) |
| Weight | |
| Main lens: | 644 gms. (23 oz.) |
| with Adapter: | 936 gms. (33 oz.) |
| Filter size: | 58mm |
| Lens case: | Semi-hard, 2 compartment case. |

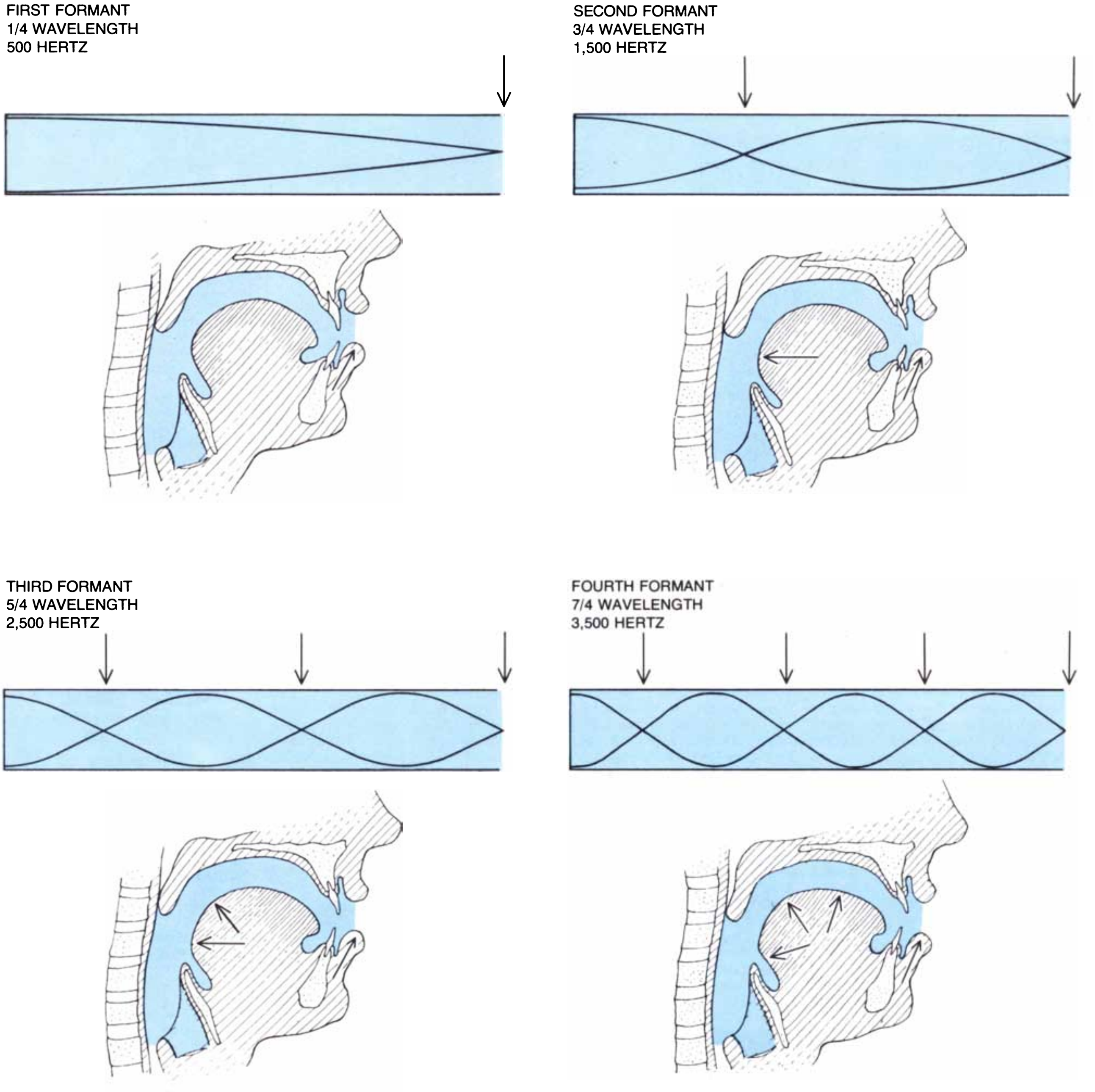
Available in mounts to fit Nikon, Canon, Minolta, Olympus OM and Universal Thread Mount Cameras.

