MODELING THE ACOUSTIC CHARACTERISTICS OF CHILDREN'S SPEECH: FUNDAMENTAL FREQUENCY

Corine Bickley

Department of Electrical Engineering and Computer Science
Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

ABSTRACT

This paper presents a model of the vibration of child-sized vocal folds. The model reflects the anatomical differences between children and adults in laryngeal structure. A scale factor, or ratio of child to adult fundamental frequency, reflects these differences. For a one-year-old child, a scale factor of 4 is derived from the model. Values of fundamental frequency are predicted and are shown to be in agreement with values measured for young children.

INTRODUCTION

Children begin to produce speech-like sounds at a very early age. Normal children communicate with speech and language skills which approximate those of adults by the age of two or three years. The changes in sound production which take place during the first few years of a child's life result from changes in the child's anatomy and in motor-control and cognitive abilities. Each of these factors constrains the sounds produced by a child. Some of the acoustic characteristics of children's sounds are a direct consequence of the size and configuration of the structures involved in speech production: the lungs, the larynx, the vocal tract. Other characteristics may be influenced most by the motor-control skills of a young child. The cognitive ability to form and manipulate mental representations of words also has a significant influence on the sound sequences produced by a child.

One aspect of an examination of children's speech is modeling the acoustic characteristics of sounds. Models which predict the acoustic characteristics of adult speech abound in the literature, but only a few models of children's vocal systems have been proposed. An approach to predicting acoustic characteristics of children's speech is uniform scaling of all vocal-tract dimensions. This simple model fails to predict spectra which are in close agreement with measured spectra of children's utterances [13]. However, vocal-tract models which incorporate more detailed anatomical constraints, such as Goldstein's [4], generate formant frequencies appropriate for children.

Analyzing the source characteristics of children's speech remains problematic. A high fundamental frequency ($F0$) is a hallmark of children's speech. The mechanisms by which children produce and control these high fundamental frequencies are not well understood. Various models, including the vibrating string and spring-mass models, have been proposed to account for the fundamental frequencies used in speech. The vibrating string model predicts the general trend of higher fundamental frequencies of children's speech than adults', due to the differences in the lengths of children's and adult's vocal folds. The spring-mass models have been successful in predicting values for the fundamental frequencies of adult speech and for airflow through the glottis during vocal-fold vibration. Difficulties arise, though, in using these models to predict an appropriate ratio of a child's $F0$ to an adult's or in predicting reasonable values for children's $F0$ as a function of anatomical measurements.

BACKGROUND

The various theories of vocal-fold vibration indicate that the frequency of vibration and the shape of the airflow waveform depend on properties of the vocal folds, including the dimensions of length, thickness, and height (vertical thickness) and the Young's modulus and effective mass of the tissue.

Measurements

Measurements have been made of length and mass of the vocal folds, the thickness of the mucosa of the folds, and the stiffness of vocal-fold tissue. The length of the vocal folds has been measured for newborns, children and adults. Hirano et al. [5] report measurements of vocal-fold length, including both the membranous and cartilaginous portions, for males of various ages. Lengths for children and adults reported by Gedgoud [3], Negus [12], and Kahane [7] are summarized by Goldstein [4]. Several values are available for newborns and adults; relatively few are reported for children between the ages of one and seven years. Hirano et al. report an average length of approximately 3 mm for one-year-old children, or approximately one-sixth as long as the vocal folds of adult males.

The thickness of the mucosa of the vocal folds has been measured by Hirano and his colleagues for newborns, children and adults. The vocal fold thickens somewhat with age, but the change in thickness is not at great as the change in length. No direct measurements of vocal-fold height are reported. We assume that the change in height is comparable to the change in thickness.

It appears reasonable to assume that the vibrating mass of the vocal folds is proportional to the combined mass of the thyroarytenoid and lateral cricoarytenoid muscles and the vocal ligament. Kahane and Kahn [8] report the mass of the vocal fold muscles: 0.87 g for adults and 0.08 g for infants. Kaneko and his colleagues[9] estimated an effective mass of 0.14 g for adult vocal folds. Based on these values for adults and infants, we calculated effective vocal-fold masses for one- and two-year-old children of 0.02 and 0.03 g, respectively.

Vocal-fold stiffness has been measured for adult humans and for young and old dogs. Kaneko and his colleagues report an effective stiffness of 7.4 × 10^4 dynes/cm for the vocal folds of adult humans. Measurements of stress/strain relationships for vocal-fold tissue of young dogs and adult dogs were performed by Perlman and Titze [14]. They found that the vocal-fold tissue of young dogs is stiffer than the tissue of adult dogs. The vocal-fold stiffness K can be determined from measurements of Young's modulus and dimensions. From Perlman and Titze's graphs of stress vs. strain, we estimated a ratio of Young's moduli of young to old tissue of 1.3. Using this ratio, the stiffness reported by Kaneko et al. and vocal-fold dimensions, we computed a value of 2.1 × 10^4 dynes/cm for the stiffness of young vocal folds. This value is consistent the range of transverse moduli reported by Kakita et al.[10].

