Making Grown Men Weep

1 author:

William O. Beeman
University of Minnesota Twin Cities

Beeman, William O. 2005. Making Grown Men Weep. In Hobart, Angela and Bruce Kapferer, eds., Aesthetics and Performance: The Art of Rite. New York: Berghahn Books, pp. 23-42.

Chapter One

MAKING GROWN MEN WEEP

William O. Beeman

Acoustics and Emotion

I once sang in a production of Puccini's *La Bohème*. In the fourth act, the heroine, Mimi, dies of consumption. The opera's final notes are delivered by her lover, the tenor Rodolfo, who bends over her lifeless body and sobs while singing her name four times on a high G. The effect is universally the same for all audiences. Almost as if a button were pushed, the scene triggers an autonomic response. Grown men and women weep openly. There is rarely a dry eye in the house. During one rehearsal for this production, our director had a problem with the soprano portraying Mimi. "My dear," he said, "you cannot cry when Rodolfo sings your name. You are already dead." "I know!" she wailed, "but I can't help it. It's *so sad*!"

It was this event some years ago that led me to think seriously about the unique ability of performance to affect the cognitive state of its audience, and the reasons why it does so. Added to this was my fascination with the brute fact that human beings in every culture are so extraordinarily engaged with performance. Not only do they enjoy it immensely but they also expend an amazing amount of energy and material resources to arrange for it to happen and to see it. Moreover, they never seem to tire of it. Specific performance experiences are revisited repeatedly—sometimes thousands of times over a lifetime, with no decrease in engagement or enthusiasm. Why this should be so somewhat defies logic and cries out for an explanation.

As I will argue below, singing is a uniquely human behavior. A particular, unusual effect occurs for listeners for particular types of singing that exhibit a special vocal acoustic property known among students of voice

Notes for this chapter are located on page 40.

as the "singer (or singer's) formant." The reaction is an autonomic physiological reaction such as occurred for our Mimi above. The singer's formant is a "spike" in the acoustic profile of the voice corresponding roughly to frequencies between 2,800–3,200 hertz (Hz). A more detailed explanation is presented below.

The singer formant is essential for trained voices singing in Western classical traditions. Without the ability to produce it, classical singers cannot have careers on the opera or concert stage. However, this vocal feature is not only a feature of Western vocal music. I hypothesize that it is also a characteristic of most professional singers in music traditions throughout the world. Although some singers produce it naturally, most have to learn to generate it in their vocal production, thus it is a performance achievement.

In addition to its aesthetic properties, the singer formant has the peculiar ability to produce an autonomic reaction from an audience. As I will show below, it appears to be an essential component in the conveyance of emotion in performance. Since expressing emotion is one of the principal functions of singing, this makes the singer formant one of the most interesting of human phenomena.

The singer formant thus presents a human mystery. Why should humans in so many cultural traditions identify a single particular preferred vocal acoustic characteristic and strive to produce it in artistic performance? Why should it produce autonomic reactions that so many identify as emotional? The answers must be hypothetical and speculative at this point, but as I will show in this chapter, the singer formant corresponds to the highest notes composed for the human voice. It also corresponds to the area of greatest sensitivity in the human hearing range. Its ability to produce an autonomic reaction suggests a primordial function. This I suggest exists in its presence in animal and infant cries, and in other human involuntary vocalizations.

Singing as an Evolutionary Act

It is noteworthy that humans especially enjoy seeing performance that emphasizes and underscores the limits of human behavior. Exemplary goodness, badness, and extremes of physical skill are among the most popular themes of performance.[1] It also shows reversal and transformation. In this regard, the mirror of performance is a fun-house mirror. It exaggerates, simplifies, and distorts in the subjunctive mode examined by Turner (1986). It holds the promise and wonder of witnessing things as they might be without the danger of the actual disruption that true change might bring. It is paradoxical that one of the most sophisticated of human cultural activities—singing—may derive its power to move people emotionally from affective expressive urges that predate our emergence as Homo sapiens. Yet our ability to sing and react to singing may be one of the most uniquely human things we are able to do as a species.

Delineating those behavioral capacities that are uniquely human has been a venerable task for students of human biology and culture for most of the twentieth century up to the present time. For a long time, tool making and linguistic communication were presented as the two activities that were the sole purview of humans.

In the last two decades we have learned much more about the behavioral and cognitive capacities of other animal species, particularly our nearest species cousins, the great apes. The research of Jane Goodall (1986), Sue Savage Rumbaugh (1994), and many others working with primate behavior in recent years has shown us that they have the capacity both for tool making and linguistic communication. Although the scientific community continues to split hairs evaluating the details, it is clear that human uniqueness is no longer defined unequivocally by these capacities. If we wish to understand human uniqueness, we may need to look to behavioral capacities that are still more complex than even these two rather complex behaviors.

The search for uniquely human behavior may depend on the exploration of the capacity of humans to engage in expressive communication, which may be defined as the overt and immediate conveyance of affect to others through public symbolic display. One of the primary means by which this is accomplished is through singing.

However, before looking at singing as a unique human activity, we should perhaps answer the more basic question: What value does expressive communication have for humans? One of the principal functions of expressive behavior would seem to be to encourage and facilitate bonding within human groups on a large scale, leading to more effective social organization.

Language itself is good at communicating information, but it is deficient in conveying affective states to others. Humans are able to accomplish a great deal of affective communication with tropic expression, such as metaphor (cf. Lakoff and Johnson 1974; Fernandez 1986), but even these structures lack immediacy. When humans really want to express interpersonal affect, language often breaks down. The deepest emotional expressions between two people, even hostile and violent ones, are usually tactile (perhaps also olfactory and gustatory) rather than linguistic, and this physical contact usually is a central component of bonding between individuals.

How is this bonding through sharing of inner states accomplished for whole groups? Most human societies find orgiastic behavior unpalatable or impractical. Untrammeled tactile intimacy leads also to social disturbance due to another factor in human social behavior: the need to establish hierarchies and the related competition for exclusive sexual partners.

Auditory and visual channels for communication have the advantage of being able to encompass and affect large numbers of individuals without the need to touch, smell, or taste every other person in the group. Normal language is of course primarily conveyed through auditory and visual channels. It is then not surprising that forms of communication conveying affect in an immediate manner have language as a component but provide significant enhancements from other dimensions of communication.

Singing is only one of a number of forms of behavior that qualify as enhanced communication. Two forms that are intermediate between speech and song are useful for understanding the power of singing. These are chant and oratory. Chant can be done by an individual, but it is most often a group behavior that strengthens solidarity among participants through shared vocal activity. It has an affinity in this regard with dance (as shared motor activity), with which it is often combined. The propositional content of chant is secondary to its acoustic properties and the fact that it is a shared activity. Chant is often a way of inducing broader shared experience in the form of trance or a deep meditative state as seen in the practices of religious orders. It often constitutes a profound experience for those engaging in it.

