Preliminaries to the body-cover theory of pitch control

Article *in* Journal of Voice · December 1988

DOI: 10.1016/S0892-1997(88)80004-3

3 authors, including:

Ingo R Titze
University of Utah
424 PUBLICATIONS **21,266** CITATIONS
SEE PROFILE

Jack J Jiang
University of Wisconsin–Madison
310 PUBLICATIONS **6,442** CITATIONS
SEE PROFILE

Journal of Voice
Vol. 1, No. 4, pp. 314–319

Preliminaries to the Body–Cover Theory of Pitch Control

*†Ingo R. Titze, *Jiaqi Jiang, and *David G. Drucker

**Voice Acoustics and Biomechanics Laboratory, Department of Speech Pathology and Audiology, The University of Iowa, Iowa City, Iowa, and †The Recording and Research Center, The Denver Center for the Performing Arts, Denver, Colorado, U.S.A.*

Summary: Preliminary theoretical and empirical work suggest that increased thyroarytenoid muscle activity may either increase or decrease fundamental frequency, depending on cricothyroid muscle activity and a new cross-sectional area parameter. This parameter is defined as the ratio of muscular tissue in vibration to total tissue in vibration. Canine laryngeal nerves were stimulated to measure vocal-fold length changes. These data, combined with previously reported tissue density, passive stress, and passive frequency data, were used to construct a set of curves predicting canine fundamental frequency from thyroarytenoid and cricothyroid muscle activity and the area ratio. The results suggest that high cricothyroid muscle activity and small area ratios tend to cause fundamental frequency lowering with increased thyroarytenoid muscle activity. **Key Words:** Body—Cover—Thyroid cartilage—Cricoid cartilage—Vocal folds—Cricothyroid muscle—Thyroarytenoid muscle—Fundamental frequency—Passive and active stress—Area ratio—Depth of vibration.

It is generally understood that the fundamental frequency of oscillation (F_0) of the vocal folds is controlled primarily by the cricothyroid (CT) muscles, secondarily by the thyroarytenoid (TA) muscles, and tertiarily by subglottic pressure. Increased activity in the CT muscles and increased subglottic pressure (P_s) raise F_0, whereas increased activity in the TA muscles may either raise or lower F_0 (1). The lack of a clear picture of how F_0 changes with TA activity is likely to be a result of the complex way in which the effective length, tension, and mass change when the TA muscles contract. Although the effective tension always increases when the CT muscles elongate the folds, or when P_s increases the amplitude of vibration, the effective tension may decrease when the TA muscles shorten the folds. This is because the increased slackness of the vocal fold cover may dominate over the increased stiffness of the body.

This leads us to the body–cover hypothesis proposed by Hirano (2). Simply stated, it suggests that the morphology of the vocal fold lends itself (mechanically) to a division between two groups of tissue layers that are set into vibration by the glottal airstream. The groups are coupled in the lateral direction, but have rather different mechanical properties. The outer group, or cover, is composed of the epithelium (skin) and the superficial and intermediate layers of the lamina propria. The cover is very pliable and has no contractile properties. The cover, being heavily irrigated with liquid, can propagate a surface wave that facilitates energy transfer from the glottal airstream to the vocal-fold tissues (3–5). Its tension is controlled by vocal-fold length. The inner group, or body, is composed of the deep layer of the lamina propria and the TA muscle. It is less deformable and has active contractile properties. Its tension is therefore not only determined by length, but also by active stiffening of the muscle internally. The combined tension of the portions of the body and cover in vibration would seem to regulate the F_0.

This article is a second attempt to quantify the

Presented at the Fifteenth Symposium: Care of the Professional Voice, New York, June 1986.

Address correspondence and reprint requests to Dr. Ingo R. Titze at 127A SHC, The University of Iowa, Iowa City, IA 52242, U.S.A.

body–cover hypothesis. A first attempt was offered by Fujimura (6), who related the elongation of the vocal folds to CT and TA contractile forces. No data on effective stiffnesses were available at the time, however, and the model was in fact more conceptual than quantitative. Since then, enough studies have emerged on active length–tension characteristics of isolated laryngeal muscles (7,8; Alipour-Haghighi et al., unpublished data) to warrant another look at the mechanics of F_0 control. Questions to be answered here are the following: (a) Can a simple body–cover model predict both F_0 rises and falls when the TA is activated? and (b) Under what conditions does F_0 rise, fall, or stay constant with TA activity?

