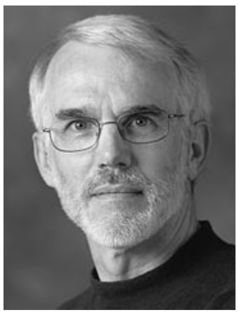
A Case for Voice Science in the Voice Studio

By Kenneth Bozeman



Kenneth Bozeman

Journal of Singing, January/February 2007
Volume 63, No. 3, pp. 265–270

If one accepts Manuel Garcia's invention of the laryngoscope as a plausible starting point, the singing community now has had over a century and a half of experience with the explorations and contributions of the field of voice science. University faculty are increasingly renaming and adjusting voice pedagogy courses to be courses in voice science and pedagogy. While teachers in earlier centuries relied primarily on the authority of historic schools of pedagogy and teacher lineage, the situation today is somewhat different. Though few would claim to base their teaching exclusively on voice science, many now seek to compare historic pedagogy, or at least their own personal pedagogic histories, with information emerging from the investigations of voice science in order to confirm or adjust their approach, for new input, and for help in articulating more accurately vocal function and strategy for their students.

There are nonetheless those who claim that voice science has not helped voice instruction, that it has been a distraction, or still worse, a substitute for effective teaching. While this no doubt has been true in some instances, the same criticism could be leveled at exclusive reliance on historic pedagogy. Excellence in the teaching of singing in either case is not an easy pursuit. The prudent course of action would be to consider every potentially fruitful resource in the quest for effectiveness. There is no necessary conflict between the art of teaching and the investigations of voice science. If approached with sufficient care, patience, and humility, they can be tremendous allies.

This is not to say that emerging voice science always has been on the right track, nor that everything it observes is of use in the studio. That is not how science proceeds nor is pedagogy its only or even its primary purpose. Furthermore, there is still too little collaboration between excellent teachers and excellent voice scientists. There is and probably always will be some skepticism from both sides. Can teachers ever be more than dabblers in science who, because of their less sophisticated understanding of the principles involved, risk making naive errors in their tentative hypotheses? And how sophisticated are the assessments of voice scientists in matters of vocal quality? Do voice scientists ask and explore the questions that would be most beneficial to the pedagogic community? The best hope for effective collaboration lies both in increased communication and cooperation between excellent teachers and excellent scientists and in those few individuals who are truly conversant in both fields. The more that teachers and scientists honestly and respectfully seek to understand each other, acknowledging the value of each other's perspective, the more progress we will make with fewer misguided or marginally relevant tangents.

A SIMPLE EXAMPLE

Let me give an example of how very simple science can assist pedagogy. Though we all have heard excellent singers with notably noisy breathing, most voice teachers recommend a noiseless inhalation. Noiseless inhalation is not only aesthetically more pleasing, it is evidence of prephonatory tuning for an open throat. Noisy inhalation is caused by some obstruction or narrowing of the vocal tract, usually in the pharynx but possibly also at the glottis, that generates an increase in air speed through that narrowing with resultant audible turbulence. Because of our skewed perception of what feels open throated, merely telling the student to open her throat rarely eliminates the noise. She probably will form a yawny /ɑ/ shape, perhaps lowering the pitch of the noise somewhat but not eliminating it. This is not her fault. Ask a crowd of uninitiated singers which vowel *feels* the most open throated and most will say /ɑ/, or which feels the most closed throated and most will say /i/. These are accurate perceptual answers to the question of how it *feels*, but the opposite of the actual physical circumstance, as any voice examination reveals. (The general practitioner asks for /ɑ/ to see the back of the mouth while the otolaryngologist pulls on your tongue and asks for /i/ to get a look at the vocal folds!) If, on the other hand, the teacher points out the wind chill effect that is increased in the vicinity of the vocal tract narrowing, progress on noiseless inhalation can be quickly made.

Wind chill effect is the increased cooling caused both by evaporation of the moist surfaces of the throat and by the difference between air temperature and body temperature. This cooling will be sharply localized where air speed is the quickest, precisely where the vocal tract is the narrowest. Noisy inhalations are therefore almost always coolest in the throat at the back of the mouth. Instructing the student to experiment with mouth, throat, and vowel shape in order to move the cooling effect to the front of the mouth, without closing the front so much as to make noise there, will effectively open the throat and reduce or eliminate the noise.