Oratory enhances communication through performative and poetic overdetermination of linguistic features (cf. Bauman 1977). The acoustic properties of speech (volume levels, pitch contours, pauses) and the structural properties of language (word choice, sentence structure, and logical interrelationship of expressions) are enhanced and exaggerated.[2] At the same time, there is a purposeful underdetermination of the "noise" found in ordinary face-to-face communication (errors, repetitions, interruptions). Oratory creates shared experience for a group by creating a common activity for audience members receiving the message of the orator. The underdetermination in noise in communication allows the audience to appreciate the overdetermination in the acoustic and structural properties of speech. These overdetermined features in oratory are the factors that communicate the affective messages of the speaker and allow the audience to share them as a group.

When it is understood that it is the nondiscursive and shared elements of vocal performance that bring about the transmission of affective content, it is a short leap to understand the power of singing—and singing is very powerful indeed.

Our most elementary vocalizations as a species both ontogenetically and phylogenetically relate to our affective states. Fear, pleasure, and discomfort rank as the most elementary of these expressions. Our nearest animal relatives possess call systems that articulate these affects. These calls are extended to express concern not just for the individual issuing the call but also for the group (cf. Fitch and Hauser 1995; Owren, Seyfarth, and Cheney 1997; Owren and Rendall 1999).

Humans also express themselves through powerful elementary vocalizations. Both the ability to produce these vocalizations and the ability to understand them may be genetically encoded, as may be seen from the first instant of birth with a baby's lusty crying on emerging from the womb. Babies are wonderful vocalizers. Their cries are penetrating, and they are able to continue for hours without getting hoarse or damaging their vocal apparatus. Their breath support system is perfect and natural. Sounds of pain or distress are deeply disturbing to adults; vocalizations of pleasure from babies are likewise directly communicative of joy to adults who are in contact with them (Drummond et al. 1999; Huffman et al. 1994; Lester 1978; Lester and Boukydis 1992).

Adults, too, have a remarkable repertoire of elementary vocalization forms, all tied to affect. Involuntary shouts of surprise, anger, or fear are as penetrating and physically efficiently produced as a baby's cries. Sounds of pleasure are equally involuntary on the part of the producers and recognizable on the part of hearers. A relaxing massage, a hot bath, eating a favorite food, or sexual activity all have characteristic, virtually involuntary vocalizations associated with them.

I hypothesize that singing is built on this human system of vocalizations. Singing produces an emotional response in listeners precisely because listeners are genetically programmed to respond to the acoustic properties of song, as they respond to other prelinguistic vocalizations, but in a more directed and differentiated fashion. Whereas response to elemental cries and vocalizations may be limited and diffuse, response to singing is directed and nuanced. The specific response is controlled through a delicate set of interactions between singer and audience that aims for fine, specific communication of affect. This additional power of singing comes from the ability to combine these powerful elemental vocal contours with other symbolic, discursive, poetic elements through text and visual means, including metaphor and narrative. This provides for a "double enhancement" of language that is effective and powerful as communication (Sundberg 1993, 1998; Sundberg, Iwarsson, and Hagegärd 1995; see also Feld 1982 and Feld 1988 for work in a non-Western context).

Singing in a way that produces an affective response in others is by no means an easy matter. Sadly, our inborn abilities as vocal artists are trained out of us early in our socialization. Children's exuberant cries are toned down by parents who exhort their offspring to use their "indoor" voices. Children with loud voices are often placed at a disadvantage in society. Added to the modulation of the voice in childhood training is the narrowing of accepted emotional expression during socialization. Display of affect is confined within narrow boundaries in most cultures, and only allowed full, untrammeled public expression on limited occasions.[3]

The Singer Formant

Singers in virtually every music tradition must learn to produce a culturally meaningful acoustic envelope for their vocalizations. This means finding a means of regulating their vocal apparatus to produce an acoustically pleasing sound. For some societies this sound will involve more resonance in the nasal cavities, for some it will involve more throat constriction. In modern times singers may get help from electronic enhancement to achieve the aesthetic ideal of their culture.

For singers in the Western classical tradition, this aesthetic involves three essential elements aside from the musical requirements that one sing on pitch and with rhythmic accuracy. The first, which I will deal with in this section, is the production of the so-called singer formant, a consistent high-frequency sound wave component to the voice that allows the voice to "carry" in a large hall over a symphony orchestra. This is the vocal feature that allows singers to perform at the Metropolitan Opera and other large houses without amplification.

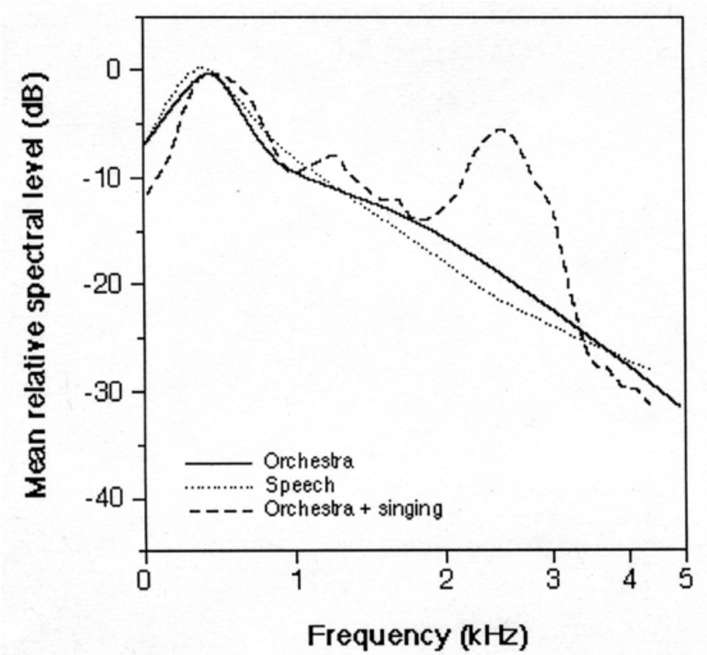
Second, the singer must execute "line." This is the ability to sing long overarching musical phrases without perceptible interruption due to the consonants within words. Third, the singer must sound as if he or she is singing with the same vocal quality on every pitch and on every vowel. This final ability is one of the most difficult goals to achieve, because different muscle groups control pitch and resonance in different parts of the voice, and their coordination to make the voice sound uniform is challenging.

