

Publications of Dr. Martin Rothenberg:

Source-Tract Acoustic Interaction in the Soprano Voice and Implications for Vocal Efficiency

Originally titled: *Così Fan Tutte* and What it Means, or, Nonlinear Source-Tract Acoustic Interaction in the Soprano Voice and Some Implications for the Definition of Vocal Efficiency

Paper presented at the 4th International Conference on Vocal Fold Physiology, New Haven Ct, June 3-6, 1985. Reprinted in the proceedings of the conference, *Vocal Fold Physiology: Laryngeal Function in Phonation and Respiration*. T.Baer, C. Sasaki, and K.S. Harris, eds., College Hill Press, San Diego, pp. 254-263, 1986.

Introduction

It is now generally accepted that interaction between the laryngeal sound source and pressure variations in the vocal tract during the glottal cycle can be an important factor in determining voice quality. This interaction can affect both the vibratory pattern of the vocal folds and the glottal airflow pattern for a given vocal fold vibratory pattern. My work has centered around the effect on the glottal airflow for a given vocal fold motion (i.e., the source-tract acoustic interaction), and this will be the focus in this paper as well. I could also mention that the emphasis in this paper will be on an intuitive development of the concepts involved, not on a precise analytic or numerical solution of the differential equations.

A number of studies, including my own, have investigated the ways in which the inertance of the vocal tract airflow tends to modify the waveform of the glottal airflow pulse so as to increase the proportion of energy generated at the closing of the vocal folds. However, the inertance of the vocal tract airflow becomes the clearly dominant factor in source-tract acoustic interaction only when the lowest formant frequency, F1, is at least about three times the voice fundamental frequency, F0.

F1 Tuning in the Soprano Voice

What happens when F1 is much less than three times F0? To help answer this question, let us consider the upper range of a soprano, in which F0 approaches F1 for the vowel /a/. According to Sundberg (1975), professional sopranos tend to alter their vocal tract as they sing in this range so as to keep F1 close to F0. According to the standard linear, noninteractive source-tract acoustic theory, this type of F1 tuning would improve the efficiency of voice production, since the fundamental frequency component of the source waveform would be greatly amplified. The resulting radiated sound would be quite strong, provided that the vibratory pattern of the vocal folds was not weakened by the pressure pattern within the vocal tract caused by F1 although the spectrum of the resulting vowel would be rather sinusoidal and devoid of the coloration caused by higher harmonics. However, if we consider the acoustic interaction between the glottal source and the vocal tract acoustic impedance, the picture changes significantly.

In order to understand the relationships between glottal area, flow, and pressure, which explain the effect of the source-tract interaction, we can start with the diagram of the projected glottal area in [Figure 19-1](#).

Note that if F_1 and F_0 are matched, the ac pressure variations just above the glottis will be almost exactly in phase with the waveform of projected glottal area, independent of the assumption we make for source-tract acoustic interaction. This phase relationship stems from a basic property of a resonant system, namely, that the input impedance tends to be purely dissipative (nonreactive) at the resonance frequency. For a dissipative acoustic system, the airflow and applied pressure will be in phase. To relate the phase of the glottal airflow to the phase of the area waveform and complete the argument, we must also assume that the F_0 component of the pressure variation just above the glottis, caused by the F_1 resonance, is much larger than the other pressure components that might affect the glottal flow pattern. These other components are, namely, the sub glottal pressure variations and the supraglottal variations caused by any acoustic impedance factor not related to the F_1 resonance (such as higher resonances, the radiation impedance at the mouth, and any inertive components due to the flow pattern within or near the glottis). This assumption can be justified if the damping of the first formant is very low, as would be the case for a nonnasalized vowel with a high first formant produced with a nonbreathy voice, conditions that appear to hold for a good soprano singing an open vowel in the upper part of her register.

Under these conditions, we would have a variation in transglottal pressure (subglottal minus supraglottal) similar to that in the figure. Pressure during the most-closed portion of the glottal cycle would be increased by the "resonance pressure," while that during the open portion of the cycle would be decreased. The average transglottal pressure would be approximately the average subglottal pressure.