**Vivitar®**

Marketed in the U.S.A. by
Ponder & Best, Inc.
Corporate Offices: 1630 Stewart St.,
Santa Monica, CA 90406.
In Canada:
Vivitar Canada, Ltd./Ltée., Quebec.

ered a criterion of quality; the extra peak has been designated the "singing formant."

What is the origin of the singing-formant peak? The peaks in the spectrum envelope of a vowel normally stem, as I have explained, from the presence of specific formants. The insertion of an extra formant between the normal third and fourth formants would produce the kind of peak that is seen in the spectrum of a sung vowel [*see bottom illustration on page 89*]. Moreover, the acoustics of the vocal tract when the larynx is lowered are compatible with the generation of just such an extra formant. It can be calculated that if the area of the outlet of the larynx into the pharynx is less than a sixth of the area of the cross section of the pharynx, then the larynx is acoustically mismatched with the rest of the vocal tract; it has a resonance frequency of its own, largely independent of the remainder of the tract. The one-sixth condition is likely to be met when the larynx is lowered, because the lowering tends to expand the bottom part of the pharynx. I have estimated on the basis of X-ray pictures of a lowered larynx that this lowered-larynx resonance frequency should be between 2,500 and 3,000 hertz, that is, between the frequencies of the normal third and fourth


FIRST FORMANT
1/4 WAVELENGTH
500 HERTZ

SECOND FORMANT
3/4 WAVELENGTH
1,500 HERTZ

THIRD FORMANT
5/4 WAVELENGTH
2,500 HERTZ

FOURTH FORMANT
7/4 WAVELENGTH
3,500 HERTZ

**FORMANTS** correspond to standing waves, or static patterns of air-pressure oscillations, in the vocal tract. Here the first four formants are shown as standing waves in cylindrical tubes, the schematic equivalent of the vocal tract (*colored areas in drawings*). The sine waves represent the amplitude of the pressure differential, which is always maximal at the glottal end and minimal at the lips. For the lowest formant a quarter of a wavelength is within the vocal tract and, if the tract is 17.5 centimeters long, the formant's frequency is about 500 hertz (cycles per second). The second, third and fourth formants are 3/4, 5/4 and 7/4 of a wavelength, and their frequencies vary accordingly. If the area of the vocal tract is decreased or increased at a place where the formant's pressure amplitude is at a minimum (*arrows*), that formant's frequency is respectively lowered or raised; the same change in area has the opposite effect if it is at a pressure maximum.

Noise—clamorous companion of man's progress—is becoming a significant environmental problem.

At the General Motors Research Laboratories we are seeking to reduce noise at its source through increased understanding of the mechanisms of noise generation. Simultaneously, we are studying human responses to traffic sounds.

Considerable effort has been focused on tires, a major source of noise in both cars and trucks. Interestingly, air flow around the tire is not a significant noise source. But aerodynamic pumping between treads can be, depending upon tread pattern. Also important are tread vibrations in the vicinity of the contact patch.

In another study, we are using signal coherence analysis to relate cylinder combustion pressure to noise radiation. This is part of an overall effort to learn in detail how engine structures transmit combustion-related noise.

Is it possible to quantify the annoyance associated with traffic noise? Psychological studies in one suburban area established that an $L_{eq}$ measurement (average amount of sound energy reaching the ear per unit time) of 60 dB was the approximate threshold above which people were willing to pay for decreased annoyance.

These research programs—and others being conducted at the Laboratories—are aimed at restoring one of life's more precious qualities . . . a quiet environment.