To answer these questions, we begin with an analysis of vocal-fold elongation based on rotational movement of the thyroid cartilage around the CT joint. This will lead to a brief experiment to support the torque equilibrium equation derived for length change, and finally to a calculation of F_0 based on different levels of activation in both CT and TA. This is not a final report on the subject because measurements are incomplete at this stage, but we believe that preliminary reporting of the approach taken here may be of interest.

MECHANICS OF THYROID CARTILAGE ROTATION

Vocal-fold length is regulated by the combined actions of the CT and TA muscles. A simplified view of this mechanism is shown in Fig. 1. In parts (a) and (c), sketches of the lateral view of the larynx show how the thyroid cartilage can rotate with respect to the cricoid cartilage. The point of rotation is the CT joint J. A simplified conceptualization of this rotation is shown in parts (b) and (d) with the aid of the numeral "4." The cricoid cartilage and the two arytenoid cartilages are represented by the anchored, vertical stem of the numeral. The three cartilages are considered fused together into one solid stationary rod. The thyroid cartilage, represented by the 90° elbow, assumes all the motion in this model. (The motion is relative and could equivalently have been assigned to the stem, or some combination of both, as may be the case in the larynx.)

We see that contraction of the CT muscle rotates the elbow downward, thereby elongating the vocal fold by an amount ΔL. Contraction of the TA muscle, on the other hand, shortens the vocal folds by creating the opposite rotation.

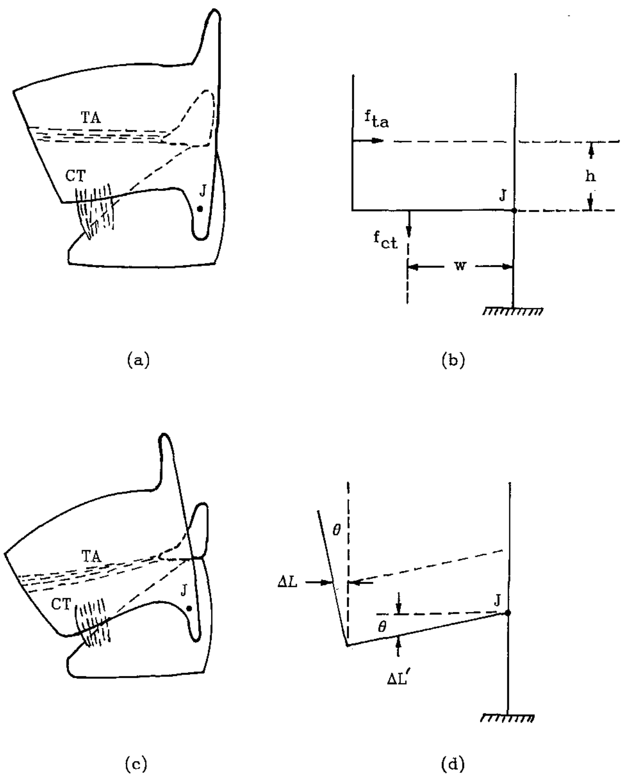

FIG. 1. Lateral view of the framework of the larynx. (a) and (c) Thyroid cartilage rotation due to cricothyroid muscle contraction. (b) and (d) Simplified sketch for calculation of torque equilibrium.

Assume the perpendicular distance from the point of rotation J to the line of action of the CT force to be w (Fig. 1b). Similarly, assume the perpendicular distance from the point of rotation to the line of action of the TA force to be h. Summation of the torques around the joint yields, for static conditions,

$$w f_{ct} = h f_{ta} + k\theta \tag{1}$$

where f_{ct} and f_{ta} are the active CT and TA forces, respectively, θ is the angle of rotation, and k is an unknown torsional stiffness that accounts for all of the passive restraining and restoring forces in the vocal folds and around the joint. In general, none of the quantities in Eq. 1 is a constant. The forces f_{ct} and f_{ta} vary with muscle activity (number of recruited motor units and firing rate of each unit) and with muscle length. The moment arms w and h can vary slightly with the angle of rotation, but for small angles may be considered approximately constant. The torsional stiffness k is also likely to be a function of θ.