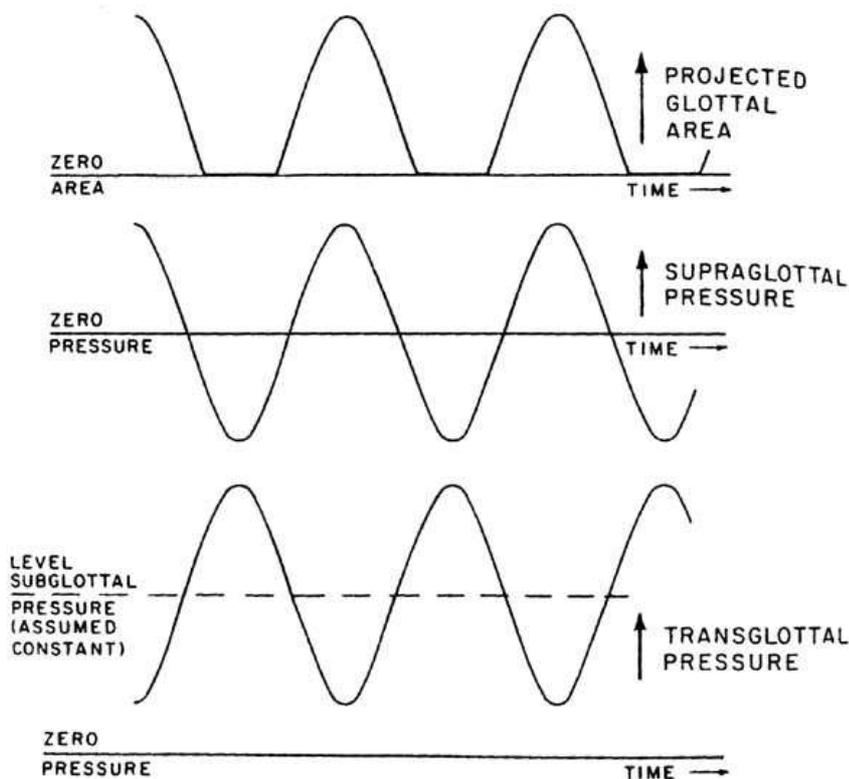


Figure 19-1. Approximate variation in transglottal pressure when F_1 is close to F_0 , assuming low F_1 damping and neglecting variations in subglottal pressure.

Now let us look at the implications for glottal airflow. Under the non-interactive assumption, the glottal airflow would approximately follow the variation in glottal area and be unaffected by the oscillations in transglottal pressure, as indicated in the sketch in [Figure 19-2](#).

If linear acoustic interaction is assumed, the fundamental frequency component of the glottal flow waveform would be suppressed by the variations in transglottal pressure. As indicated by the shaded areas in [Figure 19-2A](#), the result would be a suppression in the glottal source waveform by an amount that would have the waveform of a sinusoid at F_0 . The average flow and higher harmonics would not be affected. Thus, with a linear interactive model, the enhancement of the F_0 component is much less than in the noninteractive model, though some enhancement does occur. Also, the acoustic power (integral of flow times pressure) dissipated at the glottis decreases, even though average flow and pressure (determining the power supplied by the respiratory system) remain the same. Thus, the voice becomes acoustically more efficient.

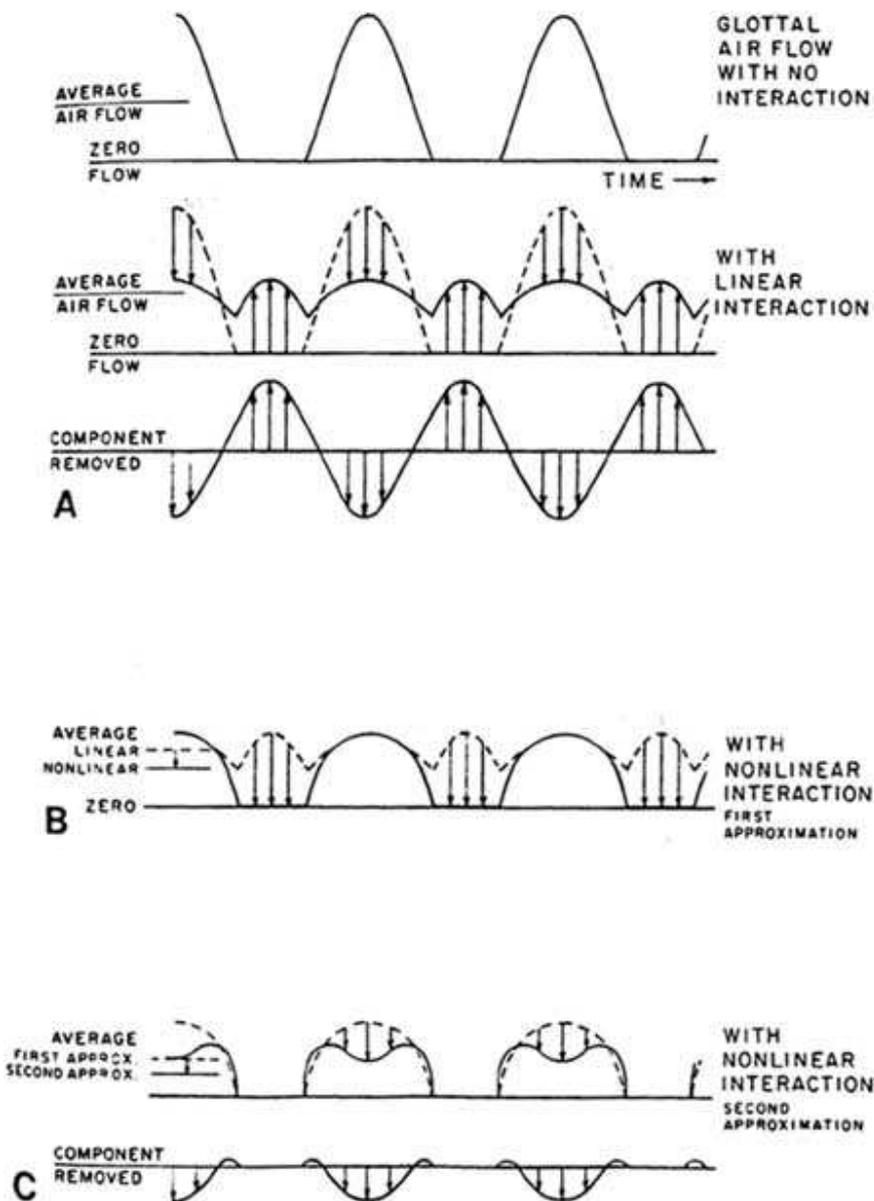


Figure 19-2. The glottal air flow that would result from the situation in Figure 19-1, assuming a linear model (A), and two approximations to a nonlinear model (B and C).