# We're working on several hush-hush projects.

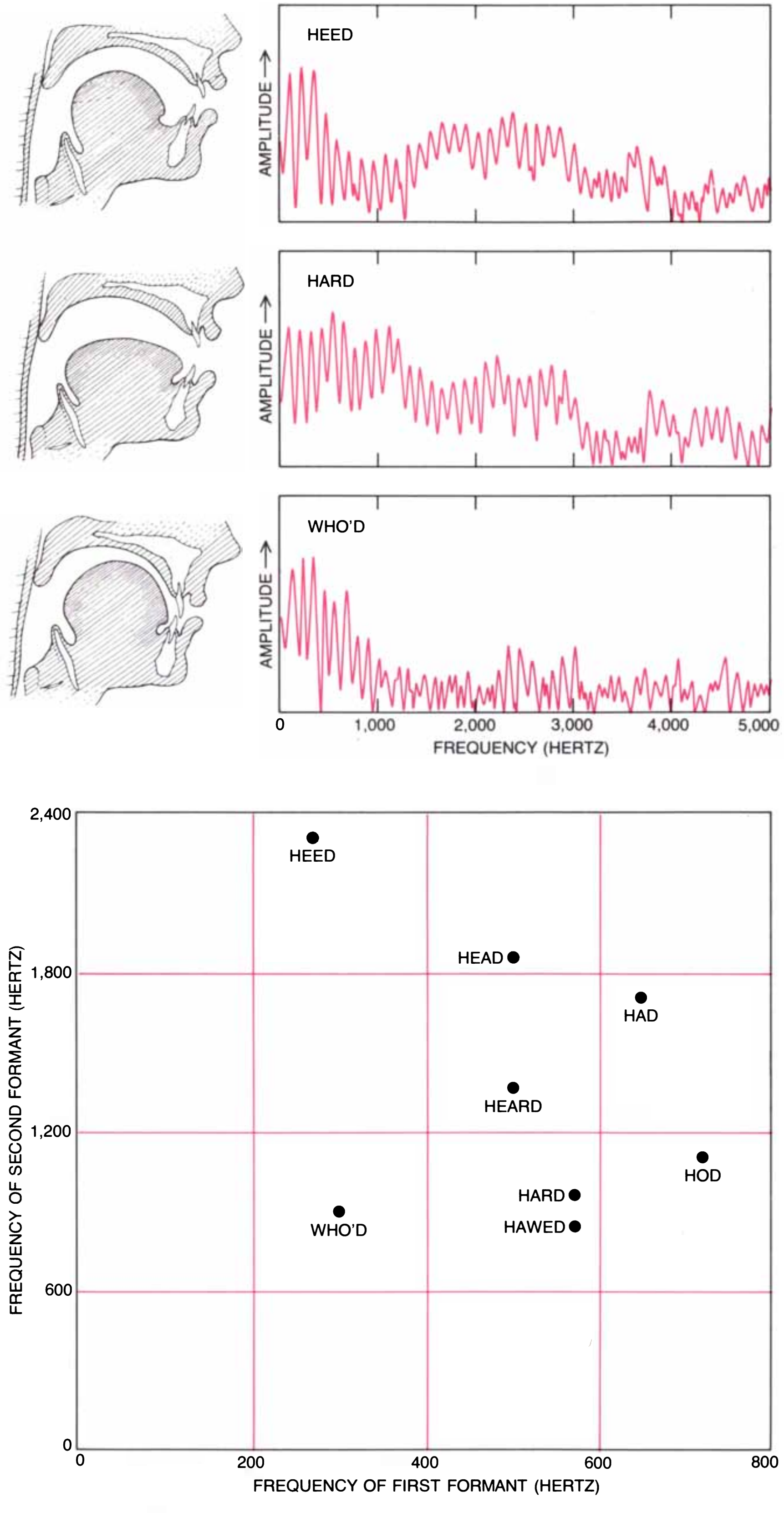
**General Motors**
**Research Laboratories**
Warren, Michigan 48090

formants and just where the singing-formant peak appears. The lowering of the larynx, in other words, seems to explain the singing-formant peak.

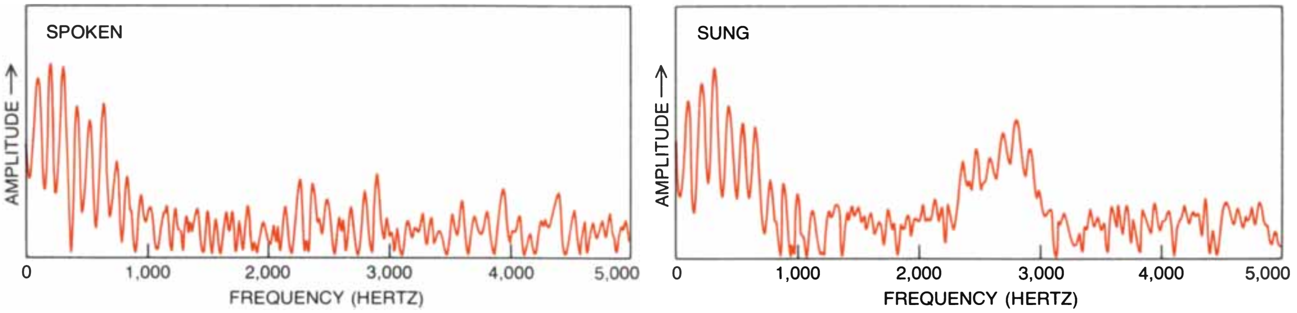
It also accounts for something else. Acoustically the expansion of the lowermost part of the pharynx is equivalent to an increase in the length of the vocal tract, and the lowering of the larynx adds still more to the length. The result is to shift downward all formant frequencies other than the larynx-dependent extra formant. This lowering of frequency is particularly notable in formants that depend primarily on the length of the pharynx. Two examples of such formants are the second formant of the vowels in "beat" and "head," and a drop in the frequency of those formants moves their vowels respectively toward those of "*für*" and "heard." The lowering of the larynx, then, explains not only the singing-formant peak but also major differences in the quality of vowels in speech and in singing.