Notwithstanding this complexity, it is useful to begin with the simplest possible model and add higher-order corrections only when theory and observation do not agree. Let us assume, therefore, that w, h, and k are constants, at least over some range of θ. Let us also recognize that θ is a small angle because the CT space is small in comparison with the dimensions w and h. We can then write

$$\Delta L \approx h \sin\theta \approx h\theta \tag{2}$$

and Eq. 1 becomes

$$\Delta L = \left(\frac{h^2}{k}\right)\left(\frac{w}{h} f_{\mathrm{ct}} - f_{\mathrm{ta}}\right) \tag{3}$$

This provides us with a basic beginning relationship between vocal-fold elongation and the two major opposing muscle forces f_{ct} and f_{ta}. Preliminary anatomical measurements of w and h on two canine larynges suggest that the w/h ratio is on the order of

$$w/h = (1.4 \text{ cm}/2.1 \text{ cm}) = 2/3 \tag{4}$$

If this ratio holds up for a larger number of larynges to be tested, one would conclude that the CT muscle may have a slight mechanical disadvantage over the TA muscle in changing vocal-fold length, since w/h is somewhat less than 1 and this ratio affects the balance of forces in Eq. 3. The w/h ratio also governs the relative elongations of the two muscles. If we designate $\Delta L'$ to be the shortening of the CT muscle, then by comparing the sine of equal angles θ in Fig. 1d, we see that

$$\Delta L' = (w/h)\Delta L \tag{5}$$

This relationship will be important for relating future length–tension measurements on isolated CT muscle tissue to control of vocal-fold length.

In a preliminary experiment to be described below, it has been possible to measure length changes in vivo in animals (canines) as a function of peak CT and TA muscle contraction. To relate the results to the discussion presented here, let a_{ct} and a_{ta} represent the muscle activity levels of the CT and TA, respectively, ranging between 0 and 1 (100%) of maximum activity. Furthermore, let the elongation ΔL be divided by an appropriate vocal-fold reference length L_o so that size differences between larynges can, to first order, be eliminated by normalization. Equation 3 can then be written in terms of the vocal-fold strain ϵ, such that

$$\epsilon = \frac{\Delta L}{L_o} = \frac{h^2 F_{\mathrm{ta}}}{kL_o}\left(\frac{wF_{\mathrm{ct}}}{hF_{\mathrm{ta}}} a_{\mathrm{ct}} - a_{\mathrm{ta}}\right) \tag{6}$$

where F_{ct} and F_{ta} are the maximum forces (corresponding to maximum muscle activities) developed by the CT and TA. Note that the activity of the CT muscle is weighted by the ratio of the maximum torques (wF_{ct} and hF_{ta}) developed by the two muscles around the CT joint.

Equation 6 has two coefficients to be determined experimentally, the factor in front of the parentheses and the torque ratio. We next describe an experiment conducted to estimate those coefficients.

EXPERIMENTAL EVALUATION OF COEFFICIENTS IN STATIC TORQUE EQUATION

Two large mongrel dogs weighing more than 25 kg were anesthetized and placed on a laboratory table. The superior and recurrent laryngeal nerves were exposed surgically and custom-designed electrodes were fitted around the nerves for subsequent stimulation by a Grass S88 dual-channel stimulator. Tissues between the hyoid bone and the thyroid cartilage were cut to expose the vocal folds. A video camera was placed so that it imaged the larynx perpendicular to the plane of the vocal folds. Extremely fine sutures (6–0) were used to mark two locations on the vocal folds, ~1 cm apart. The distance between the markers filled most of the video screen so that the resolution of the length change measurement was essentially limited by the resolution of the video screen.

Three conditions of supramaximal stimulation were used to obtain two extreme values and one intermediate value of vocal-fold length. The results are shown in Table 1. In the three all-or-none conditions, either the CT or the TA alone received supramaximal stimulation, or both received supramaximal stimulation. No phonation was attempted in the experiment.

Two of the three animals in Table 1 behaved similarly enough to justify averaging the strains for the

TABLE 1. *Vocal-fold length change with muscle activity*

Animal	a_{ct}	a_{ta}	$\Delta L/L_o$
1	1	0	0.37
	1	1	0.30
	0	1	−0.12
2	1	0	0.39
	1	1	0.30
	0	1	—
3	1	0	0.29
	1	1	0.13
	0	1	−0.19

three stimulation conditions. The third larynx had more negative strains, on the whole. The following empirical relationship describes the average response of the first two larynges to one significant figure,

$$\epsilon = 0.4a_{ct} - 0.1a_{ta} \quad (7)$$

$$= 0.1(4a_{ct} - a_{ta}) \quad (8)$$

which suggests that the maximum-torque ratio is 4 for these animals (see Eq. 6). The CT muscle is four times more effective in lengthening the vocal folds than the TA muscle in shortening the vocal folds. For the third animal,

$$\epsilon = 0.2[(3/2)a_{ct} - a_{ta}] \quad (9)$$

which suggests a considerably smaller maximum-torque ratio. Using these data, and previously reported data on the stress-strain relationships of the vocal-fold cover, we now attempt to model the regulation of pitch.