However, there is one serious deficiency in the linear, interactive model, even though it is significantly better at predicting voice quality and efficiency than the noninteractive model. This deficiency is illustrated by the fact that it predicts a nonzero glottal airflow when the vocal folds are closed, even if we assume a complete vocal fold closure during this phase. But, if we assume that the greatly increased transglottal pressure during the closed phase does not disturb the pattern of complete closure (that the closure is firm enough to withstand the increased pressure), then the flow must be forced to zero during this period. By merely forcing the flow to go to zero as the closed period is approached, we get a first approximation to the flow predicted by the nonlinear interactive model, as shown in [Figure 19-2B](#).

Note that in this first approximation to the nonlinear model, the F0 component has been greatly strengthened as compared to the linear model; the waveform looks more like a sinusoid at F0. This would strengthen the radiated SPL at F0. But, more significantly, the waveform components at frequencies other than F0 have also been altered. The high frequencies have been changed in a rather complex way, which would depend greatly upon the duty cycle of the glottal pulse, but not grossly increased or decreased in total for the rather typical duty cycle assumed here. In addition, the component at zero frequency, which is the average airflow, shows a significant decrease. This reduced average airflow not only causes a reduced glottal power dissipation as compared to the dissipation predicted by the linear model but decreased respiratory power as well.

To get a more accurate estimate of the glottal flow waveform in the nonlinear, interactive case, we should take into account that the first-approximation waveform has a strengthened F0 component as compared to the waveform with linear interaction. This would increase the supraglottal F0 component ([Figure 19-1](#)) and, therefore, increase the degree of interaction with the F1 resonance. In [Figure 19-2C](#) this is indicated by a second-approximation waveform, in which another F0 component is removed from the first-approximation waveform during the open-glottis segment. This results in a further reduction in the average airflow. However, note that, if the open quotient is more than 50 percent (it is about 60 percent in the figure), the supraglottal pressure becomes negative as the vocal folds begin to open and as they approach closure, so as to increase the glottal airflow at the onset and offset of the glottal pulse. The net result will be a sharper onset and offset of the airflow pulse for the duty cycle assumed in the figure. This sharper onset and offset would increase the energy in the higher harmonics.

That our theoretical model of the effect of nonlinear acoustic interaction is plausible is illustrated by the actual glottal airflow waveform shown in [Figure 19-3](#), from a professional soprano singing F^{#5} with a fairly high level of vocal effort during the vowel [a]. For a note in this vicinity, F0 is naturally close to F1 for the vowel [a], and therefore, a match between F1 and F0 is probably easiest to achieve for that vowel. The flow waveform was obtained from a circumferentially vented wire screen mask having an acceptable frequency response to about 3000 Hz. The mask output, measuring oral volume velocity, was inverse-filtered using a manually adjustable three formant filter, while observing simultaneous airflow and EGG waveforms during the repetitive playback of a short segment, using a two-channel transient recorder. The inverse filter was adjusted to make the filter output equal to zero during the closed-glottis period, as indicated by the EGG. The adjustment thus obtained was unambiguous and repeatable, though it should be emphasized that it required a subject having a clear period of complete glottal closure at this pitch and a moderately strong EGG waveform, as ours did.

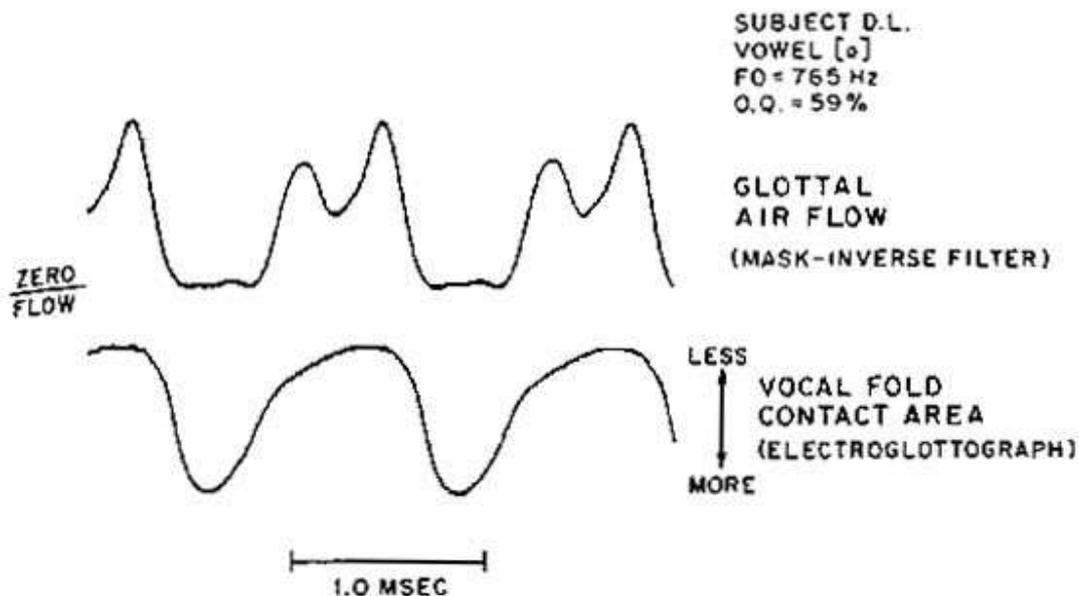


Figure 19-3. Glottal air flow as obtained by inverse filtering oral air flow for a trained soprano singing at F_0 with a high level of vocal effort. The EGG signal was used to corroborate the inverse filter settings by indicating the approximate glottal open and closed periods. To preserve time alignment, the EGG signal was electronically delayed by a time interval equal to the delay in the glottal flow signal.