To explain the singing formant's articulatory and acoustic origin is not enough, however. Why, one wonders, is it desirable for singers to lower the larynx, producing the singing formant and darkening the quality of their vowels? A plausible answer to the question has been found. It is related to the acoustic environment in which opera and concert singers have to work: in competition with an orchestra. Analysis of the average distribution of energy in the sounds of an opera or symphony orchestra shows that the highest level of sound is in the vicinity of 450 hertz; above that the amplitude decreases sharply with frequency. Now, normal speech develops maximum average energy at about the same frequency and weakens at higher frequencies. A singer who produced sounds with the energy distribution of ordinary speech would therefore be in trouble: the orchestra's much stronger sounds would drown out the singer's. The average sound distribution of a trained singer, on the other hand, differs from that of normal speech—and of an orchestra—mainly because of the singing-formant effect. We have shown that a singer's voice is heard much more easily against recorded noise that has the same average energy distribution as an orchestra's sound if the voice has a singing formant. Not only is the formant almost invariably audible, because its frequency is in a region where the orchestra's sound is rather weak, but also it may help the listener to "imagine" he hears other parts of the singer's spectrum that are in fact drowned out by the orchestra.

The singing formant is at an optimal frequency, high enough to be in the region of declining orchestral-sound energy but not so high as to be beyond the range in which the singer can exercise



**MOVEMENT OF ARTICULATORS** changes the cross section of the vocal tract, shifting formant frequencies. Three articulatory configurations are shown (*top*) together with the spectrum of the vowel sound produced by each; the peaks in the spectrum envelope reflect the formant frequencies. The chart (*bottom*) gives the frequencies of the first and second formants in some English vowel sounds as spoken by an average male. For a female or a child the envelope pattern would be about the same but the peaks would be shifted somewhat higher in frequency.