The first of these three elements—the production of the singer formant—is the most basic of the three. It must be mastered first. Line and uniformity of production depend on the muscle control needed to produce the singer formant, which carries, I hypothesize, the basis for essential affective expression in song and allows singing to communicate to the largest number of people. This is also the one element of singing that seems to be common to the vocal traditions of the world. It seems to be the aspect of vocal art that unites the adult singer with the crying infant, the person screaming in grief or fear, the individual doubled over in involuntary laughter, and the child squealing in delight at a pleasant surprise.

The "singer formant"[4] gives the trained voice its acoustic "ping," or *squillo* in the Italian term used in singing instruction. It has been ably described in an extensive literature, spearheaded by the pioneering research of Johan Sundberg (Sundberg 1972, 1977, 1987, 1998; see also Sataloff 1992; Hong 1995; Titze, Mapes, and Story 1994). This acoustical property, as mentioned above, is a "spike," or high decibel feature, in the acoustic profile of a singer occurring at between 2,800 and 3,200 Hz. As some researchers have pointed out, and as I will cite below, the spike may be produced at higher frequencies, up to 4,000 Hz. In lay terms, the singer formant may be thought of as a consciously reinforced cluster of overtones.[5]

Sundberg (1972, 1977) studied recordings of the famous operatic tenor Jussi Björling, who was particularly admired for the brilliant "ring" in his voice. Sundberg discovered the singer formant in Björling's voice at 3,000 Hz and determined that it was especially strong when he was singing over a loud orchestra. The graph in figure 1.1 represents Sundberg's findings. The solid line is the average frequency spectrum of the

Figure 1.1 Average frequency spectrum of tenor Jussi Björling plotted against the average frequency spectrum of the orchestra accompanying him



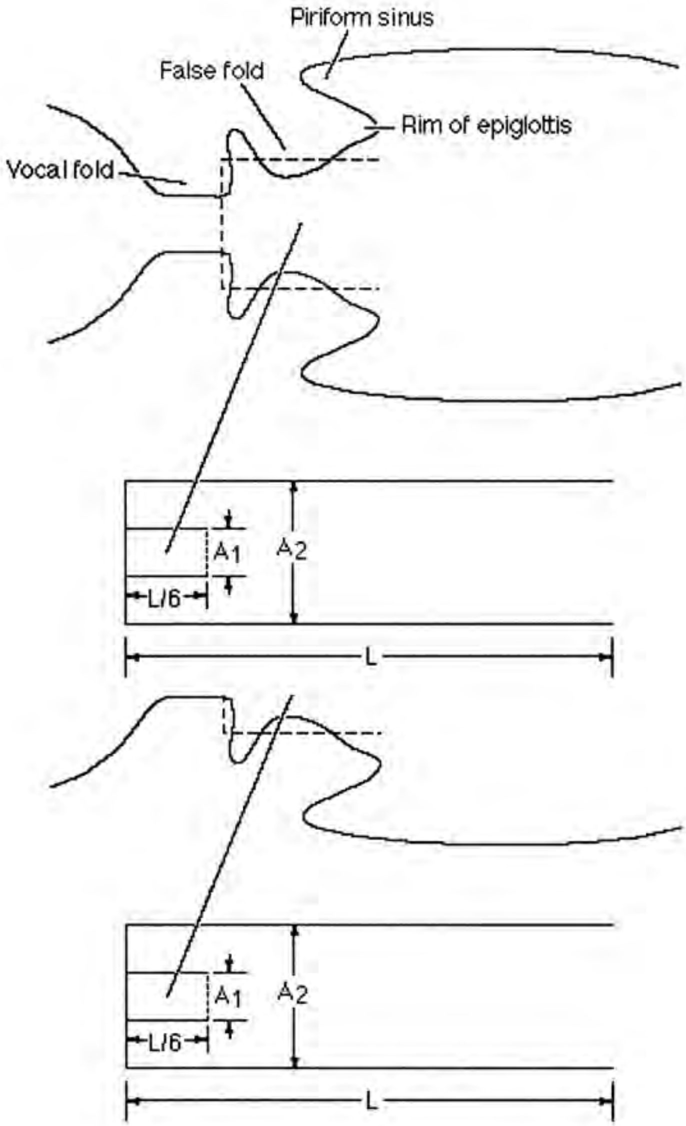
Source: After Sundberg (1972, 1974).

symphony orchestra, and the dotted line, the average frequency spectrum of Björling's voice.

Because of this increased energy clustering at 3,000 Hz, the singer can easily be heard above the orchestra. The singer's voice is said to "cut through" the orchestra. It is noteworthy that the ability of the singer to be heard above the orchestra can be seen as independent of the volume of the fundamental frequency. A trained singer can be heard above an orchestra even when singing softly, because the energy at the point of the singer formant can be made higher than that of the orchestra. As I will show below, the singer formant gets an additional boost, because the human ear hears better at 3,000 Hz than at lower frequencies.

Sundberg discovered that the origin of the singer formant is the presence of a second resonator in the vocal tract, in addition to the vocal folds. The singer generates this second resonator through muscle control of the vocal tract. To generate the singer formant, this second resonator must be about one-sixth as long as the entire vocal tract and also have about one-sixth of the cross-sectional area of the vocal tract. Figure 1.2 shows Sundberg's model for the operation of this mechanism.

Figure 1.2 Model of primary and secondary resonators in vocal tract producing the “singer formant”



Source: Sundberg (1972).

As can be seen from figure 1.2, the "false folds" are used to create the secondary resonator. These are controlled by involuntary muscles. This secondary resonator is naturally activated when humans scream in fright or fear, and as I hypothesize, they are also naturally brought into play in infant cries. Learning to control them voluntarily to create performative artistic expression is part of the aim of vocal training.

Singer Training

Singers must train long and hard to develop control over this secondary resonator and to produce the even, uniform vocal line and vocal registration required for aesthetically pleasing singing. Normal training takes a number of years[6] and is extremely demanding for most singers with professional aspirations in most classical traditions, including Western classical music.

Instruction is athletic in nature—akin to Olympic training in its demands for perfection. It involves the training of muscles throughout the body. The primary muscles are those of the mouth, throat, and breathing apparatus, but in effect every part of the body must be involved in the singing process. In order to produce smooth and even tones, the singer must deliver a steady stream of air at precisely the right pressure to both the vocal folds and the secondary resonators. The muscles that control the resonators themselves must be trained to engage them with exactly the right tonus. Too strenuous an application can create vocal injury; too light, and the proper acoustic profile will not emerge.

As with an oboe reed or trumpet mouthpiece, the singer cannot over- or underblow or the vocal tone will lose its essential quality. In order to achieve even more resonance and color, the singer must activate other parts of the vocal tract, including the nasal cavities. In some vocal traditions, such as those of Tuva (Levin and Edgerton 1999), even more complex formant structures can be produced with proper manipulation of the tongue and throat.