The F_0 value shown in the figure, $762 \text{ Hz} \pm 10 \text{ Hz}$, was measured from the cycles caught by the transient recorder. (Some vibrato was present in the production.) The resulting inverse filter settings, also shown in the figure, indicate that the subject closely matched F_1 ($749 \text{ Hz} \pm 10 \text{ Hz}$) to F_0 for this production. The glottal flow waveform indicates that the airflow during the open-glottis period was strongly suppressed by the supraglottal pressure variation at F_1 , which lagged the airflow pulse very slightly. An attempt to reproduce this result a few weeks later with the same subject produced a similar waveform except that the dip in the airflow pulse was not obvious; i.e., the flow waveform approximated a "square wave."

It should be mentioned that one effect of the flow resistance of the pneumotachograph mask used (roughly about $0.5 \text{ cm H}_2\text{O-sec/liter}$) is to increase the damping of the vocal tract formants. Thus, the effect of F_1 tuning shown in [Figure 19-3](#) would probably be stronger without the mask. (The mask also results in a slight reduction of formant frequency; however, the singer may have taken this detuning into account in her production.) This would imply that with no mask in place, the supraglottal pressure variation can be strong enough to drive the flow during the center of the open-glottis phase even closer to zero than shown in [Figure 19-3](#). For this to occur, the peak of the ac variation in supraglottal pressure would need to be similar in magnitude to the average subglottal pressure. Supraglottal pressures of this magnitude have been measured recently by Schutte and Miller (in press) using dual catheter-mounted miniature pressure transducers: one below the glottis and one above the glottis.

Peak pharyngeal pressures as high as the subglottal pressure, though not explained by a linear interactive model, are entirely consistent with a nonlinear model. In fact, the nonlinear model indicates that if the resonance were sufficiently underdamped (efficient), the net transglottal pressure could actually reverse for part of the glottal cycle, to cause a brief period of negative airflow (from the pharynx to the trachea). Whether a resonance that efficient can be attained, or whether proper vocal fold oscillatory behavior could be maintained with such a flow pattern, is not clear at this time, though Schutte and Miller's measurements appear to show at least one case in which this has occurred.

Implications for the Soprano Voice

We have shown that a significant reduction in average airflow can result from proper F1 tuning, given a non-nasalized production with a complete vocal fold closure for some appreciable portion of the glottal cycle, and assuming that the proper vocal fold vibratory pattern can be maintained under these conditions. This conclusion is supported not only by Sundberg's (1975) formant measurements and Schutte and Miller's measurements of supraglottal and transglottal pressure, but also by the strong feeling held by at least two sopranos I have talked to (based on their introspection) that average airflow can be affected significantly by vocal tract posture. For example, one of these singers (Jo Estill) shared with me her intuition of an increased airflow during nasalized vowels. This could be due to an increase in the damping of F1 and a resultant decrease in the nonlinear, interactive effect.

Our model also indicates that, if in addition to these conditions the glottal duty cycle is in the proper range, the reduced airflow due to vocal tract tuning can be accompanied by a strong, harmonic-rich tone. From the qualitative analysis sketched earlier, the optimal duty cycle or open quotient required for the production of strong higher harmonics, as controlled primarily by the degree of vocal fold adduction, is somewhat greater than 50 percent. With the open quotient greater than 50 percent, transglottal pressure increases at the onset and termination of the glottal pulse. This increased pressure will cause a more abrupt onset and offset of glottal airflow and, therefore, somewhat stronger higher harmonics. Open quotients much greater than 50 percent, though theoretically producing a harmonic-rich tone, may not result in a complete glottal closure and thus violate the assumptions of the model. This duty cycle requirement appears to be different than the requirement in modal voice, in which the strength of the higher harmonics tends to increase monotonically with a decrease in open quotient for a given level of average airflow. However, a better specification is required of the effect of glottal duty cycle in the soprano voice and its relation to the optimal vocal tract tuning.

Aside from the optimization of the duty cycle and an essentially complete glottal closure, the details of the waveform of projected glottal area do not appear to be important.

In examining our analysis for possible implications related to vocal abuse, it should be noted that our model indicates the importance of a fairly complete glottal closure during the closed phase of the glottal cycle for an efficient soprano singing voice. Conversely, it is possible that if a singer on a particular day cannot produce the essentially complete vocal fold closure required for a strong interactive effect, she may experience an excessively high airflow and a resulting increased risk of vocal abuse. Any attempt to compensate by the use of increased vocal fold adduction to reduce airflow might bring its own risk of vocal abuse through fatigue of the adductory musculature, as well as a possibly unacceptable tonal balance due to too small an open quotient.

Since the model described indicates that, for a given F_0 , small changes in the frequency or damping of the first formant, or in the degree of vocal fold adduction, can greatly affect the relative strength of all the harmonics of F_0 , these factors can conceivably have a significant effect on vowel quality and perceived vowel identity. The possible sensitivity of vowel identity to these factors is not predicted by a linear model, nor, apparently, is it present in male singing.

Another tentative conclusion for the soprano voice might be that nasalized vowels or notes at lower pitches may require a technique other than supraglottal vocal tract tuning to reduce the average airflow during the open glottal phase while maintaining a high SPL. The inertive acoustic loading mechanism apparently used by the bass or baritone singer is one candidate for such a mechanism.

