ARTICLE IN PRESS

Interactive Augmentation of Voice Quality and Reduction of Breath Airflow in the Soprano Voice

*Martin Rothenberg and †Harm K. Schutte, *Syracuse, New York; †Groningen, Netherlands

Summary: In 1985, at a conference sponsored by the National Institutes of Health, Martin Rothenberg first described a form of nonlinear source-tract acoustic interaction mechanism by which some sopranos, singing in their high range, can use to reduce the total airflow, to allow holding the note longer, and simultaneously enrich the quality of the voice, without straining the voice. (M. Rothenberg, "Source-Tract Acoustic Interaction in the Soprano Voice and Implications for Vocal Efficiency," Fourth International Conference on Vocal Fold Physiology, New Haven, Connecticut, June 3–6, 1985.) In this paper, we describe additional evidence for this type of nonlinear source-tract interaction in some soprano singing and describe an analogous interaction phenomenon in communication engineering. We also present some implications for voice research and pedagogy.

Key Words: Singing–Voice–Nonlinear interaction–Source-tract interaction–Reduce airflow–Soprano.

INTRODUCTION

In a 2014 interview by Jian Ghomeshi, the noted singer Barbra Streisand made the following comment on her ability to maintain a strong Eb5, or similar, high note: that she did not know how she did it, but would like to know the physiological mechanism. She said, "I want to understand the physiology, because that always interested me. What was I doing? People ask me 'How come you hold the note so long?' And I answer (simply) because I want to." This paper summarizes some attempts to explain the articulatory mechanisms underlying the effect Barbra Streisand was referring to. Also outlined is an analogous principle in radio engineering technology that is well known to radio engineers, but not known in the field of singing research and pedagogy, except, perhaps, for those of us who also have a background in radio engineering.[a]

In 1985, at a conference sponsored by the National Institutes of Health, Martin Rothenberg described a physiological/acoustic mechanism by which some sopranos, singing in their high range, can reduce the total airflow, to allow holding the note longer, and simultaneously enrich the quality of the voice, without straining the voice.[1] It is likely that this mechanism accounts for the strong maintained vowels that Barbra Streisand referred to.

Since his 1985 paper, there have been other research reports corroborating Rothenberg's findings. However, there has not been a wide acceptance of the use of this type of nonlinear source-tract interaction in soprano voice, with most of the literature tending to employ a linear analysis that does not explain the power and timbre of a strong soprano voice. In this paper, we will put forth the evidence for nonlinear source-tract interaction in some soprano singing and describe an analogous interaction phenomenon in radio technology. We also present some implications for voice research and pedagogy.

An analysis of this mechanism leads us to the conclusion that not all singers can achieve this interactive reduction of airflow. It appears that to do this, a singer's vocal fold vibration pattern should have a "closed quotient" of roughly 50%, or at least about 30% (meaning that the vocal folds remain closed during at least a third of the vibratory period), and also have a good vocal fold closure (no air leakage) during most of the closed phase. The precise value of the closed quotient is not important, as long as a good closure is attained to prevent airflow during the closed phase. (Apparently Barbara Streisand's voice has these characteristics.) The singer would then adjust the vowel being sung so that the lowest vocal tract resonance (usually termed the "first formant") falls at a frequency that is approximately the pitch frequency of the note. An example of this type of interaction in a soprano voice is shown in Figure 1, taken from figure 19-3 from Rothenberg's 1985 paper.

The explanation for the reduced air consumption is that each airflow pulse through the glottis after the first is suppressed in amplitude by a peak in the oral pressure just above the glottis that occurs each time the glottis is open. (This tends to give the glottal airflow pulses a double-humped appearance in some cases.) Figure 2 shows diagrammatically how the pressure pulse from a previous glottal airflow pulse (rising arrow A), when returning from the mouth (descending arrow B), can suppress the airflow in the succeeding airflow pulse. The timing of the returning pressure pulse will be correct for suppressing the succeeding airflow pulse if the frequency of the first formant F1 is approximately the same as the frequency of the glottal pulses, F0.