The student's experimentation to "cool in the front" can be guided by our scientific knowledge of the actual tube and tongue shapes that accompany each vowel and by our knowledge of the acoustic effects of vowels on tube shape. For example, having the student attempt to lower the "pitch" of the noise as much as possible also will tend to open the throat, since the noise element in a moderately opened vocal tract tends to "play" the first formant which is lower in open throated vowels. (It plays the second formant when the throat has been closed somewhat as in typical whispering.) Another common strategy takes advantage of our knowledge of the effect of a "hint of a yawn" on the soft palate and pharyngeal muscles. *Note bene:* The sensation of a noiseless inhalation that cools more near the front of the mouth initially can be disconcerting, since it *feels* almost throatless and certainly not like what one is used to thinking of as open throated. It also eliminates the "sucking" sensation that previously reassured the student that air was being taken in, making her wonder if she "got enough air." Less resistance to inhalation equals less perception of having drawn air in. Nonetheless, a quick, noiseless inhalation is the most efficient inhalation; you get a lot of air in the least amount of time with the least inhalatory effort.

Now this is not high science and, with the exception of basic tube acoustics, hardly even *voice* science, but it *is* the productive application of knowledge from science (evaporation, heat exchange, effects of conduit size on air speed, etc.) to the teaching of singing.

A LESS SIMPLE EXAMPLE

A more complex application of voice science to voice pedagogy occurs in the registration events and acoustics of male *passaggio*. For most of the last century the majority paradigm for vocal registration maintained that it was an exclusively laryngeal and muscular phenomenon. This understanding observed that the two primary modes of muscular action that shape the vocal folds for singing and for pitch adjustment accounted separately for each range extreme and cooperated through the middle voice in varying proportions to blend the two registral poles for a seamless range. In male singers the *zona di passaggio* or transitional zone defined the area of transition from chest register dominance to head register dominance (often referred to as falsetto in voice science literature). In voice pedagogy literature as documented by Richard Miller this transition zone is described in the male voice as spanning about a perfect fourth whose location in the upper middle voice varies by voice type.[1]

This paradigm of registration was in harmony with the prevailing acoustic paradigm, the linear source-filter theory of vocal acoustics, in which acoustic factors seemed not to affect laryngeal registration to any great degree. Though there were minority opinions to the contrary (for example, Berton Coffin),[2] and though acoustic factors were acknowledged in pedagogies that observed and/or advocated vowel modification in the upper voice, registration still was considered to be primarily about the relative proportion of thyroarytenoid (TA, chest) and cricothyroid (CT, head) muscle involvement.

We now increasingly realize that registration in singing is not exclusively laryngeal; it is both muscular and acoustic. While a gradual, or at least smooth shift from dominance of the thyroarytenoid to the cricothyroid muscles *does* occur in the trained male singer ascending the scale, acoustic factors either facilitate or impede this coordination to a very significant extent. As our understanding of the nonlinearity of vocal acoustics and production increases, this becomes more evident, with resonance adjustments and phenomena heavily influencing, if not determining vibrator adjustments. This article will address the particular interactions of pitch and vowel acoustics in the male voice and their relationship to negotiating the *zona di passaggio.*

I distinctly remember hearing a synthesized voice for the first time at the 1981 NATS convention in Minneapolis. In a presentation on the singer's formant, vowel and singer's formants had been synthesized to represent a tenor voice singing /ɑ/ over an ascending major scale. Presetting the formants implied a particular, unchanging tube shape with specific resonance characteristics (formants). The changing pitches of the scale caused a different set of voice source partials to be resonated by that stable tube shape with each ascending step. It was remarkable to hear the synthesized voice modulate through all of the color changes associated with a trained tenor negotiating his *passaggio*. In fact, I asked whether any other parameters had been programmed into the synthesizer to generate such color changes. On the contrary, they were purely the result of the interaction between the changing source spectrum partials and the fixed formant settings, in other words, of singing the scale while simultaneously maintaining the *same* vocal tract shape. I previously had assumed that the color changes (vowel modifications?) of the *passaggio* were the result of the gradual muscular shift from chest voice (TA dominance) to head voice (CT dominance). What began to emerge from this experience was an understanding that the timbral changes of the *passaggio* were brought about primarily by acoustic interactions, regardless of the muscular registration adjustments that might accompany them.