Implications for Modal Voice

We have argued that, for sopranos, a crucial factor in producing vocalization at high sound pressure levels without vocal abuse is the ability of the larynx and vocal tract, working together, to maintain a reasonably low average airflow at elevated levels of lung pressure, without the strain of excess vocal fold adduction. There is considerable evidence that at lower values of F_0 , for both males and females and in speech as well as singing, a similar function can be performed by inertive vocal tract loading. It is now generally accepted that, when F_0 is much smaller than F_1 , the inertance of the vocal tract airflow, by creating an appropriate variation in transglottal pressure during the glottal cycle, suppresses the buildup of airflow during the glottal opening phase and maintains a higher airflow during the glottal closing phase, thus skewing the glottal airflow pulse to the right and concentrating an increased generation of high frequency energy near the instant of vocal fold closure. If lung pressure is held constant and vocal tract inertance varied, this type of source-tract acoustic interaction will lead to an increase in energy in the third formant region of about 5-10 dB, depending on the degree of interaction and the model assumed for the calculations (Fant, 1982; Rothenberg, 1981).

However, increasing vocal tract inertance also decreases the average airflow for a given lung pressure. If when calculating the effect of varying vocal tract inertance, we consider the average glottal airflow to be constant and average lung pressure as a dependent variable that assumes the value necessary to maintain average airflow, then the increase in higher-formant energy that can conceivably be caused by this type of interaction increases to as much as 20 dB (Rothenberg, draft manuscript). Thus, the large differences in vocal efficiency in modal voice that are observed among otherwise normal voices could be explained by interaction with vocal tract inertance. However, we must break away from the concept that voice is produced by a fixed reservoir of pressure, which is relatively constant between speakers for a given "vocal effort," and change to the concept that the respiratory system is a source of airflow, which is relatively constant between speakers for a given "vocal effort," with lung pressure being a secondary variable in voice production.

This new concept is at least as supportable physiologically as the concept of a fixed reservoir of pressure. Excessive airflow can dry out the mucosa, lead to too-frequent breath pauses, and possibly even hyperventilation in a speaker or singer who is vocalizing continuously over a long period of time. On the other hand, measurements with a manometer and a tube at the lips will easily show how easily attainable are lung pressures much higher than commonly reported for speech. Moreover, high values of lung pressure are produced by relatively large abdominal and intercostal muscles that usually are not felt

to be fatigued, even in stressful vocalization situations, and that could be developed further, if necessary, by the professional vocalist.

What is more likely to be fatigued as subglottal pressure is increased are those elements of the laryngeal musculature that must hold the vocal folds sufficiently adducted to maintain an acceptable rate of airflow at the increased lung pressure. The parameter of adductory tension is undoubtedly an important contributor to the total concept of vocal effort and should not be confused with subglottal pressure, even though the two tend to co-vary in a single individual, for a given flow rate.

The question of which is the primary aerodynamic variable in the generation of voice-average lung pressure or average airflow-may not have a perfectly clear answer because of the difficulty in defining and equating among speakers the degree of "vocal effort." It is an important question, nevertheless. For example, a measure of vocal efficiency proposed by Isshiki(1981), namely the ratio of ac to dc airflow, would be supported by a conclusion that flow is primary. The question is surely worthy of future research and debate.

Epilogue

The decreased average flow brought about by proper vocal tract tuning should come as no surprise to an electronic engineer familiar with radio transmitter amplifiers or to an experienced ham radio operator. It would only be necessary to point out that a soprano singing in the upper part of her range is analogous to the final amplifying stage of a radio transmitter. This amplifier must supply a maximum electrical power to the antenna while drawing a minimum average electrical current from the power supply. In the so-called class C amplifier, commonly used for this purpose, the power supply current is allowed to flow to a tuned electrical circuit, and thence to the antenna, for only a short interval during each cycle of the transmitter signal (Terman, 1947). The transmitter engineer can check for the proper tuning (proper resonance frequency) of this antenna circuit by adjusting the tuning for a minimum average power supply current, just as the trained soprano can adjust her vocal tract tuning for a minimum average expenditure of lung air.

From a mathematical - not an aesthetic - point of view, the primary difference in operation between the transmitter amplifier and the soprano is that the amplifier, for proper operation, uses a duty cycle (an open quotient) much smaller than 50 percent. Looking to the amplifier analogy for lessons, a soprano might well note that an output amplifier that operates over a period of time during which it is improperly tuned can overheat and blow its fuse.

ACKNOWLEDGMENT

This work was supported by a research grant from the National Institutes of Health.

Discussion

Dr. Scherer: When you look at subglottal pressure measurements made with a wide bandwidth, like those of Miller, the subglottal formant shows that the subglottal pressure falls to about 50 percent of the mean value somewhere in the subglottal cycle. I'm not sure what are the subglottal resonances for females, but if they are tall singers with long tracheas, the formant frequencies might be relatively low,

so this negative-going part of the subglottal pressure might coincide in time with the open portion of the vibratory cycle. This would be a negative pressure added to the pressure in the pharynx, and that would help to drive this flow back. Do you want to comment on that?

Dr. Rothenberg: A factor that's important in creating this high supraglottal pressure is the efficiency of the resonance. For example, it may sometimes be necessary to keep the pharyngeal wall stiff to keep the resonance efficient. In addition, as mentioned in the paper, our results might explain Jo Estill's intuition that nasality increases airflow. If a female singer singing in the high soprano range opens the nasal port slightly, the vocal tract resonances will be damped, and this damping might cause airflow to increase. The efficiency of the resonance is very important.

