Normal Voice in Children Between 6 and 12 Years of Age: Database and Nonlinear Analysis

Richard Nicollas, Renaud Garrel, Maurice Ouaknine, Antoine Giovanni, Bruno Nazarian, and Jean-Michel Triglia

Marseille, France

Summary: This study was carried out using a transversal design. It aimed to investigate possible changes of the normal voice in children before mutation, to create a database for the parameters used in the study, and to examine the use of fractal dimension and the largest Lyapunov exponent (LLE) in the assessment of nonpathological phenomena. Two hundred twelve children were enrolled: 111 females and 101 males; and 9 six-year-olds, 24 seven-year-olds, 18 eight-year-olds, 25 nine-year-olds, 27 ten-year-olds, 55 eleven-year-olds, and 54 twelve-year-olds. Fundamental frequency (Fo) decreased with age and was lower in boys than in girls. Jitter and shimmer did not significantly differ with age or gender. Fractal dimension and LLE were significantly lower in boys; LLE decreased with age. The present series confirmed the established findings that Fo is lower in boys than in girls, even before mutation, and decreases with age; two other classical voice analysis parameters, jitter and shimmer, also showed the same behavior as described in the literature. The study of nonlinear parameters (fractal dimension and LLE) showed that laryngeal dynamics is more stable in boys than in girls, and that stability is correlated with age.

Key Words: Voice—Children—Database—Fractal dimension—Largest Lyapunov exponent.

INTRODUCTION

Many authors have studied childhood voice abnormality, but there have been fewer studies of the normal childhood voice, and the evolution of the voice before larynx maturation has very rarely been investigated. The present study sought first to look for objective modifications in the normal voice of children aged 6–12 years, with voice samples recorded on a digital audio tape (DAT), and to implement a normal voice database. The classical objective parameters Fo, shimmer, and jitter were studied: Fo, because it has been shown to decrease with age,[1,2] and shimmer and jitter because objective voice analysis is usually based on the assessment of the regularity of intensity and frequency. Jitter is a statistical index that quantitatively assesses irregularities and studies the evolution of Fo. These parameters, however, simply characterize

Accepted for publication January 23, 2007.

Laboratoire d'Audiophonologie Expérimentale et Clinique du Laboratoire Parole et Langage, Université de la Méditerranée, Marseille, France.

Address correspondence and reprint requests to Richard Nicollas, Service d'ORL Pédiatrique, Hôpital d'Enfants de la Timone, 264 Rue Saint-Pierre, 13385 Marseille Cedex 5, France. E-mail: richard.nicollas@ap-hm.fr

Journal of Voice, Vol. ■, No. ■, pp. ■
0892-1997/$32.00

doi:10.1016/j.jvoice.2007.01.009

the vocal signal and do not shed light on the original dynamics. We therefore employed two nonlinear tools, fractal dimension and the largest Lyapunov exponent (LLE), sometimes used in pathological investigations, and examined their value in the exploration of nonpathological physiological phenomena.

A detailed introduction to nonlinear dynamics and its applications in voice analysis would be beyond the scope of this article and can be found elsewhere.[3,4] Here we shall only briefly sketch the basic concepts. The phase portrait technique, widely used in nonlinear dynamics, can be applied to voice signal analysis and consists in reconstructing the state of a dynamic dissipative system (eg, voice signal) in an embedding space or phase space. A great deal of information about the dynamics of the system can be obtained, because each experimental variable depends on interactions with the others. If the external parameters of the system are held constant, an orbit describing an asymptotic path will be observed in the phase space after a period of transition. The geometric figure thus formed is called an attractor. The attractor formed by sinusoidal signals is called a limit-cycle. A composite signal containing two independent frequencies forms a torus. Chaotic and strange attractors can be observed in cases of nonperiodic behavior. Quantitative analysis of attractors has been the focus of extensive research in theoretical physics and particularly in the field known as ''theories of deterministic chaos.'' The Lyapunov exponent (LE) is a parameter that could find clinical application: Berge et al demonstrated that the LEs obtained from an ''image'' phase space for a single variable (with varying delay coordinates) correlated with the original phase space obtained from the full set of variables.[5] The general principle of LE calculation involves computing the divergence between two initially nearby trajectories. In a stable system, neighboring trajectories remain similar, and LE is low: LE in a limit-cycle has a value of 0. In a poorly stable system that is highly sensitive to initial conditions, the two trajectories are similar at first but rapidly diverge, and LE increases. The LLE is a very useful parameter for determining the instability of a signal-generating system.[4,6–8]

The notion of dimension is very hard to define rigorously: it is not unequivocal and, importantly, different definitions may lead to different numerical results for a given set. When the results differ, the set is said to be fractal. In the words of Benoit Mandelbrot, the father of fractal geometry, ''A fractal is by definition a set for which the Hausdorff-Besicovitch dimension strictly exceeds the topological dimension.'' Laryngeal dynamics exhibits its phase states in a multidimensional embedding phase space. These states are included in a set of trajectories known as an attractor. The fractal dimension of the attractor corresponds to the complexity (ie, degrees of freedom) of the system.