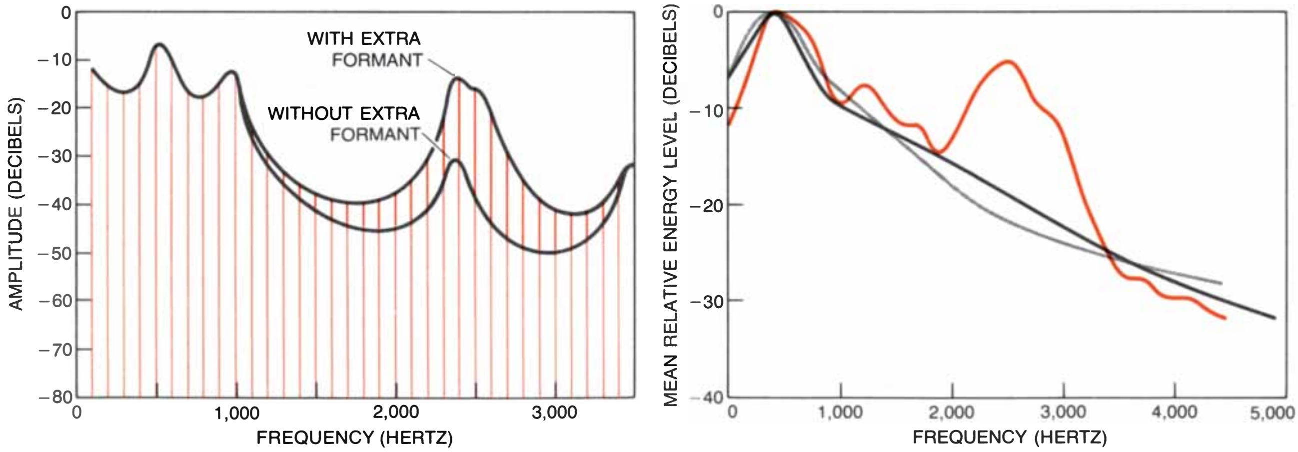


**VOWELS SOUND DIFFERENT in speech and in singing and the difference is visible in their recorded spectra. Here the spectra of the vowel in "who'd" as spoken (*left*) and as sung (*right*) by a male opera singer are compared. What is significantly different about the sung spectrum is the spectral-energy peak that appears in it between about 2,500 and 3,000 hertz. The new peak is called the singing formant.**

good control. Because it is generated by resonance effects alone, it calls for no extra vocal effort; the singer achieves audibility without having to generate extra air pressure. The singer does pay a price, however, since the darkened vowel sounds deviate considerably from what one hears in ordinary speech. In some kinds of singing that price is too high: the ideas and moods expressed in a "pop" singer's repertoire, for example, would probably not survive the deviations from naturalness that are required to generate the singing formant. And pop singers do not in fact darken their vowels; they depend on electronic amplification to be heard.

In cartoons a female opera singer is almost invariably depicted as a fat woman with her mouth opened very wide. In a study of female singers I have found that the way in which the jaw is manipulated is in fact quite different in ordinary speech and in singing. In speech the size of the jaw opening varies with the particular vowel, but in female singing it tends to depend also on the pitch of the tone that is being sung: the higher a soprano sings, the wider her jaw is opened. This suggested to me that a soprano must vary the frequency of her first formant according to the pitch at which she is singing. Analysis of formant frequencies confirmed that the articulation was being varied in such a way as to raise the first-formant frequency close to the frequency of the fundamental of the tone being sung. I noted such a frequency match whenever the frequency of the fundamental was higher than the frequency of a vowel's first formant in ordinary speech.

The reason becomes clear when one considers that the pitch frequency of a soprano's tones is often much higher than the normal frequency of the first formant in most vowels. If a soprano sang the vowel *ee* at the pitch of her middle C and with the articulation of ordinary speech, her first formant would be in the neighborhood of 270 hertz and the pitch frequency (the frequency of her lowest spectrum partial) would be almost an octave higher, at 523 hertz. Since a sound is attenuated in proportion to the distance of its frequency from a formant frequency, the fundamental would suffer a serious loss of amplitude. The fundamental is the strongest partial in the voice-source spectrum, and the higher its pitch is, the more important the fundamental is for the loudness of the tone, and so the singer's *ee* would be rather faint. Assume that her next sound was the *ah* sound of "father," to be sung at the pitch of high *F*. The fundamental, at 698 hertz, would be very close to the frequency of the first formant, about 700 hertz, and so the tone would be loud. The loudness of the singer's tones would vary, in other words, according to a rather unmusical determinant: the frequency distance between first formant and fundamental. In order to modulate the loudness accord-



**SINGING FORMANT'S ORIGIN (*left*) and its utility in singing (*right*) are demonstrated. An extra formant was inserted between the usual third and fourth formants in an experiment with an electronic resonator that behaves like the vocal tract (*left*). The new formant increased the amplitude of the partials near it by more than 20 decibels; similarly, an extra formant (achieved by lowering the larynx) supplies the high-frequency peak in the spectrum of a sung vowel. The three curves (*right*) show the averaged distribution of energy in the sound of orchestral music (*black*), of ordinary speech (*gray*) and of the late tenor Jussi Björling singing with an orchestra (*colored*). The distribution is very similar for speech and the orchestra at all frequencies; it is the singer's voice that produces the peak in the colored curve between 2,000 and 3,000 hertz. In that frequency region a singer's voice is loud enough, compared with an orchestra's sound, to be discerned.**

ing to the musical context, the singer would need to continually vary her vocal effort. That would strain her vocal folds. (Experiments with synthesized vowel sounds suggest that it would also produce tones more characteristic of a mouse under severe stress than of an opera singer!)