As pointed out by Rothenberg,[1] this principle of tuning to reduce airflow is well known in the field of radio engineering, but until now, is not known in singing voice analysis. (This is documented below.)

The observation that many sopranos tend to tune F1 close to F0 at higher pitches is supported in the literature. Johan Sundberg[2] has found that a professional soprano he recorded tended to place F1 near F0 at higher pitches, supporting the thesis of this paper. More recently, Andrea Deme[3] showed formant measurements that lead to the same conclusion. Deme justified this F1 placement as required for voice strength, because placing F1 above F0 would mean that F1 would amplify no glottal harmonic

[a]Dr. Rothenberg received a First Class Radiotelephone Operators license at the age of 18 and subsequently serviced transmitters and receivers in the US Army Signal Corps while serving in Korea, where he learned to "tune for the dip." Dr. Schutte, MD-ENT, PhD, was trained as a radar engineer in the Dutch army in 1960–1962. He has been an amateur radio operator since 2003, with the call letters PC5BV, and enjoys building and working with vacuum tube transmitter amplifiers.

Accepted for publication September 23, 2015.

From the *Syracuse University, Syracuse, New York; and the †Groningen University, Groningen, Netherlands.

Address correspondence and reprint requests to Martin Rothenberg, Syracuse University, 1201 E. Fayette St., Syracuse, New York 13210. E-mail address: mrothen@glottal.com

Journal of Voice, Vol. ■■, No. ■■, pp. ■■-■■
0892-1997

http://dx.doi.org/10.1016/j.jvoice.2015.09.016

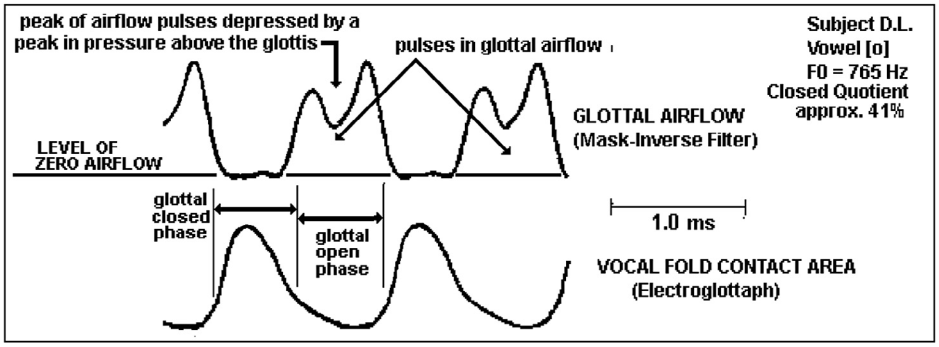

FIGURE 1. Adapted from figure 19-3 in Rothenberg.[1] Glottal airflow as obtained by inverse filtering oral airflow for a trained soprano singing F#5. The electroglottograph signal was used to corroborate the inverse filter settings by indicating the approximate closed and open intervals. To preserve time alignment between EGG and airflow, the EGG signal was electronically delayed by a time interval equal to the delay in the glottal flow signal with respect to the EGG signal.

(because F1 could not reach the second glottal harmonic at a high pitch). However, neither author comments that for F1 near F0, a linear acoustic interaction model would predict a vocal quality with a dominant fundamental component, such as in breathy voice or falsetto voice. The possibility of nonlinear acoustic interaction in which the amplitude of the fundamental component in the glottal airflow wave was depressed by the interaction was not considered.

In the year following the presentation of Rothenberg,[1] Gunnar Fant, as well as Harm Schutte and Donald Miller published data corroborating the theory proposed by Rothenberg in 1985.[1]

CORROBORATION BY FANT

Figure 6 from Fant's paper "Glottal Flow: Models and Interaction,"[4] showing a computer simulation result, and some of his conclusions, are reproduced here.