Singers must also learn to overcome inhibitions that prevent control of the delicate vocal musculature and that keep them from producing full and expressive sounds in public. The singer is exposed and defenseless before the public. Thus, there may be involuntary reflexes such as "fright or flight" adrenaline production that need to be brought under control through the training process lest they hinder the free and relaxed working of the voice. Meditative techniques that involve mental focus and bodily flexibility such as ta'i chi and yoga are used by many to aid in overcoming these difficulties (cf. Helfgot and Beeman 1993).

Vocal line and uniform vocal registration make singing seem speechlike. We talk in unbroken phrases, and when speaking our voices have the same characteristics throughout our speech range. The difference for the singer is that the need for articulation is greater than that of the orator, and the pitch ranges that are used in singing are three or four times more extensive than those of the normal speaking voice. Achieving conversational smoothness while operating in these nonconversational acoustic parameters is a great artistic challenge.

The length of training for a singer is extreme because, paradoxically, singers cannot hear themselves. The singing voice is filtered through the bones of the head and is misleading to the singer. From the singer's own perspective, a seemingly big sound may in fact lack the essential singer formant; an apparent legato line or acoustically equalized vowels may sound jerky or uneven to an audience. Eventually singers learn to work not by hearing themselves but by being trained by expert teachers to monitor the physical sensations of their bodies to produce the best sounds. They in fact learn to feel their voice rather than to hear it.

The effects of singing for the singer are far more visceral than for the audience. When singing correctly, the singer achieves a physical release that might best be compared to the rush of an Olympic athlete executing a winning floor routine in gymnastics or achieving a record-setting high jump. Moreover, as with Olympic athletes, these peak experiences may be rare, but when they occur they are powerful enough to motivate the singer to try and achieve them again and again, even at great personal sacrifice.

When the singer reaches these goals, adding to them the emphatic nuances of the orator and the facial and bodily expressions of the best actor, the experience is overwhelming for an audience. The communication of affect is complete, and the audience is united in a reaction of empathy and common understanding. So all in attendance cry at Rodolfo's grief, audience and singers alike (even the recently deceased Mimi). Everyone leaves the theater with the certain knowledge that for a brief moment they have shared something remarkable—something both elementary and primitive and yet at the same time sophisticated and uniquely human.

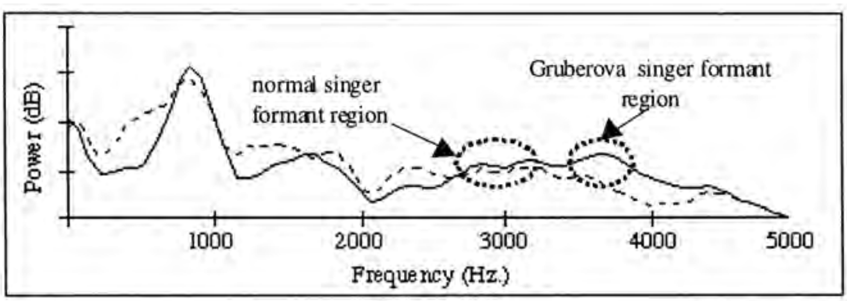
Thus, for singers the training is severe and difficult. The chances for success are extremely slim. The personal rewards for achieving vocal skill are great, but perhaps not in proportion to the effort. Nevertheless, as I have already noted several times, both singers and audiences find the activity itself physically as well as mentally fulfilling to the point that both expend enormous effort to experience it. So, the question remains, Why should humans so highly privilege vocal production that emphasizes the singer formant? I will try to answer this question in the next sections of this discussion.

The Singer Formant and Emotion

There is some evidence that the singer formant plays an important role in the registering of emotion on the part of an audience. Although most researchers have associated the conveyance of emotion in singing with prosody, phrasing, and tempo variation (Sundberg 1998; Sundberg, Iwarsson, and Hagegärd 1995; Flanagan 1981), a few studies have shown the singer formant to be essential in emotional expression (Johnstone and Scherer 1995; Siegwart and Scherer 1995; Scherer 1986; Berndtsson and Sundberg 1995).

Johnstone and Sherer compared recordings of the cadenza in the aria "Ardi gli insensi" from Donizetti's *Lucia di Lammermoor.* Working from earlier work by Siegwart and Scherer, they wanted to find out why, of five famous sopranos sampled, listeners preferred the rendition by soprano Editha Gruberova, finding that her rendition expressed more "tender passion" and "sadness" than the other singers. To the surprise of the researchers, they found that Gruberova did not exhibit high energy in the area of the singer formant. Rather, she exhibited energy at a higher frequency—namely, at approximately 3,600 Hz. A comparison of Gruberova's sound spectrum and that of Joan Sutherland is contained in figure 1.3.

Figure 1.3 Vocal spectrum of Editha Gruberova (*solid line*) and Joan Sutherland (*broken line*) rendering the cadenza in the aria "Ardi gli incenti" from Donizetti's *Lucia di Lammermoor*



Source: After Johnstone and Sherer (1995).

As Johnstone and Scherer (1995) concluded, "the singer's formant is not a single formant as such, but rather a clustering of formants around a predicted frequency of about 3,000 Hz (in sopranos). When clustered sufficiently closely, individual formants tend to reinforce each other, leading to a spectral region with increased overall resonance." Gruberova, in effect, had "relocated" her singer formant region to a higher frequency, and this was read by the judges as conveying greater emotion.[7]

Johnstone and Scherer have exhibited two important principles. The first is that the singer formant has a complex interaction with harmonics (overtones) and that it can be created at frequencies higher than the canonical 2,800–3,200 Hz and convey even more emotional content.

We will see these two principles at work when analyzing the cries of infants below, but first we need to ask why humans react to frequencies in this range with particular attention. For this we need to look at classic research into human hearing.