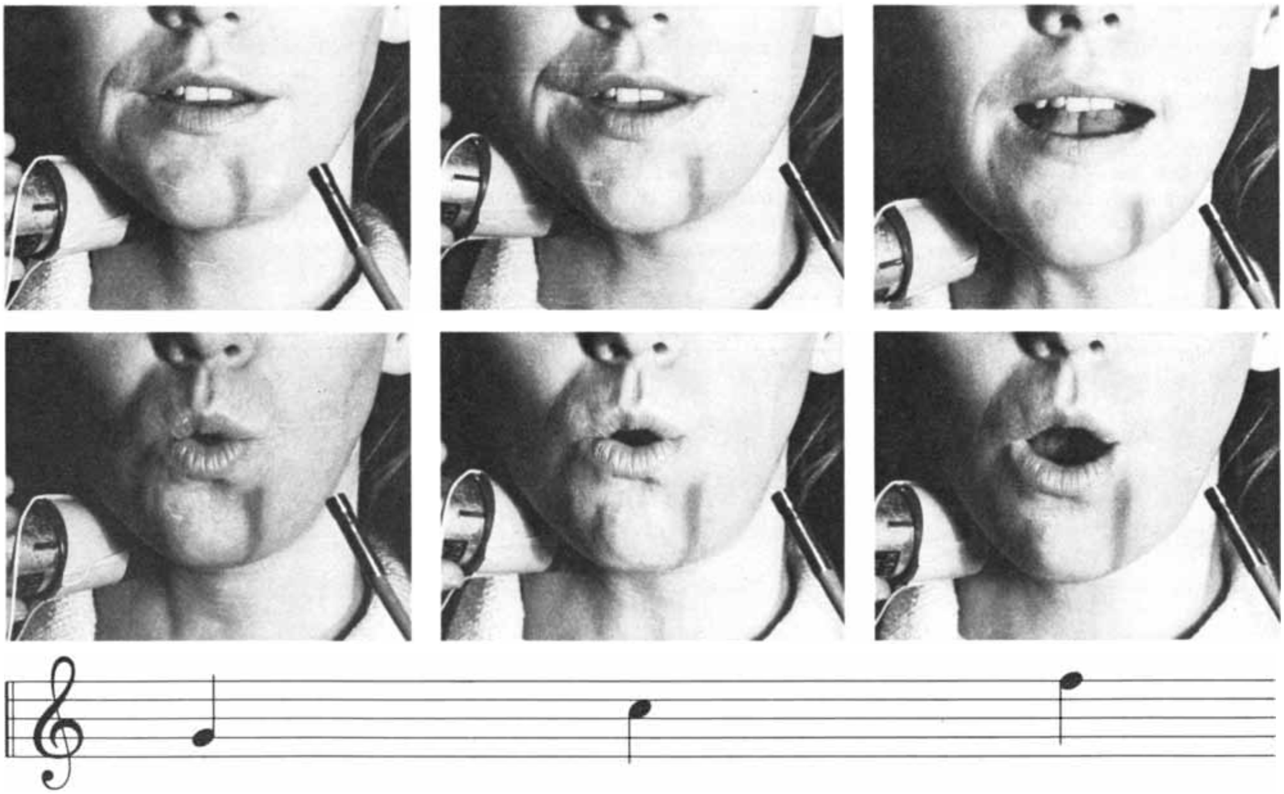
The soprano's solution is to move the first formant up in frequency to match the frequency of the fundamental, thus allowing the formant always to enhance the amplitude of the fundamental. The result is that there is minimal variation in loudness from pitch to pitch and from vowel to vowel. Moreover, changing the size of the jaw opening in this way provides maximum loudness at the lowest possible cost in vocal effort. The strategy is probably resorted to not only by sopranos but also by other singers whose pitch range includes frequencies higher than those of the first formants of ordinary speech: contraltos, tenors and occasionally even baritones.

It can be hard for a student of singing to learn this special way of regulating the jaw opening, and particularly hard if the jaw muscles are under constant tension. That may explain why many singing teachers try to get their students to relax the jaw. Another frequent admonition is: "Hear the next tone within yourself before you start to sing it." That could be necessary because proper manipulation of the jaw opening requires some preplanning of articulation for particular vowels and for the pitch at which they are to be sung. Opening the jaw, however, is not the only way to raise the first-formant frequency. Shortening the vocal tract by drawing back the corners of the mouth serves the same purpose, and that may be why some teachers tell their students to smile when they sing high tones.