In the simulation shown in Figure 3, the formant value chosen resulted in a strong, harmonic-rich output volume and low airflow consumption caused by the suppression of glottal airflow by the pressure wave above the vocal folds.

Fant concludes:

> "One aspect of such superposition interaction at high F0 is the situation of a soprano singing at an F0 coinciding with F1. As shown in fig. 6, the peak value of the supraglottal F1 component pressure is slightly greater than the subglottal pressure and occurs just before the instant of maximum glottal aperture thus highly reducing the maximum rate of airflow, whilst retaining a high flow pulse steepness at closure. This simulation supports the suggestion of Rothenberg (1985)[1] that an F0/F1 coincidence minimizes the air consumption of soprano phonation."

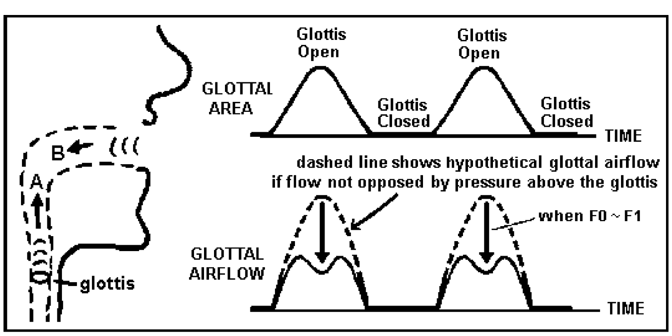

FIGURE 2. Diagrammatic explanation of airflow suppression when F1 is approximately equal to F0.

CORROBORATION BY SCHUTTE AND MILLER

Using miniature pressure transducers suspended just above and below the vocal folds with two sopranos, Schutte and Miller[5] corroborated the Fant simulation results, namely that when the lowest formant is tuned close to the pitch of the voice, the resulting pressure oscillations in the pharynx (p-supra in Figure 4) can result in a pressure peak at the approximate center of the glottal open phase, with the open and the closed glottal phases interpreted approximately from the electroglottograph (EGG) waveform. This would cause a reduction in the transglottal pressure (p-trans) that drives the air through the glottis. Although airflow was not measured directly in the present study, the sharp drop in transglottal pressure during each glottal open phase could

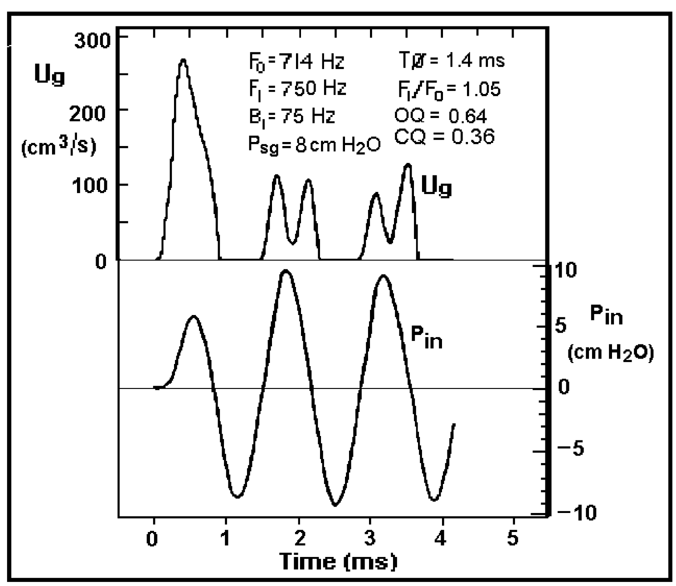

FIGURE 3. Figure 6 from Fant,[4] showing simulated glottal airflow and supraglottal pressure, assuming a one-formant vocal tract with the formant frequency (750 Hz) close to F0 (714 Hz).