PITCH REGULATION WITH VOCAL-FOLD LENGTH

The most difficult problem that is faced in predicting F_0 from basic biomechanic principles is the lack of knowledge about the effective mass in vibration. It is clear that the amplitude of vibration is maximum at the medial (glottal) surface and decreases laterally into the tissue. Vibration in the thyrovocalis portion of the body is less than vibration in the cover, and vibration in the thyromuscularis portion appears to be near zero. In the vertical dimension, it is also difficult to predict what portion of tissue is in vibration. Given this difficulty, some intermediate steps are necessary to eliminate the need for direct measurement of effective mass.

Consider the vibrating portion of the vocal fold to be a fat string as shown by the dashed outline in Fig. 2. Let the cross section of the fat string be A, and let it include portions of the cover and body, as indicated. Extending the simple string formula for the lowest mode of vibration to the fat string of finite cross section, we have

$$F_0 = (1/2L)(\sigma/\rho)^{1/2}, \quad (10)$$

where σ is the mean longitudinal stress (mean tension per unit cross-sectional area A) and ρ is the mean tissue density. (Recall that in the more familiar string formula, density is expressed in mass per unit length. Here we have divided both numerator and denominator in parentheses by the cross-sectional area, which now yields the conventional tissue density in mass per unit volume.) The volumetric density can be considered constant across the body and cover. Measurements we have made on various layers of tissue (9) indicate values ranging between 1.02 and 1.04 g/cm^3, with an average value of 1.03 g/cm^3.

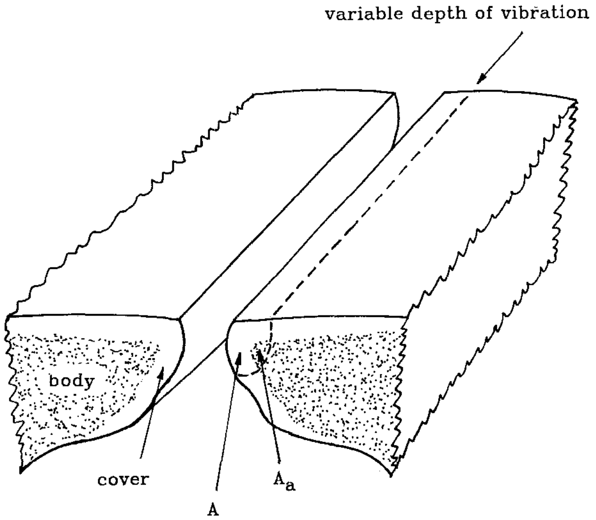

FIG. 2. Schematic of vocal fold showing body and cover. The effective depth of vibration within the vocal fold is indicated by the dashed line.

With the density ρ constant in Eq. 10, only two variables remain that affect F_0 in this simple model, L and σ. We have already quantified L in the previous section, allowing us now to focus solely on σ, the mean tissue stress. Let us assume that this mean stress is the sum of an active stress and a passive stress, such that

$$\sigma = \sigma_p + (A_a/A)\sigma_{am}a_{ta} \quad (11)$$

where σ_p is the mean passive stress over the entire vibrating cross section A, σ_{am} is the maximum active stress in the muscular cross section A_a, and a_{ta} is the TA muscle activity level, as defined previously. The equation is a direct result of summing longitudinal forces (tensions) along the vocal fold, as can be verified by multiplying both sides of the equation by A.

The passive stress can be determined indirectly from F_0 measurements on excised dog larynges. According to data reported by Titze and Durham (10), an empirical relationship is

$$F_{0p} = 38 + 175(\epsilon + 0.1) + 2450(\epsilon + 0.1)^2 + 2P_s \quad (12)$$

where F_{0p} is the fundamental frequency in hertz for the passive (excised larynx) tissue at strains ϵ ranging between -0.1 and $+0.1$, and P_s is the subglottal pressure in cm H_2O. No data are presently available for $\epsilon > 0.1$, and it should be mentioned that the rise in F_0 with P_s has been found to be >2 Hz/cm H_2O for a number of other larynges. Using Eq. 10 we now estimate the passive stress as

$$\sigma_p = 4L^2F_{0p}{}^2\rho \tag{13}$$

where

$$L = L_0(1 + \epsilon) \tag{14}$$

and combine it with the active stress in Eq. 11. The combined stress σ is then substituted back into Eq. 10 to yield the final result

$$F_0 = F_{0p}\left(1 + \frac{A_a}{A}\frac{\sigma_{am}}{\sigma_p}a_{ta}\right)^{1/2} \tag{15}$$