Since the subglottal resonances tend to be more damped than the supraglottal resonances, they may not be as effective in altering glottal airflow. Nevertheless, it is true that the peak-to-peak subglottal pressure variation can be significant compared to the average subglottal pressure. However, if the soprano uses "tuning" of the supraglottal resonance to reduce the glottal airflow, the reduced airflow will also suppress the subglottal pressure variations. So during a properly executed tuning maneuver, the subglottal pressure may be of much less importance than the supraglottal resonance.

Dr. Scherer: Thank you. A long time ago Bartholomew gave a paper in JASA deal in with male phonation¹ saying that, for a good quality, male singers have the singer's formant region enhancement, as well as a low frequency enhancement around the first formant, around 500 Hz or so. Do you think that, if you were to take spectra of the double humped waveforms you showed, you would in fact have some energy from the spectra that would allow a relative enhancement of a low region and a high region? I'm trying to figure out a laryngeal, rather than vocal tract, acoustic reason for the enhancement of these two regions.

Dr. Rothenberg: In looking at the soprano voice, I neglected the inertive effect and I also examined pitches at which the resonance was very important. Conversely, in the male voice I have looked at just the inertive effect and did not look at resonance effects. There is a big range between sopranos and basses, and there's a lot of complexity in between. You can have both effects in one voice, and in the future we should look at how they interact.

Ms. Estill: Would you like to know how I do this tuning?

Dr. Rothenberg: Go ahead.

Ms. Estill: I can describe my tuning maneuvers. To make an /i/ constriction close the port, tighten the ari-epiglottic ring, and then anchor the whole mechanism, perhaps by tightening the pharyngeal wall. But these three different kinds of constrictions above the glottis may contribute to the supraglottic pressure that we're talking about.

Dr. Rothenberg: Some of those maneuvers would tend to make a hard wall tube with a sharp resonance, and maybe that would reduce the damping.

Dr. Baer: Martin, I want to clarify a point. When you suggested we think of the larynx as responding to flow rather than pressure, you meant the control system for the larynx rather than the mechanical

structures. I'm glad you flagged that point, because our paper claims exactly the opposite: that over breath groups, it's pressure that's controlled. We are not talking about control within a glottal cycle now but control over the level of syllables. When you have obstruent syllables, it appears more that the pressure is being dynamically controlled over sentence length intervals.

Dr. Rothenberg: There's no conflict really. I was not really talking about the control during a syllable; because, obviously, if you change the pressure on top, you may have to change the pressure on the bottom to compensate. I was talking about the control as it relates to loudness.

Dr. Baer: In singing, for instance, which is different from reading sentences, your task is to get the loudness set up and to get a rich quality.

Dr. Rothenberg: Or vocal effort. If I want to increase my vocal effort, I am really increasing the airflow. And, to do this, I have to have a higher pressure. But, I can also talk louder by increasing the airflow, not by increasing the pressure, even though I am doing both. It's not dynamic control during the syllable that I'm talking about.

Dr. Stevens: Would you have any comment about how people control loudness when they are at say, 10,000 feet, where the relation between pressure and velocity is different? Do people who live at high altitudes control pressure the same way as we do, or do they control velocity?

Dr. Rothenberg: We have reported an experiment, using a helium-oxygen mixture, in which we found that the flow was increased by the helium for a given lung pressure^{2a} However, it appeared that after using the helium for just a short while the subject adapted by increasing vocal fold adduction to reduce flow. Speakers may do something analogous at high altitudes.

Editor: Relevant observations have been made by Wathen-Dunn and Michaels (1968).²

Dr. Cranen: I would like to stress the point I made this morning, comparing normal speakers and singers for glottal closure. Is there a leak area or not? When you look at the back slope of the singer's glottal flow waveform, you see that, although the top of the amplitude of the glottal form waveform is the same, the back slope is much steeper; and this difference is related to vocal efficiency. So, when you want to compare singers with un-trained speakers like me or Lou, it is important to consider the leak area as an important factor.

Dr. Rothenberg: I agree.

Dr. Titze: I'm a little nervous about your going through all this rationale without paying attention to what happens to the tissue in this strong nonlinear interaction. I was wondering if you can make your case, as you do, without involving changes in the driving of the tissue. Either the tissue is insensitive to these pressure changes, moving essentially in its normal mode pattern - which is something that I have believed for some time - or, if that isn't the case, you would have to consider changes in vibratory movement and basic glottal configuration to determine if the vocal folds oscillate. Considering the high-frequency part of the source spectrum, I can't see how you can do that independent of what the pressure does to the vibration.

Dr. Rothenberg: Yes, I'm glad you reminded me of that. You could test your hypotheses easily by looking at the electroglottograph waveform. Even though we may not know exactly what it means in terms of the movement, we know that if it stays the same then the movements stay the same, grossly. However, I would do this experiment with a subject that had a stronger electroglottograph waveform than those I've shown, since a weak electroglottograph waveform has two components. One is the component due to vocal fold contact area, but there is also a component caused by other F0-synchronous vibrations of different parts of the anatomy, perhaps far from the glottis, such as tongue surface vibrations. If you vocalize loudly you can feel the tongue surface vibrating. When the EGG waveform is very strong, you can assume that it's coming mostly from the vocal folds, but sometimes the signal from the vocal folds is relatively weak.

Dr. Titze: If the larynx is not protruding enough or if the angle is wide, then the electric field pattern goes way out instead of straight across. This often happens with women and children; you can't get good EGG because the field goes in a wide arch instead of directly between the electrodes.