Over the years since that beginning, other questions arose and more pieces of the puzzle fell into place. In my own vocal journey, during a lesson with Richard Miller, in fact, I had the surprising experience of a close /e/ vowel "turning over" rather lower in the *passaggio* than expected, making it noticeably less effortful. Being startled by my internal perception of the sound, I asked if it was acceptable. Richard confirmed that the timbre and vowel sounded fine. I previously had been working under the assumption that all vowels would shift at about the same place, close to the so-called *secondo passaggio* of the Italian international school, as would be logical if the shift were due purely to muscular adjustments of the larynx. It began to dawn on me that if the timbral changes we hear in the *passaggio* are not muscular in origin but are caused by source and tube acoustic interactions, since the acoustics of the tube—especially of the first two formants—differ from vowel to vowel, there were reasons not to expect all vowels to "turn over" at the same pitch in the scale. But what factors were causing the change, and would knowing more about them assist me in working out my *passaggio* and in training my students?

Over the next fifteen years, with the help of a Kay Elemetrics Sonograph in my studio (a fortuitous development), I was able to observe at least informally that in trained male singers, vowels "close" or "turn over" in a parallel relationship to the location of their first formant. Initial investigations would indicate that the change occurs when the second voice source partial crosses the first vocal tract formant, in other words an octave below the location of the first vowel formant. Bear in mind that since singers "tune" their vocal tracts for resonance, these formants are not necessarily located exactly as in speech, nor are they fixed, but can be adjusted by changes in the vocal tract shape. Nonetheless, vowels will seem to "turn over" at frequency locations that parallel the locations of the vowels' first formants an octave lower. A sung scale therefore will go through these vowel color changes or modifications at fairly predictable locations. Closer vowels with low first formants turn over lower; more open vowels with higher first formants turn over higher. In a lyric tenor voice these occur approximately as follows: the close vowels /i/ and /u/ modify (close) quite low in the range, at about E_3–$F^{\#}_3$. (Since these vowels are inherently close in pronunciation, this event is more commonly perceived as the need to *open* them *below* that frequency for resonance and carrying power.) The close forms of /o/ and /e/ (as in German or French) modify at or just below the *primo passaggio*, or about C_4-D_4; /ɛ/ and /ɔ/ near mid-*passaggio,* at about $E^{\flat}_4$-F_4; and /ɑ/, the most open vowel, at F_4-G_4, the location of the *secondo passaggio*. Though a synthesized voice can behave as if purely linear, that is, unlike in real singing the source and filter can be independently controlled, the color changes (vowel modifications) through the *zona di passaggio* are nonetheless quite like those of a living, nonlinear (i.e., acoustically interactive) male singer.

PEDAGOGIC APPLICATION

Of what pedagogic value is this observation? It is helpful to anticipate that, without encouragement, very few young males instinctively will allow these changes from speech timbre to occur as they ascend the range. Typically, in an attempt to avoid what to them seems a disconcerting vowel change, they will gradually shorten the tube by raising the larynx, constricting the deep throat, and *over* opening the mouth in order to raise the first formant locations and postpone the modification. These adjustments artificially preserve the relationship between the first two voice source partials and the first formant locations, avoiding "cover" modifications, and result in the typically "reachy" belt timbre (*voce aperta*). They also usually result in a short range, since raising and constricting the larynx inhibits more appropriate muscular registral adjustments from occurring. The more sophisticated strategy of the trained singer is to maintain the settled, open throat, even while opening the mouth somewhat as the pitch ascends, thus allowing each vowel to migrate and eventually "turn over" or "cover" at its appropriate frequency. Once the vowel has begun to turn, increased mouth opening can follow (rather than lead) so that the degree of modification is minimized and vowel intelligibility is preserved.