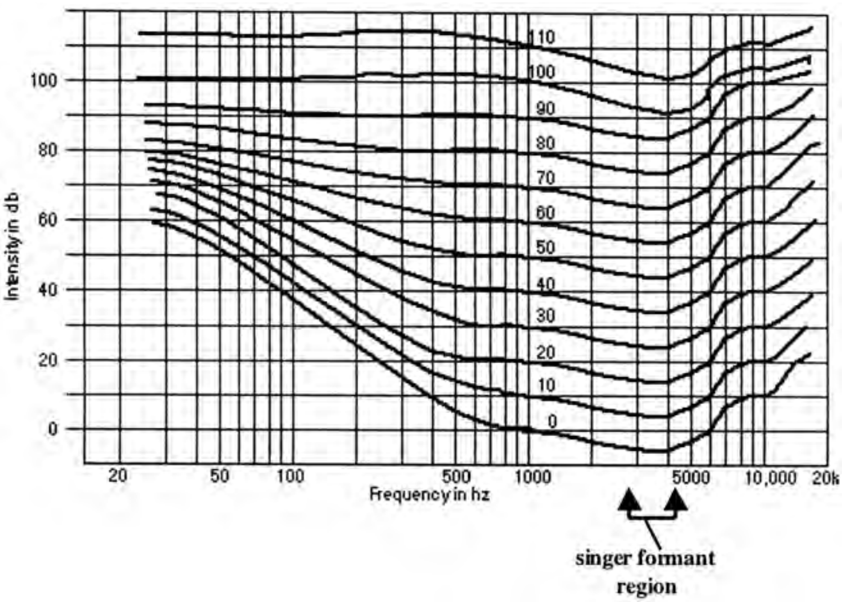
The Fletcher-Munson Curves

One part of the answer to the above question involves the properties of human hearing. Humans do not hear all frequencies of sound at the same level. The measurement of loudness of sound is a highly subjective phenomenon, but it was first described in a spectacularly successful manner by Fletcher and Munson (1933). The Equal Loudness Contours reveal the average human hearing sensitivity at frequencies in the human hearing range over various listening volumes. They were determined something like this: Fletcher and Munson played a reference tone at, for example, 1,000 Hz, at a designated sound volume. They would then play a different frequency, adjust the volume, and ask the listener to indicate when the new frequency sounded like it was at the same volume as the original frequency. They continued this process through several frequencies to come up with a "contour" of human hearing sensitivity at that one reference sound volume, say at 80 decibels (dB) sound pressure level. Next they would change to a reference volume of 85 dB and repeat the process.

They averaged this data over several listeners to come up with their published Equal Loudness Contours—commonly called today the Fletcher-Munson Curves. In this way, they demonstrated that human ears are more sensitive to some frequencies and less sensitive to others and that the sensitivity changes with the sound pressure level. More importantly, they were able to calibrate the actual energy in a sound as measured in decibels with the perceived loudness of the same sound. Figure 1.4 shows the changes in the sensitivity level of the ear over a range of frequency levels; energy is expressed here in decibel intensity along the vertical access and frequency on the horizontal axis shown as a logarithmic scale.

Each curve in figure 1.4 represents a gradient of ten *phons,* a subjective measure of loudness corresponding to decibel levels when measured by physical measurement equipment. From about 500 Hz to roughly 1,500 Hz, the 10 phon line roughly corresponds to 10 dB. This means that for humans to perceive the sound being a loudness level of 10 phons, frequencies from 500 Hz to 1,500 Hz must have an energy level of 10 dB. However, at 5,000 Hz the 10 phon line dips. This shows that the human ear perceives 4,000 Hz to be 10 phons when the source is actually only approximately 5 dB. One gets twice the effect for the energy expended. At 4,000 Hz, the phon curves turn up. To perceive 10,000 Hz at the same level of 10 phons, the energy level needs to be approximately 20 dB.

Figure 1.4 Fletcher-Munson equal loudness contours



Note: Each curved line on the chart represents a gradient of ten *phons*, a subjective measure of loudness.

Source: Fletcher and Munson (1933).

Fletcher and Munson's research shows that the ear is less sensitive for low frequencies and at frequencies above 5,000 Hz. For this reason manufactures of sound equipment put "bass boosters" on their amplifiers to compensate for the lower sensitivity of human hearing. Subwoofers and tweeters (to boost the extreme high frequencies) as specialized loudspeaker equipment accomplish the same end.

For the purposes of this discussion, however, the most interesting fact revealed by the curves is that human hearing is more sensitive in the 2,800 kHz to 4,000 kHz range. This range of intense sensitivity in hearing corresponds to the region of the singer formant. Thus we can conclude that, at least in terms of hearing sensitivity, humans are uniquely attuned to hearing the singer formant.

This cannot be mere happenstance. The coincidence of the primary acoustic feature of the primary form of human music making and the dominant characteristic of human auditory sensitivity suggests an evolutionary adaptation. Any conclusions I might draw, like all statements about the evolution of human behavior, must be hypothetical and speculative. Nevertheless, I believe a compelling case can be made, as I suggested at the beginning of this discussion, for the origin of the singer formant, if not singing itself, in infant cries, primate calls, and involuntary human vocalizations. All of these nonspeech vocal behaviors have one thing in common: the expression of emotional states, and this is the essence of singing as a communicative performative behavior.

Infant Cries

The study of infant crying has produced an exceptionally large body of literature as cited above (cf. Drummond et al. 1999; Huffman et al. 1994; Lester and Boukydis 1992; Lester 1987; Lester 1978; Thompson and Olson 1996). Much of the clinical literature on infant cries has focused on the acoustic characteristics of the cry and its possible relationship to brain damage as a result of smoking and alcohol consumption (Nugent et al. 1996), infant drug addiction (Lester, Boukydis, and Twomey 2000; Lester 1999; Lester and Tronick 1994) and sudden infant death syndrome (Corwin et al. 1995, Lester, Corwin, and Golub 1988).

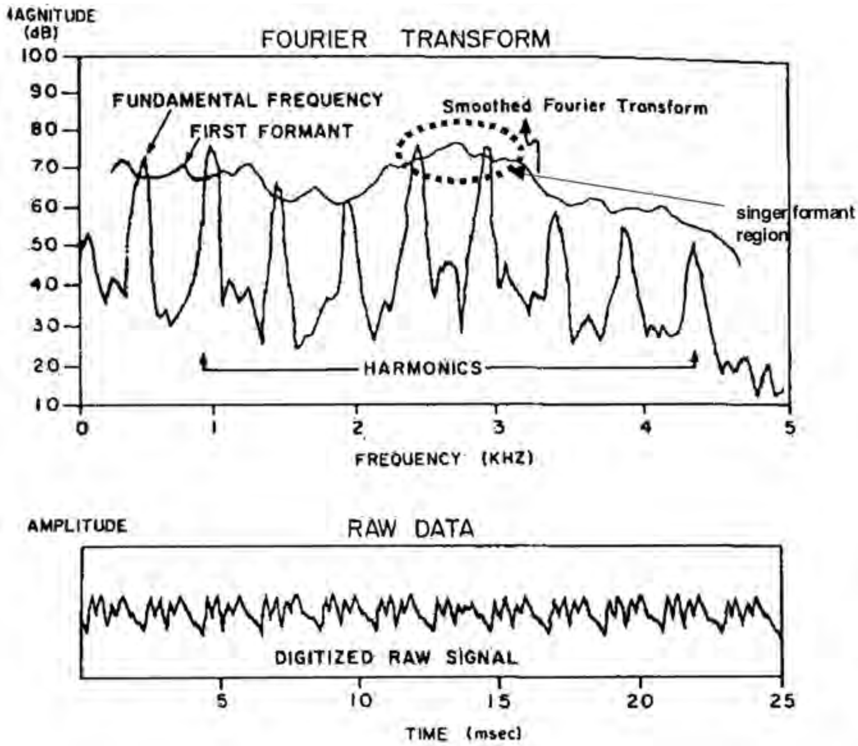
Infant crying causes extraordinary reactions in adults. As mentioned above, infants are astonishing vocalists. They never tire and can produce piercing cries for hours on end. Few adults can tolerate these vocalizations for a long period of time, and this can lead to parental distress. In relatively benign cases it causes parents to lose sleep as they repeatedly rise in the middle of the night to comfort and calm their baby. However, the adult reaction can have effects that are more deleterious. As parents' nerves become frayed, their reactions to their child's crying can lead to infant battering and even infanticide. For this reason, a number of researchers have focused on the structure of bouts of crying and adult reactions to them (Lester 1984; Bisping et al. 1990; Lester, Garcia-Coll, and Valcarcel 1989; Huffman et al. 1994; Thompson and Olson 1996; Green, Gustafson, and McGhie 1998).