Models

Vibrating string and spring mass models have been proposed to describe vocal-fold vibration. For each of these models, the fundamental frequency of vibration of the vocal folds can be determined. The vibrating string model is a one-dimensional model whose parameters are vocal-fold length and tension. Various spring-mass models have been proposed (for example, [6]) which model the vocal folds in terms of lumped elements representing the mass, stiffness and losses of the vocal-fold structure. In order to predict the fundamental frequency of the vocal folds, only the effective mass and stiffness of the model are needed.

A scale factor, or ratio of child to adult male fundamental frequency, reflects the differences in anatomical parameters between children and adults. For the vibrating string model, the scale factor SF_{string} depends on vocal-fold length (L) and tension (T):

$$SF_{string} = \frac{F0_c}{F0_a} = \frac{L_a}{L_c}\sqrt{\frac{T_c}{T_a}} \quad . \tag{1}$$

The subscripts $_a$ and $_c$ refer to adult and child values, respectively. Assuming that the tensions T_c and T_a of child and adult vocal folds are approximately the same, we find that $SF_{string} \approx 6$ for a one-year-old child.

The scale factor for the fundamental frequency predicted by a spring-mass model is

$$SF_{spring-mass} = \sqrt{\frac{K_c}{K_a}\frac{M_a}{M_c}} \quad , \tag{2}$$

where K is the stiffness of the vocal-fold tissue and M represents the effective mass of the vibrating vocal fold. Solving for K in terms of the Young's modulus E and the dimensions of the vocal folds gives

$$SF_{spring-mass} = \frac{FO_c}{FO_a} = \sqrt{\frac{E_c}{E_a}\frac{h_cL_c}{b_c}\frac{b_a}{h_aL_a}\frac{M_a}{M_c}} \quad , \tag{3}$$

where h and b are the vocal-fold height and thickness, respectively. Assuming the same ratio of child to adult value for both cross dimensions h and b, the scale factor for the spring-mass model reduces to

$$SF_{spring-mass} = \sqrt{\frac{E_c}{E_a}\frac{L_c}{L_a}\frac{M_a}{M_c}} \quad . \tag{4}$$

For the values listed above, $SF_{spring-mass} \approx 1.3$ for the fundamental frequency of a one-year-old child compared to an adult.

Both the vibrating string and spring-mass models predict that the $F0$ of a child's speech is greater than the $F0$ of an adult's speech. Neither prediction, however, gives a ratio which is in good agreement with the values of $F0$ of children reported by various researchers. Typical values of $F0$ for one- to two-year-old children are in the range of 300 – 500 Hz, or 3 – 4 times the $F0$'s reported for adult males.

THEORY

The vibrating string and spring-mass models capture important aspects of vocal-fold vibration, but fail to adequately model some aspects of the vocal-fold anatomy. For instance, the vibrating string model does not take into account the effect of the cross dimensions of the vocal folds on the stiffness of the structure. Another shortcoming of this model concerns the boundary conditions. The vibrating string model allows discontinuities in slope at the juncture of the cartilages and the vocal-fold tissue. The spring-mass model allows for discontinuities in both position and slope at the endpoints of the vocal folds. The specification of boundary conditions is important in analyses of the vibration of children's vocal folds; children's vocal folds are relatively shorter and thicker than adults', as shown in Fig. 1a. The attachments of the vocal folds to the arytenoid and thyroid cartilages can be expected to play a significant role in the vibration of children's vocal folds.

A model of vocal-fold vibration which reflects the anatomical structure of children's vocal folds is the bending beam model.

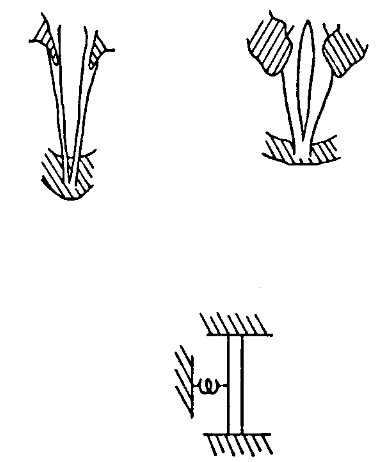

Figure 1: (a) Adult's and child's vocal-fold structures (not drawn to scale) (adapted from Bosma, 1986); (b) Bending beam model of vocal folds.