Anticipating where these events typically occur can assist the teacher in guiding the student through the *passaggio*. Knowing that their locations vary by vocal weight and type can be of eventual assistance in determining voice classification. Knowledge of these behaviors also can provide more than one acceptable solution to *passaggio* issues. Some may prefer to postpone the event by moving to a more open neighbor vowel, accepting the compromise in pronunciation to buy power or for some other objective; for example, opening an /e/ toward an /ɛ/ upon ascent. Alternatively, if a vowel is getting stuck in *aperta* production and not closing or tipping soon enough upon ascent, the student can encourage it to shift earlier by moving toward its closer neighbor, which would naturally turn over lower. Moving to the closer vowel also draws the student's attention to the need for internal space, possibly helping to keep the larynx settled and the palate high. In working out a young voice both of these strategies have been productive, providing some useful variety of approach until the voice settles into a more automatic adjustment. And finally, though the effects of these transitions are primarily the result of acoustic and not muscular factors, as stated above it is increasingly recognized that they are highly interactive with the muscular adjustments of the larynx. Even pedagogies from opposing camps (for example, Cornelius Reid and Berton Coffin) have acknowledged the differing influences of vowel choice on laryngeal adjustment. Open vowels such as /ɑ/ generally are observed to encourage chest register, while close vowels such as /u/ and /i/ encourage head register. Why this should be the case is perhaps not yet fully understood, but that it *is* the case is widely experienced. A significant part of the answer lies in the fact that close vowels increase vocal tract inertance and thereby lower phonation threshold pressure, as reported by Titze.[3] There is probably also some influence due to the fact that close vowels close or turn over lower in the range. Since /i/ and /u/ close well below the traditionally understood location of the male *passaggio,* they pose less risk of getting the student stuck in a *voce aperta* adjustment through the *passaggio.* This can be fashioned into a strategy for influencing challenging vowels. For example, an /ɑ/ is often resistant to transitioning toward a lighter adjustment in the upper *passaggio*, becoming instead increasingly blatant. An exercise can be improvised that surrounds or even temporarily replaces it with vowels that turn over sooner, such as /o/ or /u/ or the glide equivalent /w/. If a teacher understands the principles involved of how the behaviors of vowels differ by frequency, he or she can be rather creative in fashioning helpful corrective strategies for particular situations.

QUESTIONS RAISED

Given the observation that vowels vary significantly in timbre and behavior across range as described above, further questions suggest themselves relative to effective *passaggio* training.

1) What exactly is meant by vowel modification? Is it:
    - the deliberate substitution of a different shape or sound than that to which we are accustomed in speech? (Vowel substitution: sing /ʊ/ instead of /ɑ/ at that pitch.)
    - the slight coloration of a vowel that we still perceive to be the spoken vowel? (Vowel coloration: think more /ʊ/ in your /ɑ/ at that pitch.)
    - the fact that the *same* shape of tube yields vowels that differ as a function of frequency? (Vowel migration: the shape for /ɑ/ in typical speech range yields more of an /ʊ/ in the upper voice. Therefore leave the shape basically the same, but realize [anticipate and allow] that the vowel will migrate somewhat towards an /ʊ/ *on its own* as you ascend.)
    - simply the need to open the mouth more with ascending pitch, while continuing to think the intended vowel?
    - perhaps a combination of some or all of the above?
2) Given the aesthetic and historic pedagogic goals of vowel purity and intelligibility, what then is *acceptable* vowel modification versus disturbing vowel *distortion*?
3) And of course, how do vowel modifications differ in training the female *passaggio*?

But these are questions for another time.

CONCLUSION

There is much to be gained by careful consideration of the information emerging from the voice science community. Teachers needn't burden the voice student with great detail, but carefully can choose what is truly applicable and helpful. Well informed voice teachers also have much to offer the voice science community in terms of identifying and carrying out pedagogically fruitful research agendas and by providing context and access to excellent subject pools. By working together, well designed and executed studies can be undertaken that will fulfill that which motivates both communities: our mutual love of beautiful, liberated singing.

Some Definitions

Prephonotory tuning—the shaping of the vocal tract for successful resonance *prior* to phonation, typically during inhalation.

Passaggio—Italian for a transition in the voice from one register to another; sometimes used as a shorter reference to the *zona di passaggio*.

Zona di passaggio—Italian for a transition zone between the first and second *passaggi*. In the male voice it has been traditionally understood to cover about a perfect fourth, varying in location by vocal category.

Primo passaggio—the lower vocal register pivotal point, or entry into the *zona di passaggio*.

Secondo passaggio—higher vocal register pivotal point, or exit of the *zona di passaggio* and entry into the upper voice.

Sonograph—an instrument that does real time analysis of sound, capable of displaying wave form, power spectrum, and spectrographic images of sounds.

Voce aperta—Italian for "open voice:" a more open timbred, brighter, potentially blatant resonance adjustment that may be appropriate lower in the range for power and projection, but is inappropriate higher in the range. Where the voice can open acceptably varies by vowel: the closer the vowel, the lower the location.