Several observations are in order here. First, we see that the passive F_0 is augmented by an active component that depends on an area ratio, a stress ratio, and the level of activity in the thyroarytenoid muscle. The area ratio can vary between 0 and 1, indicating that either none or all of the vibrating cross section is muscle. The stress ratio can vary over larger ranges. Alipour-Haghighi et al. (9) report a maximum active tetanic stress (σ_{am}) of 100 kPa at $\epsilon = 0$ and 125 kPa at $\epsilon = 0.4$ for the thyrovocalis muscle. The passive stress (σ_p) was reported by Perlman and Durham (11) to range between 0 and 100 kPa for a thin section of the cover (1 mm thickness) over the same range of strain. The passive stress for the muscle, on the other hand, varied over half that range (0–50 kPa). Indirect calculations of σ_p from Eqs. 12 and 13 at $\epsilon = 0.1$ and $P_s = 10$ cm H_2O reveal a passive stress of 41 kPa with a corresponding F_{0p} of 191 Hz. Given these stress ranges, it is clear that the active contribution to F_0 in Eq. 15 can range from being completely negligible (when a_{ta} is near zero) to being dominant when a_{ta} is near 1 and A_a/A is near 1. Whether or not F_0 will rise or fall with increasing a_{ta} will depend on whether a drop in F_{0p} due to a length decrease will be overcome by a rise in the active term in parentheses of Eq. 15. We will now address this question quantitatively.

CALCULATIONS OF F_0 BASED ON MUSCLE ACTIVITIES

The combined set of Eqs. 8, 12, 14, and 15 now give predictions of F_0 as a function of the CT and TA muscle activity levels a_{ct} and a_{ta}, as well as on the P_s. This is the result we were seeking. One parameter, A_a/A, is yet undetermined. This area ratio parameter, which quantifies the degree to which the body is in vibration, can be varied over its range of 0 to 1 to determine the effect on F_0. Results of a number of calculations with the four equations mentioned above are shown in Fig. 3. This is a composite graph of the dependence of F_0 on the two variables a_{ct} and a_{ta}, with A_a/A being the parameter in each of the three families of curves. The P_s is held constant at 8 cm H_2O and L_o is held constant at 1.5 cm for all cases in Fig. 3 to avoid a more complex picture. The effect of P_s on F_0 was explored in a previous article (12) and will not be considered further in this study. Effects of L_o will also be studied in a future article.

Examining Fig. 3, we first note that F_0 is a strong function of a_{ct}, as expected. The F_0 ranges from about 100 Hz to more than 300 Hz across the three families of curves that represent 0, 25, and 50% CT activity. ($a_{ta} > 0.5$ was avoided in the calculation because Eq. 12 is uncertain in this region.) Within each family of curves, F_0 can go up or down with increasing a_{ta}. If A_a/A is high, i.e., if much of the muscle is in vibration, then F_0 will rise with TA activity. On the other hand, if only the cover is in vibration ($A_a/A = 0$), the fundamental frequency drops with TA activity. For some value of A_a/A, there is very little change in F_0 with TA activity.

We have imposed the maximum active stress σ_{am} in the calculations. With the obvious limitation of $A_a/A \leqslant 1$, we see that a rising pitch with increased a_{ta} is less likely at high a_{ct} than at low a_{ct}. In other words, at high pitches, the TA is less able to match the passive stress in the cover, which is primarily localized in the epithelium (10). The TA has to let go (relax) in order for the CT to impose its full force on the epithelium (or perhaps the vocal ligament in the human larynx). This may have profound implications on register changes (12), where the TA is gradually released (or more abruptly in untrained singers) to facilitate the higher pitches. More will be reported about register phenomena in a future article.