The second aim of the present study was to establish a database of the normal voice in children aged 6–12 years (with a nonmature larynx) and to examine the value of two nonlinear tools (fractal dimension and LLE) for studying possible slight modifications such as premutation.

Lastly, it is now well known that many histological changes occur in the larynx during growth, from birth to puberty:[9–12] any differences between age-groups on acoustic analysis can thus be compared with the histological data in the child larynx literature.

SUBJECTS AND METHOD

Voice samples from 310 children aged 6–12 years, with no history of voice disorder and free of upper respiratory infection or ENT pathology, were studied. All parents provided written informed consent before their child was included and, in the case of girls, were asked to confirm that their daughter was strictly premenarchal and without breast growth at the moment of the study. The children were instructed to produce a sustained /a/ for several seconds, with self-selected habitual pitch and loudness. A single operator recorded all voices in a single day, in a soundproof room in the children's school. Voice samples were recorded on a Sony digital audio tape (DAT) (Sony Corporation, Tokyo, Japan) using a Sennheiser microphone (Sennheiser Electronic Corporation, Old Lyme, CT) placed 30 cm in front of and slightly above the mouth. Gain was adjusted to optimize the dynamic performance of the recording system. The signals were stored as a computer file, using the *Cool Edit 96* software. The sampling rate was 22,050 Hz, 16 bits. A 1-second stable portion (ie, excluding onset and offset) was selected for study. Recordings were first assessed by a panel of three experienced listeners (one pediatric otolaryngologist, and two pediatric speech therapists), who were instructed to select normal voices: ie, free of breathiness and roughness and (for boys) without mutation. From the recordings of 310 children, 212 samples, unanimously considered normal, premutational, and with good recording quality, were thus selected for study. The age/gender distribution was as follows: 9 six-year-olds (5 females [F]/4 males [M]), 24 seven-year-olds (12 F/12 M), 18 eight-year-olds (11 F/7 M), 25 nine-year-olds (8 F/17 M), 27 ten-year-olds (13 F/14 M), 55 eleven-year-olds (27 F/28 M), and 54 twelve-year-olds (33 F/21 M).

Acoustic analysis was performed on *.wav files using the *Diana* software. Fundamental frequency (Fo), shimmer, and jitter were calculated for each sample. The fractal dimension was calculated using the Grassberger-Procaccia algorithm.[13] After obtaining these data, the LLE was calculated by a research code derived from the study of Wolf et al[14] and used in several previous reports.[4,6–8] The present study differed from the above in that delay was calculated by first zero-crossing of the autocorrelation, the sampling rate was 22,050 Hz, the analysis window was 22,050 points, and the embedding space was 11. The reasons for those changes will be explained in the discussion. LLE evolution was correlated with the histological data found in the literature.

Statistical analysis was performed using the *Systat* software. Nonparametric statistical tests (Kruskal-Wallis and Mann-Whitney) were performed to compare indices obtained in age and gender groups. Probability values below 0.05 were considered significant.

RESULTS

Fundamental frequency decreased with age from 268.9 to 234.42 Hz in boys and from 260.92 to 239.43 Hz in girls ($P = 0.0012$); Fo was also significantly lower in boys than in girls, as shown in Table 1 ($P = 0.014$). Jitter and shimmer did not significantly vary with age (Table 2) or gender. Fractal dimension and LLE were statistically lower in boys than in girls (mean fractal dimension = 3.501 in girls vs 3.220 in boys [$P = 0.010$] and mean LLE = 4910.97 in girls vs 4423.61 in boys [$P = 0.018$]) (Table 3). LLE decreased with age from a mean of 6427.12 in 6-year-olds to 3605.84 in 12-year-olds ($P = 0.035$) (Table 4 and Figure 1).

Correlation with histology[11] found that LLE decreased with increasing type I collagen fiber in arytenoid cartilage and with increasing type III collagen in thyroid cartilage (Figure 2A and B).

TABLE 1. *Mean Fo in Males and Females According to Age-Group*

Age-Group	Males	Females
6-year-old	268.9 Hz	260.92 Hz
7-year-old	252.7 Hz	255.98 Hz
8-year-old	259.42 Hz	282.83 Hz
9-year-old	233.04 Hz	264.28 Hz
10-year-old	255.23 Hz	251.73 Hz
11-year-old	228.7 Hz	244.84 Hz
12-year-old	234.42 Hz	239.43 Hz

Notes: The difference is statistically significant ($P = 0.014$).