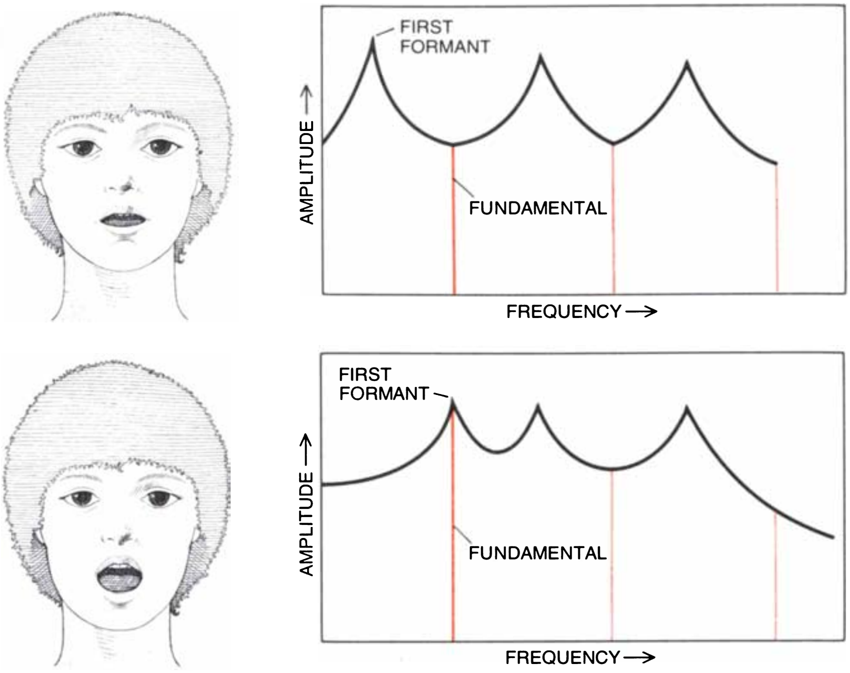
Since formant frequencies determine vowel quality, shifting the first-formant frequency arbitrarily according to pitch might be expected to produce a distorted vowel sound, even an unintelligible one. It does not have this effect, largely because we are accustomed to hearing vowels produced at various pitches in the ordinary speech of men, women and children with vocal tracts of very different lengths; if a vowel is high-pitched, we associate it with relatively high formant frequencies. The correlation is so well established in our perceptual system that we may perceive a change of vowel when we hear two sounds with identical formant frequencies but different pitches; if a singer raises her first-formant frequency with the pitch, some of that rise is actually required just to maintain the identity of the vowel. It is true that when the pitch is very high, our ability to identify vowels deteriorates, but that seems to be the case no matter what the formant frequencies are. The soprano, in other words, does not sacrifice much vowel intelligibility specifically as a result of her pitch-dependent choice of first-formant frequency. (Incidentally, composers of vocal music are conscious of the problem of vowel identification at high pitches and generally avoid presenting important bits of text only at the top of a soprano's range; often the text is repeated so that the words can be well understood at a lower pitch.)

It is clear that a good deal of the difference between spoken and sung vowels can be explained by the singer's need for economy of vocal effort. The general idea is the same, whether in being heard over the orchestra or in maintaining loudness at high pitch: to take advantage of vocal-tract resonance characteristics so as to amplify sounds. The importance of these resonances, the formants, is paramount.

Confirmation of the importance of



**SOPRANOS** and other singers of high tones tend to open their mouth wider with rising pitch. The tendency is demonstrated in these photographs of a soprano singing the vowel sounds of "heed" (*top*) and of "who'd" (*middle*) at successively higher pitches, shown in musical notation (*bottom*). When these photographs were made, the singer held a vibrator against her neck and a small microphone was placed near her lips. She began to sing each vowel at a specified pitch and then, with the vibrator turned on, she stopped singing but maintained the positions of the articulatory organs. The vibrator now supplied a steady, low-pitched sound that was influenced by the singer's vocal tract just as her own voice source would have been but that was more suitable for analysis than a high voice tone, which has few partials.