This model has been useful in predicting the vibratory motion of relatively stiff structures which are attached rigidly at their ends and vibrate a small amount in the transverse direction [15]. The traditional bending beam model can be augmented by the addition of a distributed stiffness along one side. Figure 1b shows a bending beam which is fixed at both ends and which is coupled to material on one side by means of a spring. The fixed ends model the attachment of the vocal fold to the arytenoid and thyroid cartilages. The spring models the lateral stiffness of the vocal-fold tissue.

The equation for transverse motion of the vocal-fold model shown in Fig. 1b is

$$\left(\frac{Eb^2}{12\rho}\right)\frac{d^4\xi}{dx^4} - \left(\omega^2 - \frac{K}{\rho Lbh}\right)\xi = 0 \quad , \tag{5}$$

where E represents the Young's modulus of the vocal fold, K models the stiffness of the vocal-fold tissue, b and h are the thickness and height of the vocal fold, ρ is the density of the tissue, and ξ is the transverse displacement of the fold. Four boundary conditions are imposed: continuity of displacement and of slope at both ends of the vocal fold. A solution which is a linear combination of trigonometric and hyperbolic functions is assumed. Application of the boundary conditions results in

$$\cos\beta - \frac{1}{\cosh\beta} = 0 \quad . \tag{6}$$

The variable β takes on discrete values which are found by graphical solution; the lowest non-zero value of β is approximately 4.73.

The values of ω for which equation (5) has a solution are given by

$$\omega^2 = \frac{\beta^4}{12}\frac{Eb^3h}{L^3M} + \frac{K}{M} \quad . \tag{7}$$

The first term is the square of the natural frequency of the beam model of the vocal fold assuming no lateral stiffness, and is called ω_b^2. The second term, ω_s^2, is the square of the frequency of the spring-mass model of the vocal fold. The solution of the general equation of motion shows the combined contributions of the beam and the spring character of the vocal fold structure:

$$\omega = \sqrt{\omega_b^2 + \omega_s^2} \quad . \tag{8}$$

Numerical values for ω can be found by substitution of the dimensions and tissue properties of the vocal folds.

If boundary conditions of position continuity and no stress at the endpoints are assumed (instead of continuity of position and slope), the solution of equation (5) reduces to the solution of the equation of motion for a vibrating string.

RESULTS AND DISCUSSION

The frequency of vibration, ω, of the vocal fold is a combination of the terms ω_b and ω_s. The first term of equation (7) represents the frequency of vibration due to the characteristics of the beam, where

$$\omega_b = \sqrt{\frac{\beta^4}{12}\frac{Eb^3h}{L^3M}} \quad . \tag{9}$$

For small L, as in the case of a child's vocal folds, this term dominates the expression for ω, and $\omega \approx \omega_b$. The vibration of a child's vocal folds appears to be most like that of a bending beam. For a child-sized vocal fold with length 0.35 cm, height and thickness 0.23 cm, and mass and stiffness as above, we find $\omega_b = 2810$ and $\omega_s = 1000$. The child's fundamental frequency is thus 470 Hz.

In the adult case, or for large L, the ω_s term dominates, or $\omega \approx \omega_s$, where

$$\omega_s = \sqrt{\frac{K}{M}} \quad . \tag{10}$$

The vibration of adult-sized vocal folds is similar to the vibration of a mass coupled to a spring. For an adult whose vocal folds are of length 1.7 cm, height and thickness 0.27 cm, and mass and stiffness as above, $\omega_b = 160$ while $\omega_s = 760$. The corresponding fundamental frequency is 120 Hz.

Returning to our discussion of scale factors, we calculate a scale factor relating the $F0$ of the bending beam model (appropriate for a child's vocal folds) to the $F0$ of a spring-mass model (for an adult).

$$SF = \sqrt{\frac{\beta^4}{12}\frac{E_cb_c^2}{L_c^4}\frac{b_a^2}{E_a}} \quad . \tag{11}$$

For values listed above, $SF \approx 4$.

Measurements of $F0$ of comfort-state vocalizations of young children have been reported by several researchers. Keating and Buhr [11] report $F0$ measurements for children of ages 8 months to approximately 3 years. In a study of the acoustic characteristics of vowels produced by young children of ages one and one-half to two and one-half years, we found average values of $F0$ ranging between 350 and 400 Hz[1]. These values as well as those of Keating and Buhr are shown in Fig. 2. Overlaid on these values are predicted values at ages one, two and three years. It can be seen that the predictions of the bending beam model closely approximate the data for young children.