Dr. Rothenberg: Yes, that may be, because I've seen obviously inaccurate waveforms on both men and women, sometimes under conditions for which I couldn't identify the causal factors. In such cases, the noise component which is synchronous with the glottal vibrations can even dominate the signal. But you always have both components mixed together to some degree.

Dr. Hirano: When the soprano singer tunes the formant to the fundamental frequency, how does she differentiate different vowels?

Dr. Rothenberg: Let me give you an anecdotal example, The singer I've worked with found a way of singing the passage with an /i/ vowel that was easier for her. She was able to make something that came across as an /i/, but perhaps it had the first formant raised, So, she was able to find the vocal tract configuration that gave her a proper first formant (for singing) but was acoustically acceptable as an i/. I want to repeat this experiment again with her to see what the formants actually were, to see if she was tuning, My prediction would be that that is what was happening.

Dr. Scherer: I would like to ask Martin and the authors of previous papers to combine their ideas on these aspects of more efficient singing and other acoustic effects, plus constrictions downstream. What were the effects and what you would expect?

Dr. Titze: It's been suggested that for the so-called primary register transition, which is roughly the same for male and female speakers and occurs somewhere between 290-350 Hz, there could be an acoustically triggered change in the mechanism, the larynx, and the vibratory pattern of the vocal folds. Further, it would seem that some kind of supraglottal loading could, in fact, enhance the oscillation up to the first formant, when an abrupt transition is made from an inductive to a capacitive load. This would be an undesirable loading that kicks the system into something like a falsetto from the normal chest system. I was wondering how your analysis here would treat the slightly higher end of the resonance. Two years ago, in Stockholm, I did a paper on that, trying to show how the subglottal and supraglottal pressures would change the driving pressures of the vocal folds, and it appeared that ideal conditions were achieved right into the region of the formant. But right above the formant, the phases turned around and it seemed as though the tract pressures would not maintain the vibration very well. Have you looked at those glottal waveforms just slightly on the high side of that formant?

Dr. Rothenberg: The answer is no. We did some very gross loading measurements, such as trying to change the resonance by putting a partial obstruction near the lips and looking at average airflow, But the inverse filtering technique we use is very, very difficult under these conditions, So to modify the procedure and still get accurately inverse-filtered flow signal (it's now being done by hand just for a short segment during which the formants are assumed constant) is tedious.

Dr. Stevens: Using these kinds of mechanical changes, you can fix up something in the mouth, phonate, and change the constriction suddenly. Then you can adjust (change the constriction) So that you have tuned exactly to the fundamental. After that, you can change the constriction even more, so that the formant goes above the fundamental, and observe what happens both to the waveform and to the spectrum. At least, you can measure the spectrum of the sound that's radiated.

Dr. Rothenberg: Now, suppose you don't do it quite instantaneously. The singer is listening; she adapts and moves her formant back to compensate to where it was supposed to be because she feels uncomfortable. We've tried this technique, but we didn't know how the singer was compensating.

Dr. Fujimura: When the vocal tract is in tune, it will be more important to consider the effect of damping for the formant. Most probably, there would be a peculiar damping effect. By over-damping the first formant, it would not be crucial any more to have the F1 in the right place, even though the vowel quality wouldn't be clear. It is not misinterpreted as another vowel and, to me that seems to be exactly what's happening in singing. You can't really tell the phonetic value very clearly, but it doesn't mislead you either.

Dr. Rothenberg: I think what you are saying is, if you damp the formant, then you won't have the strong spectral differences as a function of pitch, for example.

Dr. Fujimura: Well, it may be a function of pitch, but the phonetic quality doesn't matter so much. The vowel quality may not be clear, but it doesn't indicate another vowel.

Dr. Rothenberg: I don't want to give the idea that women "tune" all the time. Probably a good singing technique requires a number of different mechanisms in different parts of the range and at different volume levels. But I would assume that when tuning is used a minimum damping is important for a loud, higher pitched vocalization, especially if it was held. But it may not be always necessary, and increasing the damping may even be necessary under other circumstances in singing.

Dr. Luschei: I understand there may be some disagreement about the exact nature of these pressure pulses. However, people are agreed that, with small transducers in place, there are substantial pressure waves at the fundamental frequency of oscillation supraglottically and subglottically. If that's the case and if there's no suspicion that those pressure waves are due to movements of the pressure transducers, then you have to conclude that any mechanoreceptors in the vicinity are also subject to those pressure modulations. If so, probably any sensitive mechano-receptors in the area must be masked properly. Their behavior is going to be dictated by the pressure variations, if they are at all physically sensitive to the actual phonation taking place. This would, then, provide a good explanation why one might not see pressure variations, or reflex responses to pressure variations, at low frequencies. It also means, interestingly enough however, that the afferent source has to be a potential source of feedback independent of the auditory system, if the nerves can actually respond to these pressure variations. So, it

seems to me that a useful thing might be to see whether these afferent fibers supraglottically or subglottically respond to these pressure variations. But I want to make sure it is not just some kind of turbulence in the standing wave or some funny thing like what is being measured here.

Dr. Titze: I think there have been a number of investigations of the subglottal pressure variation. The fluctuations tend to be around 40 percent of the mean value. That's just a rough value. And it shows up in all the simulations that we have done. So, if we're wrong, we are collectively wrong on some very central assumptions.

Dr. Stevens: I think we can calculate that the actual motion of the tissue due to these pressures is a fraction of a millimeter. I would guess that would certainly be enough to excite those receptors.