DISCUSSION

The purpose of this study was to assess the value of nonlinear parameters in the exploration of a slight physiological phenomenon (premutation, in this case) in normal voices. We adapted the algorithm used for exploring pathological voices in our laboratory. To determine the embedding space, we first calculated the fractal dimension from the signal, and not only from the Fo as some authors suggest.[15] Of the various methods proposed for calculating the

TABLE 2. *Mean Jitter and Shimmer According to Age-Group*

Age-Group	Jitter	Shimmer
6-year-old	0.008	0.423
7-year-old	0.013	0.519
8-year-old	0.013	0.43
9-year-old	0.029	0.559
10-year-old	0.013	0.526
11-year-old	0.017	0.456
12-year-old	0.021	0.471

Notes: No statistically significant difference is noted ($P > 0.05$).

TABLE 3. *Mean Fractal Dimension* According to Age and Gender*

Age-Group	Females	Males
6-year-old	3.967	3.907
7-year-old	3.304	2.664
8-year-old	3.491	3.006
9-year-old	3.466	3.192
10-year-old	3.412	3.138
11-year-old	3.525	3.313
12-year-old	3.342	3.317

*Its value decrease is lower in boys than in girls ($P = 0.010$).

fractal dimension, the Grassberger-Procaccia algorithm[13] was chosen for its reliability.[5] There is also discussion as to the best means of determining the embedding dimension. Many authors use Taken's theorem, which states that the embedding dimension m (with $m \geq 2d + 1$, where $d =$ fractal dimension of the attractor) guarantees a sufficient mapping. Actually, it is well known that, in an experimental series, increasing the dimension helps avoid the risk of including false neighbors. However, it has been shown by Townshend (in Aparicio-Acosta[16]) that, for adult voices, an embedding space between 3 and 5 is enough. No article in the literature provides such data in children; we therefore determined the embedding space value for LLE calculation according to Taken's theorem, making the embedding dimension equal to $2d + 1$ (with $d =$ the integer immediately above the highest calculated fractal dimension). The highest fractal dimension was 4.1, so the embedding dimension used for computing the LLE was 11.

TABLE 4. *Mean LLE According to Age and Gender*

Age-Group	Females	Males
6-year-old	6056.25	6798
7-year-old	6420.83	3890.5
8-year-old	4701.87	5362.75
9-year-old	4654.33	4632.59
10-year-old	4482.85	3282.42
11-year-old	4863.88	2984.16
12-year-old	3196.81	4014.88

LLE decreases with age ($P = 0.035$) and is lower in boys than in girls ($P = 0.018$).

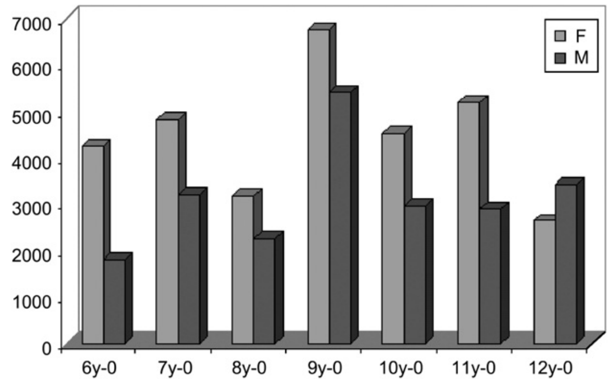

FIGURE 1. Graph showing the evolution of LLE (in bit/second) with age and gender.

There are several ways to estimate delay, two of which are often used: first zero-crossing of autocorrelation, and mutual information. The first zero-crossing of the autocorrelation function is a measure of the delay at which data become decorrelated. For a random process, the autocorrelation fluctuates randomly around zero, but for a periodic process, the autocorrelation is also periodic. Mutual information is computed in terms of Shannon entropy and can be viewed as a nonlinear generalization