In nearly all studies of infant crying to date, the fundamental frequency of the cry, that is, the basic pitch (F0), is taken as its primary acoustic characteristic. Researchers have generally tried to categorize cry types (fright, hunger, illness, etc.) and gauge parent reactions according to variations in F0 (Gustafson and Green 1989; Bisping et al. 1990; Protopapas and Eimas 1997; Lenti-Boero et al. 1998). I wish here to suggest that the salient element in infants' cries eliciting reaction from adults is not (or is not exclusively) F0 but rather the presence of prominent energy in the area of the singer formant.

Lester provides a spectral analysis of an infant cry that shows clearly the spike of energy characteristic of the singer formant. This is shown in figure 1.5.

Lester's spectrum of the infant cry in question shows the characteristic clustering of mutually reinforcing frequencies in the singer formant region noted by Johnstone and Scherer, resulting in the formant "spike" or "hump" in the smooth Fourier transform of the overall spectrum.

Figure 1.5 Smooth Fourier analysis of infant cry

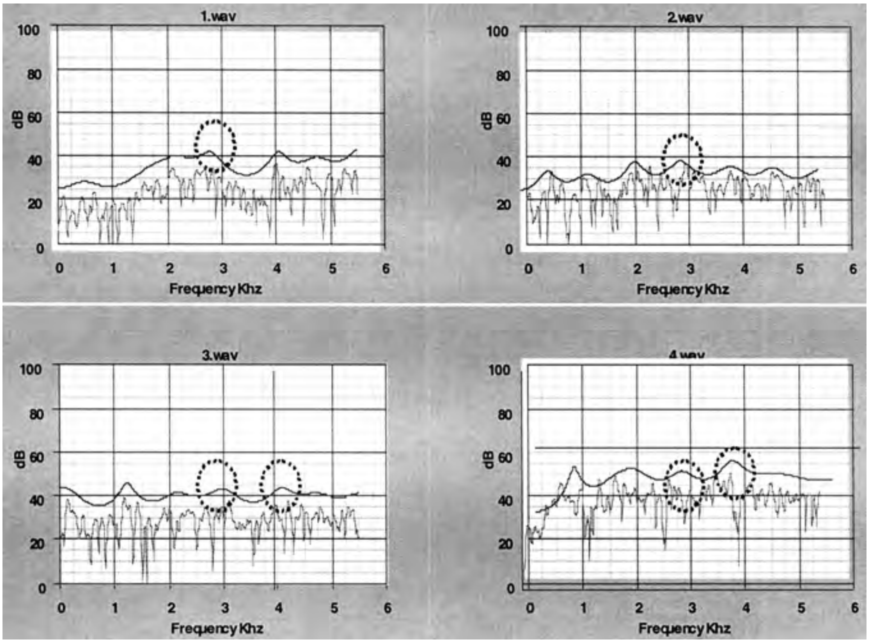


Source: Lester (1987: 531).

To test this phenomenon further, I ran four examples of infant cries (very kindly supplied by J. A. Green of the University of Connecticut, whose work on the acoustics of infant cries has already been noted in this discussion). These were provided as digital computer files and analyzed through a Fourier transform program, The BLISS program developed at Brown University.[8] Figures 1.6-1 through 1.6-4 show the results of this analysis.

In the four analyses of infant cries, we see a clustering of energy at the area of the singer formant in all four examples, thus replicating Lester's results with independent data. Additionally, in examples 1.6-3, and particularly 1.6-4, we see an additional area of intense energy corresponding to the "Gruberova" formant noted by Johnstone and Scherer. An additional, intriguing body of research on primate cries reveals that these too exhibit strong energy in the singer formant region (Owren, Seyfarth, and Cheney 1997; Owren and Linker 1995). Figure 1.7 shows this clearly. Owren, Seyfarth, and Cheney go on to point out that the frequencies in the singer formant region are even higher and exhibit stronger energy when the cry is associated with an infant (1997, 2956, 2960).

Figure 1.6-1, 1.6-2, 1.6-3, 1.6-4 Fourier transforms of four infant cries recorded by J. A. Green as analyzed by the BLISS speech analysis programs



Note: Singer formant region is indicated by a dotted oval. In the case of cries 3 and 4, a second "Gruberova" formant region is indicated at approximately 3800–4100 Hz.

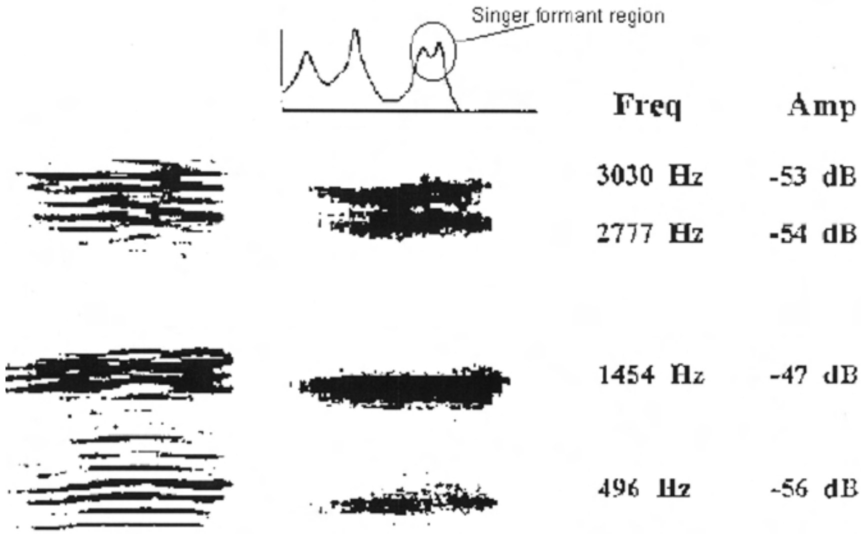
Source: **[[Please supply if applicable.]]**

Hearing, Emotion, and Performance

Joseph Ledoux has associated autonomous emotional response with particularly strong auditory stimulus (1994, 162–65). The hearing channel seems particularly hardwired to the thalamus and the amygdalar regions, particularly for reactions associated with fear. As I have shown above, the areas of hearing that are most sensitive in human beings also correspond to the singer formant region of vocal energy.