Dr. Rothenberg: These high pressures may give more credibility to the singers' characterization of their vocalizations, like "head voice." These strong pressure variations must give some sort of a sensation that a singer could identify and relate to parts of the anatomy.

Dr. Scherer: What is the source for the wall movement measurements, Dr. Stevens?

Dr. Stevens: I take the more or less accepted acoustic mass of the wall - I think people agree it's in the range 1-2 (or 1/2-2) grams per cm² - and calculate from that the motion.

Dr. Megirian: Coming back to the story of the mechano-receptors as part of a feedback loop, I think we can point out that some earlier work by Bruce Mehl in John Widdecomb's lab shows that the mechano-receptor in the upper airway is a very fast adapting receptor. I would like to ask Dr. Sasaki if the effects of local anesthetization on the singing voice have been studied.

Dr. Sasaki: I believe that Dr. Gould and his associates³ have done experiments of this sort.

Dr. Luschei: Dr. Wyke mentions the effects of topical anesthesia of the subglottal laryngeal mucosa in the discussion of his paper at the last conference (Wyke, 1985).⁴ He says that topical anesthesia of the subglottic laryngeal mucosa has little effect on the conversational speaking voice but produces perturbations of the declamatory speaking voice and renders singing almost impossible, because of lack of accurate pitch and intensity control.

Dr. Harris: There are experiments in the speech literature on the effects of trigeminal nerve block, superficial anesthetization of the oral mucosa, (Borden, Harris, & Catena, 1973; Scott & Ringel, 1974),⁵ and anesthetization of the temporo-mandibular joint (Kelso & Tuller, 1983).⁶ Although there are procedural problems with these studies, the overall result is that these proceedings have surprisingly little effect on speech.

Dr. Rothenberg: These comments lead me to an idea about how I could become a soprano. Those of us who can't sing could still experience a singer's vibratory sensations by introducing an artificial tone into the vocal tract, That's similar to what I think Sundberg (1979)⁷ has already done at lower pitches, with bass singers, measuring the vibration of the thorax and abdomen, I believe that he was trying to get some intuition about where the term "chest voice" comes from. But we can also get these sensations by having

sounds artificially introduced into the vocal tract. Until recently, the experiment could not have been performed well for the soprano voice, since he didn't know how high the pressure levels were.

Footnotes

1. Bartholemew, W; T. (1984). Physical definition of "good voice-quality" in the male voice. *Journal of the Acoustical Society of America*, 6, 25-33.
- 2a. Rothenberg, M. (1984). Source-tract acoustic interaction and voice quality. *Transcripts of the Twelfth Symposium on Care of the Professional Voice*, 25-31. New York: The Voice Foundation.
2. Wathen-Dunn, W., & Michaels, S. (1968). Some effects of gas density on speech production. In *Sound production in Man*, edited by A. Bouhuys. *Annals of the New York Academy of Sciences*, 155 368-378.
3. Gould, W. J., & Okamura, H. (1974). Interrelationships between voice and laryngeal mucosal reflexes. In *Ventilatory and Phonatory Control Systems*, edited by B. Wyke. London: Oxford University Press.
- Gould, W. J., & Tanabe, M. (1975). The effects of anesthesia of the internal branch of the superior laryngeal nerve upon phonation: An aerodynamic study. *Folia Phoniatica*, 27, 337-349.
4. Wyke, B. (1985). Reflexogenic contributions to vocal control systems. In *Vocal fold physiology: Biomechanics, acoustics and phonatory control*, edited by I. Title and R. Scherer. Denver: The Denver Center for the Performing Arts.
5. Borden, G. J., Harris, K. S., & Catena, L. (1973). Oral feedback II. An electromyographic study of speech under nerve-block anesthesia. *Journal of Phonetics*, 1, 297-308.
- Scort, C. M., & Ringel, R. L. (1971). The effects of motor and sensory disruptions on speech: A description of articulation. *Journal of Speech and Hearing Research*, 6, 369-378.
6. Kelso, J. A. S., & Tuller, B. (1983). "Compensatory articulation" under conditions of reduced afferent information: A dynamic formulation. *Journal of Speech and Hearing Research*, 26, 217-224.
7. Sundberg, J. (1983) Chest vibrations in singer, *Journal of Speech and Hearing Research*, 26, 329-340

References

- Fant, G. (1982). Preliminaries to an analysis of the human voice source. *Quarterly Progress and Status Report* (Speech Transmission Laboratory, Stockholm: Royal Institute of technology), STL-QPSRX, 4/1982, 1-27. Also, in *Speech Communication Group Working Papers*, Massachusetts Institute of Technology, Research Laboratory of Electronics.
- Isshiki, N. (1981). Vocal efficiency index. In *Vocal Fold Physiology*, edited by K. N. Stevens and M. Hirano. Tokyo: University of Tokyo Press.
- Rothenberg, M. (1981). Acoustic interaction between the glottal source and the vocal tract. In *Vocal Fold Physiology*, edited by K. N. Stevens and M. Hirano. Tokyo: University of Tokyo Press
- Rothenberg, M. (draft manuscript). The effect of flow dependence on source-tract acoustic interaction.
- Schutte, H. K., & Miller D. G. (in press). The effect of F0-F1 coincidence in soprano high notes on pressure at the glottis. *Journal of Phonetics*.
- Sundberg, J. (1975). Formant technique in a professional female singer. *Acoustica*, 32, 89-96
- Terman, F. E. (1947). *Radio Engineering*, New York: McGraw Hill.