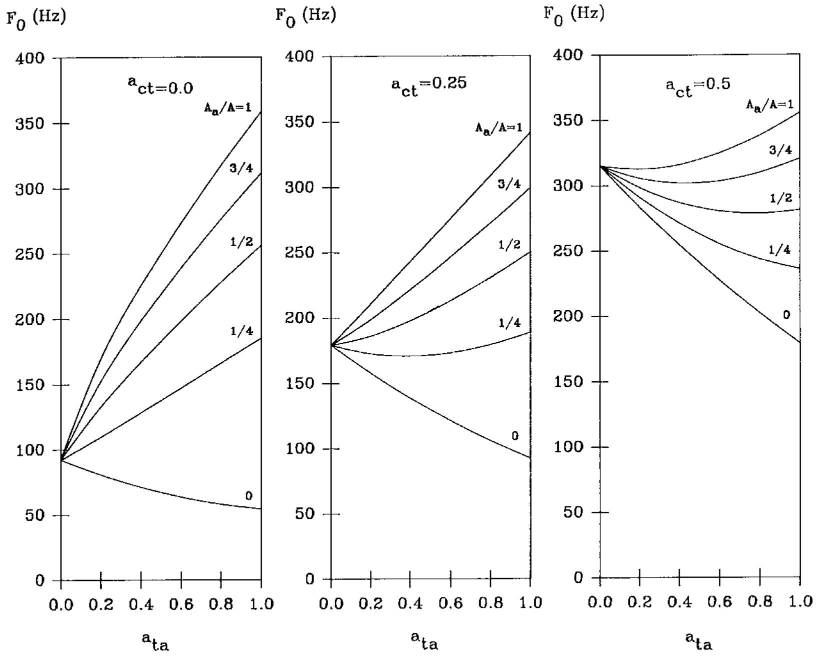

FIG. 3. Fundamental frequency as a function of cricothyroid muscle activity (a_{ct}) and thyroarytenoid muscle activity (a_{ta}) with A_a/A being the cross-sectional area ratio that defines the effective depth of vibration.

CONCLUSION

This was a preliminary study to determine if a simple body-cover model can predict both pitch rises and pitch falls with increased TA muscle activity. The answer is in the affirmative. The second question, under what conditions does F_0 rise or fall, seems to be answerable on the basis of the effective depth of vibration within the tissue. If vibration is confined to the cover only, as is generally thought to be the case in the dog larynx (because of its thick, superficial layer of the lamina propria), contraction of the TA would be a pitch-lowering mechanism exclusively. On the other hand, if the superficial layer is thin and vibration extends deep into the vocal ligament and TA muscle, TA would tend to be a pitch raiser. It is yet unknown how much control the human has over this effective depth of vibration.

Acknowledgment. Work was supported by a grant from the National Institutes of Health, grant no. NS 16320-07. We appreciate the assistance of Gail Miller in surgical procedures and data collection.

REFERENCES

1. Larson CR, Kempster GB. Voice fundamental frequency changes following discharge of laryngeal motor units. In: Titze IR, Scherer RC, eds. *Vocal fold physiology: biomechanics, acoustics and phonatory control.* Denver: The Denver Center for the Performing Arts, 1983:91–104.
2. Hirano M. Morphological structure of the vocal cord as a vibrator and its variations. *Folia Phoniat* 1976;26:89–94.
3. Baer T. *Investigation of phonation using excised larynxes* [Dissertation]. Cambridge, Massachusetts: Massachusetts Institute of Technology, 1975.
4. Stevens KN. Physics of laryngeal behavior and larynx modes. *Phonetica* 1977;34:264–79.
5. Titze IR. The physics of flow-induced oscillation of the vocal folds. Part 1. *J Acoust Soc Am* (in press).
6. Fujimura O. Body-cover theory of the vocal fold and its phonetic implications. In: Stevens KN, Hirano M, eds. *Vocal fold physiology.* Tokyo: University of Tokyo Press, 1981:271–88.
7. Alipour-Haghighi F, Titze IR, Durham PL. Twitch response in the canine vocalis muscle. *J Speech Hear Res* (in press).
8. Perlman AL, Alipour-Haghighi F. Comparative study of the physiological properties of the vocalis and cricothyroid muscles. *Acta Otolaryngol (Stockh)* in press.
9. Perlman AL, Titze IR. Development of an in vivo technique for measuring elastic properties of vocal fold tissue. *J Speech Hear Res* (in press).
10. Titze IR, Durham PL. Passive mechanisms influencing fundamental frequency control. In: Baer T, Sasaki C, Harris KS, ed: *Laryngeal function in phonation and respiration.* Boston: Little, Brown, 1987:304–19.
11. Perlman AL, Durham PL. In vitro studies of vocal fold mucosa during isometric conditions. In: Baer T, Sasaki C, Harris KS, eds. *Laryngeal function in phonation and respiration.* Boston: Little, Brown, 1987:291–303.
12. Hirano M, Vennard W, Ohala J. Regulation of register, pitch, and intensity of voice. *Folia Phoniat* 1970;27:1–20.