In considering babies' cries and baboon articulations, it seems that both contain strong concentrations of energy in the singer formant area. This suggests that evolutionarily, selectivity favored creatures who could both generate and hear accurately sound energy at this particular high frequency. What is especially interesting about the singer formant region is that it does not interfere with the comprehension of normal speech, the formants for which are at a much lower frequency level.

Figure 1.7 [[Please supply title]]



Source: **[[Please supply if applicable.]]**

This suggests that humans evolved to be able to both comprehend the speech of others and comprehend the emotional tone of speech, particularly when it consisted of strong affective dimensions such as fear.

At least one research team has suggested that there is even more going on. Babies may be actually performing when they cry. Moreover, they may be manipulating or fooling their parents, who are often powerless to resist them when they produce high-pitched wailing (Thompson and Olson 1996). Thus as students of performance we are tempted to the conclusion that crying may be the first real human performance. Moreover, it may be one of the most effective forms of performance in which humans ever engage. A crying baby nearly always gets a strong emotional reaction from its audience—perhaps not as pleasurable a reaction as that effected by the high notes of a romantic operatic tenor, but a rousing response nevertheless.

Notes

1. Wise theatrical producers who want to attract the attention of young people put children in their productions. If they are used well, the children virtually guarantee the rapt attention of other children in their audience.
2. This is the "poetic function" so ably analyzed by Jakobson (1960).
3. Performers frequently express their pleasure at performing for children's audiences because of their exuberant response. My suspicion is that children scream and shout at all-children's concerts and stage performances precisely because this is one of the few occasions where adults approve of this behavior, and do nothing to control it.
4. Also variously called the "singing formant" or "singer's formant"
5. This is a simplification. Every basic tone has an overtone series that occurs "naturally." In general the overtone sequence for a particular fundamental frequency consists of energy peaks at regular higher intervals—generally multiples of the fundamental frequency, usually with the same, or decreasing energy as the fundamental frequency. The singer formant, as I have suggested by the term "spike" is a shorthand term for an unusual cluster of high energy, expressed in decibels, seen in a particular set of wave frequencies clustering at 2800–3200 Hz [and, as noted, sometimes ranging to 4100 Hz]. As will be seen, the energy of the singer formant may actually be greater than the fundamental frequency that underlies it (cf. Sound 1999). The singer "formant" is revealed by determining the average frequency spectrum of the voice. For those with a mathematical bent, this formant is indicated by a "hump" in a graph of a smoothed Fourier transform of the entire vocal spectrum of a given sound sample. Naturally the Fourier transform will change with the size and selection of the sample. What is interesting about the singer formant is its constant presence in the voice of trained singers even over very large samplings of sound. Those who know the term "formant" from the study of speech may find this terminology confusing. In principle, however, although the singer formant is not a component of speech sounds, it is analytically identified in the same way as the formants of speech.
6. In my own case, training at a major conservatory and with private teachers, it took about seven years to develop consistent control.
7. Johnstone and Sherer report that "long term spectra of all Gruberova recordings displayed a high energy region between 2900 and 4100 Hz (1995).
8. This program is available free of charge to researchers in acoustics through the Department of Cognitive and Linguistic Sciences, Brown University.

References

Bauman, R. 1977. *Verbal Art as Performance*. Rowley, Mass.: Newbury House.

Berndtsson, G., and J. Sundberg. 1995. "Perceptual Significance of the Centre Frequency of Singer's Formant." *STL-QPSR* (KTH) (Stockholm) 4: 95–105.

Bisping, R., H. J. Steingrueber, M. Oltmann, and C. Wenk. 1990. "Adults' Tolerance of Cries: An Experimental Investigation of Acoustic Features." *Child Development* 61, no. 4 (August): 1218–29.

Corwin M. J., B. M. Lester, C. Sepkoski, M. Peucker, H. Kayne, and H. L. Golub. 1995. "Newborn Acoustic Cry Characteristics of Infants Subsequently Dying of Sudden Infant Death Syndrome." *Pediatrics* 96, no. 1, pt. 1 (July): 73–77.

Drummond, Jane, Nicole LeTourneau, Susan Neufeld, Harriet Harvey, M. Ruth Elliott, and Sandra Reilly. 1999. "Infant Crying and Parent-Infant Interaction: Theory and Measurement." *Infant Mental Health Journal* 20, no. 4: 452–65.

Feld, Steven. 1982. *Sound and Sentiment: Birds, Weeping, Poetics and Song in Kaluli Expression*. Philadelphia: University of Pennsylvania Press.

———. 1988. "Aesthetics as Iconicity of Style, or 'Lift-Up-Over Sounding': Getting into the Kaluli Grove." *Yearbook for Traditional Music* 20: 74–113.

Fernandez, James. 1986. *Persuasions and Performances: The Play of Tropes in Culture*. Bloomington, Ind.: Indiana University Press.

Fitch, W., and M. Hauser. 1995. "Vocal Production in Nonhuman Primates: Acoustics, Physiology, and Functional Constraints on 'Honest' Advertisement." *American Journal of Primatology* 37: 191–219.

Flanagan, J. L. 1981. "Emotions, Voice, and Music." In *Research Aspects of Singing,* 51–79. Stockholm: Royal Swedish Academy of Music.

Fletcher, H., and W. A. Munson. 1933. "Loudness, Its Definition, Measurement and Calculation." *Journal of the Acoustical Society of America* 5: 82–108.

Goodall, Jane. 1986. *The Chimpanzees of Gombe: Patterns of Behavior*. Cambridge, Mass.: Harvard University Press.

Green, J. A., G. E. Gustafson, and A. C. McGhie. 1998. "Changes in Infants' Cries as a Function of Time in a Cry Bout." *Child Development* 69, no. 2 (April): 271–79.

Gustafson, G. E., and J. A. Green. 1989. "On the Importance of Fundamental Frequency and Other Acoustic Features in Cry Perception and Infant Development." *Child Development* 60, no. 4 (August): 772–80.

Huffman, Lynne C., Yvonne E. Bryan, Frank A. Pedersen, Barry M. Lester et al. 1994. "Infant Cry Acoustics and Maternal Ratings of Temperament." *Infant Behavior and Development* 17, no. 1: 45–53.