Linear source filter theory of acoustics—the theory of vocal acoustics by which each element of the system modifies the contribution of the previous element leading to the sound radiated at the lips: the power source (breath) activates the voice source (vocal folds) generating an acoustic signal (laryngeal buzz) that is then filtered (resonated) by the vocal tract, which has its own acoustic characteristics (natural resonances or formants). The final product is then radiated from the lips.

Nonlinearity of acoustics—the increasing realization that the acoustic feedback of the filter (resonator) upon the vibrator (vocal folds) as well as the interactions of resonations from below (in the trachea) with the vocal folds can substantially influence the behavior of the vibrator, so that vocal acoustics are not merely unidirectional.

Voice source partials—the fundamental frequency and overtones produced by the vocal folds prior to being resonated by the vocal tract; the "laryngeal buzz" responsible for the pitch that we perceive and for providing the raw material for resonation of the vocal tract, hence contributing significantly to the ultimate color or timbral quality of the voice. The vocal tract can filter or resonate only acoustic input that is provided to it, primarily by the voice source.

Formant—a natural resonance of the vocal tract. Typically 3–5 formants are audibly present in a voice, though more may be activated given enough acoustic power. The lowest two formants are largely responsible for vowel differentiation. Higher formants may cluster to form the so-called "singer's formant."

Vowel formant—one of the lowest two natural resonances of the vocal tract. The tunable formant frequency differences from vowel to vowel enable us to decode or differentiate vowels from one another.

Thyroarytenoid—the muscle group attached just below the thyroid notch in front and to the arytenoids cartilages in the back, constituting the body of the vocal folds. When contracted they are responsible for the shorter, thicker vocal fold shape that forms the chest register and that produces chest voice quality. In this sense the vocal folds can be said to be made up of the "chest voice muscles."

Cricothyroid—the muscle group on the front and sides of the larynx between the cricoid and thyroid cartilages that tilt the thyroid and cricoid cartileges, stretching and thinning the vocal folds for head register and head voice quality. In this sense, the "head voice muscles" can be said to lie outside of the voice box.

Vocal tract inertance—the inertia (the sluggishness or resistance to acceleration) of the air column in the vocal tract.

Phonation threshold pressure—the minimum air pressure below the glottis at which the vocal folds can be brought into vibration.

NOTES

1. Richard Miller, *The Structure of Singing: System and Art in Vocal Technique* (New York: Schirmer Books, 1986), 116–117.
2. Berton Coffin, *Overtones of Bel Canto* (Metuchen, NJ: Scarecrow Press, 1980).
3. Ingo Titze, "The Use of Low First Formant Vowels and Nasals to Train the Lighter Mechanism," *Journal of Singing* 55, no. 4 (March/April 1999): 41–42.

Kenneth Bozeman, tenor, holds performance degrees from Baylor University and the University of Arizona. He subsequently studied at the State Conservatory of Music in Munich, Germany on a fellowship from Rotary International. He is the Frank C. Shattuck Professor of Music and chair of the voice department at the Lawrence University Conservatory of Music in Appleton, Wisconsin, where he teaches voice, voice science, and pedagogy. He has received both of Lawrence University's Teaching Awards (Young Teacher Award, 1980; Excellence in Teaching Award, 1996), and an endowed chair in 1999.

Mr. Bozeman has a strong interest in the application of voice science to singing and received the 1995 Van L. Lawrence Fellowship Award from The Voice Foundation for "excellence in teaching and active interest in voice science and pedagogy." He has been a member of the editorial board of *Journal of Singing* for a number of years, serving as chair since 2000.

Mr. Bozeman has been an active performer of recitals and oratorio, including singing the tenor roles in the *St. Matthew* and *St. John Passions*, the *Christmas Oratorio*, the *B Minor Mass*, the *Magnificat*, and various cantatas of Bach, Handel's *Messiah*, Haydn's *Creation*, Mendelssohn's *Elijah*, and Vaughan Williams's *Hodie*. He has performed with the Milwaukee Symphony, the Wisconsin Chamber Orchestra, the Green Lake Music Festival, the Purgatory Music Festival of Colorado, the Louisville Bach Society, the Historical Keyboard Society of Wisconsin, and on Wisconsin Public Radio's "Live form the Elvehjem."