**NEED FOR WIDE JAW OPENING arises from the fact that a soprano must often sing tones whose fundamental (lowest partial) is far higher in frequency than the normal first formant of the vowel being sung. When that is the case (*top*), the amplitude of the fundamental is not enhanced by the first formant and the sound is weak. Opening the jaw wider raises the pitch of the first formant. When the first-formant frequency is raised to match that of the fundamental (*bottom*), the formant enhances the amplitude of the fundamental and the sound is louder.**

the formants was provided by a recent study of how male voices are classified as bass, baritone or tenor. Obviously the singer's frequency range is ultimately the determinant, but even when the true range (which is established primarily by the shape, size and musculature of the vocal folds) has not yet been developed, a good voice teacher can often predict the classification after listening to a student's voice. How is that possible? Thomas F. Cleveland, who was visiting our laboratory at the Royal Institute of Technology in Stockholm and is now at the University of Southern California, analyzed vowels sung by basses, baritones and tenors with respect to formant frequencies and the spectrum of the voice source. Then he had a jury of voice teachers listen to the vowel samples and classify the voices. The teachers tended to classify vowels in which the formant frequencies were comparatively low as having been sung by bass voices and vowels whose formant frequencies were high as having been sung by tenors. Variations in the voice-source spectrum (which varied slightly with the pitch at which a vowel was being sung), on the other hand, did not provide a basis for consistent classification. In a second test the same jury judged a series of synthesized (and therefore clearly defined) sounds and confirmed Cleveland's original impression: the lower the formant frequencies of a given vowel were, the lower the singer's voice range was assumed to be.

Cleveland found that typical bass and tenor voices differ in formant frequencies very much as male and female voices do. The formant-frequency differences between males and females are due mainly to vocal-tract length, and so the bass-tenor differences are probably also largely explained by the same physical fact. Formant frequencies are determined, however, not only by the individual's vocal-tract morphology but also by habits of articulation, which are highly variable. Be that as it may, vocal-tract morphology must set limits to the range of formant frequencies that are available to a singer.

At this point the reader who knows and cares about music may be rather disappointed. I have failed to mention a number of factors that are often cited as determinants of excellence in singing: the nasal cavity, head and chest resonances, breathing and so on. These factors have not been mentioned simply because they seem to be not relevant to the major acoustic properties of the vowel sounds produced in professional operatic singing. Our research suggests that professional quality can be achieved by means of a rather normal voice source and the resonances of the vocal tract.

Our implied model may not be perfect, to be sure. It is just possible, for example, that the nasal cavity has a role in the singing of vowels that are normally not nasalized. If that is so, we have attributed its effect to the voice source, thus compensating for one error by making another. Moreover, we have dealt only with sustained vowel sounds, whose production is important but is certainly not the only acoustic event in singing.

Resonances outside the vocal tract, such as in the head or the chest, cannot contribute appreciably to the singer's acoustic output in view of the great extent to which sound is attenuated as it passes through tissues. This is not to say that such resonances may not be important to the singer, who may receive cues to his own performance not only from what he hears but also from felt vibrations. As for breathing, it is clear that the vocal folds would vibrate no matter by what technique an excess of air pressure is built up below the glottis. Breathing and laryngeal manipulation are likely to be physiologically interdependent, however, since the larynx is the gatekeeper of the lungs. Probably different ways of breathing are associated with different adjustments of the larynx, and probably some ways are effective for singing and others are inadequate or impractical.

Finally we return to the original question: What is so special about a singer's voice? The voice organ obeys the same acoustic laws in singing that it does in ordinary speech. The radiated sound can be explained by the properties of the voice-source spectrum and the formants in singing as in speech. From an acoustical point of view singers appear to be ordinary people. It is true that there is a major difference between the way formant frequencies are chosen in speech and the way they are chosen in singing, and hence between the way vowels are pronounced in singing and the way they are pronounced in speech. A man with a wide pharynx and with a larynx that will resonate at a frequency of between 2,500 and 3,000 hertz is likely to be able to develop a good singing voice more readily than a person who lacks those chacteristics. And his progress may be facilitated if his vocal folds give him a range that agrees with his formant frequencies. As for a female singer, she should be able to shift the first formant to join the pitch frequency in the upper part of her range; that requirement may bar some women with a long vocal tract from having a successful career as a coloratura soprano. There are, in other words, a few morphological specifications that probably have some effect on the ease with which someone can learn to sing well. There are other conditions that may be more important, however. It is in the complex of knowledge, talent and musical instinct that is summed up as "musicality," rather than in the anatomy of the lungs and the vocal tract, that an excellent singer's excellence lies.