Jakobson, Roman. 1960. "Closing Statement: Linguistics and Poetics." In *Style in Language*, ed. Thomas Sebeok, 350–77. Cambridge, Mass.: MIT Press.

Johnstone, T., and K. Scherer. 1995. "Spectral Measurement of Voice Quality in Opera Singers: The Case of Gruberova." In *Proceedings of the XIIIth International Congress of Phonetic Sciences*, vol. 1, ed. K. Elenius and P. Branderud. Stockholm.

Lakoff, George, and Mark Johnson. 1974. *Metaphors We Live By*. Chicago: University of Chicago Press.

Ledoux, Joseph. 1994. **[[Please supply missing info.]]**

Lenti-Boero, D., C. Volpe, A. Marcello, C. Bianchi, and C. Lenti. 1998. "Newborns Crying in Different Contexts: Discrete or Graded Signals?" *Perceptual and Motor Skills* 86, no. 3, pt. 2 (June): 1123–40.

Lester, Barry M. 1978. "The Organization of Crying in the Neonate." *Journal of Pediatric Psychology* 3, no. 3: 122–30.

———. 1984. "A Biosocial Model of Infant Crying." *Advances in Infancy Research* 3: 167–212.

———. 1987. "Developmental Outcome Prediction from Acoustic Cry Analysis in Term and Preterm Infants." *Pediatrics* 80, no. 4 (October): 529–34.

———. 1999. "To Covary or Not to Covary: What Is the Question?" *Journal of Drug Issues* 29, no. 2: 263–68.

Lester, Barry M., Michael Corwin, and Howard Golub. 1988. "Early Detection of the Infant at Risk through Cry Analysis." In *The Physiological Control of Mammalian Vocalization*, ed. John D. Newman, 395–411. New York: Plenum Press.

Lester, Barry M., Cynthia T. Garcia-Coll, and Marta Valcarcel. 1989. "Perception of Infant Cries in Adolescent and Adult Mothers." *Journal of Youth and Adolescence* 18, no. 3: 231–43.

Lester, Barry M., and C. F. Zachariah Boukydis. 1992. "No Language but a Cry." In *Nonverbal Vocal Communication: Comparative and Developmental Approaches,* ed. Hanus Papousek and Uwe Juergens, 145–73. Studies in Emotion and Social Interaction. New York: Cambridge University Press.

Lester, Barry M., and Edward Z. Tronick. 1994. "The Effects of Prenatal Cocaine Exposure and Child Outcome." *Infant Mental Health Journal* 15, no. 2: 107–20.

Lester, Barry M., C. F. Zachariah Boukydis, and Jean E. Twomey. 2000. "Maternal Substance Abuse and Child Outcome." In *Handbook of Infant Mental Health,* 2nd ed., ed. Charles H. Zeanah, Jr., 161–75. New York: The Guilford Press.

Levin, Theodore C., and Michael E. Edgerton. 1999. "The Throat Singers of Tuva." *Scientific American* 281, no. 3 (September): 80–82.

Nugent, J. K., B. M. Lester, S. M. Greene, D. Wieczorek-Deering, and P. O'Mahony. 1996. "The Effects of Maternal Alcohol Consumption and Cigarette Smoking during Pregnancy on Acoustic Cry Analysis." *Child Development* 67, no. 4: 1806–15.

Owren, Michael J., Robert M. Seyfarth, and Dorothy L. Cheney. 1997. "The Acoustic Features of Vowel-like Grunt Calls in Chacma Baboons (Papio cyncephalus ursinus): Implications for Production Processes and Functions." *Journal of the Acoustic Society of America* 101, no. 5, pt. 1 (May): 2951–63.

Owren, Michael J., and D. Rendall. 1999. "An Effect-Conditioning Model of Nonhuman Primate Vocal Signaling." In *Perspectives in Ethology*, ed. D. H. Owings, M. D. Beecher, and N. S. Thompson. Vol. 12. New York: Plenum Press.

Protopapas, A., and Peter D. Eimas. 1997. "Perceptual Differences in Infant Cries Revealed by Modifications of Acoustic Features." *Journal of the Acoustic Society of America* 102, no. 6 (December): 3723–34.

Sataloff, Robert T. 1992. "The Human Voice." *Scientific American* (December).

Savage-Rumbaugh, Sue. 1994. *Kanzi: The Ape at the Brink of the Human Mind*. New York: John Wiley & Sons.

Scherer, K. R. 1986. "Vocal Affect Expression: A Review and a Model for Future Research." *Psychological Bulletin* 99, no. 2 (March): 143–65.

Siegwart, H., and K. R. Scherer. 1995. "Acoustic concomitants of emotional expression in operatic singing: The case of Lucia in Ardi gli incensi." *Journal of Voice*.

Sound. 1999. Encyclopedia Britannica Online. http://www.eb.com:180/bol/topic?eu=117555&sctn=38: Encyclopedia Britannica.

Sundberg, Johan. 1972. "Production and Function of the Singing Formant." In *Report of the 11th Congress of the International Musicological Society*, vol. 2, ed. H. Glahn, S. Sorenson, and P. Ryom, 679–88. Copenhagen: Edition Wilhelm Hansen.

———. 1977. "The Acoustics of the Singing Voice." *Scientific American* (March): 82–91.

———. 1987. *The Science of the Singing Voice*. Dekalb, Ill.: Northern Illinois University Press.

———. 1993. "How Can Music Be Expressive?" *Speech Communication* 13, no. 1/2: 239–53.

———. 1998. "Expressivity in Singing: A Review of Some Recent Investigations." *Log Phon. Vocol*. 23: 121–27. Paper presented originally at the PEVOC-II conference, August 29–31, 1997 in Regensburg, Germany.

Sundberg, Johan, J. Iwarsson, and H. Hagegärd. 1995. "A Singer's Expression of Emotions in Sung Performance." In *Vocal Fold Physiology: Voice Quality and Control*, ed. O. Fujimura and M. Hirano, 217–32. San Diego: Singular Publishing Group.

Thompson, Nicholas S., and Carolyn Olson. 1996. "Babies' Cries: Who's Listening? Who's Being Fooled." *Social Research* 63, no. 3: 763–85.

Titze, Ingo R. 1994. *Principles of Voice Production*. Englewood Cliffs, N.J.: Prentice-Hall.

Titze, Ingo R., and Sharyn Mapes. 1992. "Acoustics of the Tenor High Voice." *Journal of the Acoustical Society of America* 91, no. 4: 24–33.

Titze, Ingo R., Sharyn Mapes, and Brad Story. 1994. "Acoustics of the Tenor High Voice." *Journal of the Acoustical Society of America* 95, no. 2: 1133–42.

Turner, V. 1986. *The Anthropology of Performance*. New York: PAJ Publications.