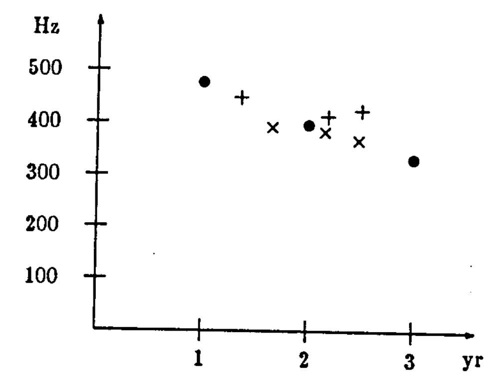

Figure 2: Predicted and measured values of $F0$. Predicted values are shown by filled circles. Averages of values reported by Keating and Buhr are shown by +'s; those of Bickley, by ×'s.

CONCLUSION

A model of vocal-fold vibration has been presented for which the expression for the fundamental frequency consists of two terms. The bending beam term depends on tissue characteristics and the connections at the ends of the vocal folds; the spring-mass term depends of the bulk characteristics of the folds. For young children, the bending beam term dominates; for adults, the spring-mass term determines the fundamental frequency.

Acknowledgement

This work was supported in part by a Whitaker Health Sciences Fellowship and in part by a LeBel Fellowship.

REFERENCES

[1] Bickley, C.A. *Acoustic Evidence for the Development of Speech.* Cambridge, MA: unpublished Ph.D. dissertation, Massachusetts Institute of Technology, 1987.

[2] Bosma, J.F. *Anatomy of the Infant Head.* Baltimore: Johns Hopkins University Press, 1986.

[3] Gedgoud, V.A. Anatomical peculiarities of the respiratory organs in children. Translated by S. Pelvoy, 1957. St. Petersburg: thesis. Referenced in U.G. Goldstein (1980) *Articulatory Model for the Vocal Tracts of Growing Children.* Cambridge, MA: unpublished Ph.D. dissertation, 1900.

[4] Goldstein, U.G. *Articulatory Model for the Vocal Tracts of Growing Children.* Cambridge, MA: unpublished Ph.D. dissertation, Massachusetts Institute of Technology, 1980.

[5] Hirano, M., Kurita, S. and Nakashima, T. The structure of the vocal folds. In M. Hirano and K.N. Stevens (Eds.), *Vocal Fold Physiology.* Tokyo: University of Tokyo Press, 1980.

[6] Ishizaka, K. and Matsudaira, M. What makes the vocal cords vibrate. In Y. Kohasi (Ed.), *Reports of the 6^{th} International Congress on Acoustics.* Tokyo: Maruzen, 1968.

[7] Kahane, J.C. The developmental anatomy of the prepubertal and pubertal larynx. Pittsburgh: unpublished Ph.D. dissertation. Referenced in U.G. Goldstein (1980) *Articulatory Model for the Vocal Tracts of Growing Children.* Cambridge, MA: unpublished Ph.D. dissertation, 1975.

[8] Kahane, J.C. and Kahn, A.R. Weight measurements of infant and adult intrinsic laryngeal muscles. Folia phoniat. 36:129-133, 1984.

[9] Kaneko, T., Masuda, T., Shimade, A., Suzuki, H., Hayasaki, K. and Komatsu, K. Resonance characteristics of the human vocal fold in vivo and in vitro by an impulse excitation. In T. Baer, C. Sasaki and K. Harris (Eds.), *Laryngeal Function in Phonation and Respiration.* Boston: College-Hill Press, 1987.

[10] Kakita, Y., Hirano, M. and Ohmaru, K. Physical properties of the vocal fold tissue: measurements on excised larynges. In M. Hirano and K.N. Stevens (Eds.), *Vocal Fold Physiology.* Tokyo: University of Tokyo Press, 1980.

[11] Keating, P. and Buhr, R. Fundamental frequency in the speech of infants and children. J. Acoust. Soc. Am. **63(2)**:567-71, 1978.

[12] Negus, V.E. *The mechanics of the larynx.* London: Heinemann Medical Books. Referenced in U.G. Goldstein (1980) *Articulatory Model for the Vocal Tracts of Growing Children.* Cambridge, MA: unpublished Ph.D. dissertation, 1929.

[13] Nordström, P.-E. Attempts to simulate female and infant vocal tracts from male area functions. Speech Transmission Laboratory Quarterly Progress and Status Report, Royal Institute of Technology, Stockholm, Sweden. **2-3**:20-33, 1975.

[14] Perlman, A.L. and Titze, I.R. Measurements of viscoelastic properties in live tissue. In I.R. Titze and R.C. Scherer (Eds.), *Vocal Fold Physiology: Biomechanics, Acoustics and Phoniatory Control.* Denver: Denver Center for Performing Arts, 1983.

[15] Woodson, H.H. and Melcher, J.R. *Electromechanical Dynamics.* New York: John Wiley and Sons, 1968.