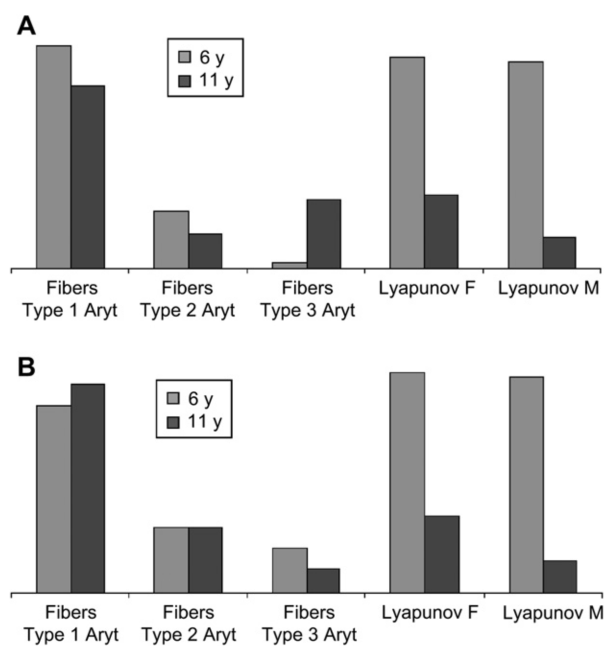

FIGURE 2. A. Compared evolution of arytenoid cartilage collagen fibers and LLE with age (nondimensional scale). **B.** Compared evolution of thyroid cartilage collagen fibers and LLE with age (nondimensional scale).

of the autocorrelation function. In the case of voice, the signal is pseudoperiodic, which is the main reason why, like many other authors, we adopted this means of estimating the delay; other methods (1/e, or mean least square) are more difficult to apply and not much more reliable in the case of a pseudoperiodic signal.

We decided to perform analysis on a sustained \a\ because speech varies a lot from one language to another, hampering comparison with other studies, and because other large series have also used sustained vowels.[2,12,13] Boys and girls under 12 years of age were selected to study the nonmature larynx, and data were screened by yearly age-group. The criteria for prepubescence were, for girls, that their parents affirmed that their daughter was premenarchal and without breast growth (two of the main criteria of puberty) and, for boys, that their voice was not discerned to be mutating by the listeners (five cases of exclusion). After this selection, the age range was nearly the same as that of Glaze et al[17] (5 years to 11 years 2 months for Glaze et al, and 6–12 years for us), but could not be objectively compared with most of the other databases or studies published, which include children from 12 to 14 years[18] or girls over 12 years.[19] Another point is the occurrence of breathiness and roughness, especially around 10 years of age, as shown by Sederholm et al:[20,21] in their study of a 10-year-old population, 9 out of 55 children (14.5%) presented such vocal features. In our series, 93 out of 310 children (30%) were excluded for breathiness and/or roughness, a rate close to that found by Campisi et al (26%).[19]

In the present study, mean Fo decreased between 6 and 12 years, as also found by Bennett and Glaze.[1,2,17] The reported values are significantly different from those of Woisard et al[22] and Kent,[23] who used other means of recording. Jitter did not significantly differ with age or gender (this index is usually calculated for pathological voices, but some authors have used it for studying normal voices,[24] which is why we tested it). Shimmer did not significantly differ between age-groups in our study, in contrast to the findings of Glaze et al.[2]

The high discriminatory power of the LE may be an advantage in quantifying physiological phenomena; the LLE, which is more reliable, was used in the present study. For the same reason, calculating LLE for normal voices is of greater interest in groups than in individuals. As described above, LLE enables instability, and not only chaos, to be assessed, and explores the stability of the whole signal, whereas jitter[4] explores only the Fo. LLE has been often used for clinical voice applications: Jiang et al used LLE to show that a chaotic mechanism of above-range phonation was suggested in a symmetric vocal-fold model.[25] Other authors used it in laryngeal paralysis[6] or dysphonia,[4,7] whatever the etiology. It has, however, never been used for exploring physiological modifications of the voice. In the case of the larynx, vocal changes are the consequence of anatomical, histological, and neurological modifications.

In the present study, LLE seemed to decrease with age. This can be explained by several factors and leads to a general concept: the main issue in the pediatric larynx is evolution. The topographic and intrinsic anatomy as well as the histology of the pediatric larynx are not fixed. As shown by Eckel et al,[9] the subglottic airway rapidly increases during the first 2 years and then follows a linear mode. During the same period, Fo decreases from 450 to 300 Hz, while membranous vocal fold increases by 1 mm (30%).[26] A growth spurt begins with puberty, at about 12 years for girls and 13 years for boys. Histologically, it is well known that collagen distribution in the vocal folds varies with age:[21] there are great differences in the distribution of immature elastic fibers in the vocal-fold structures.[11] Very great variations in collagen fiber type occur during the period between 6 years and mutation,[12] corresponding to the "premutation" identified by Hacki and Heitmuller.[27] These histological changes are correlated with LLE changes in Figure 3A and B. It appears that the quantity of type I fibers decreases with age in the posterior and mobile insertion of the vocal folds, as does LLE. The same evolution is seen for type III fibers in the thyroid cartilage, the anterior and nonmobile insertion of the vocal folds. To our knowledge, no such correlation has been noted in the literature. Increasing neurological control and feedback with age may be another important factor in voice signal stabilization. The development of laryngeal function (apprenticeship) improves coordination and decreases

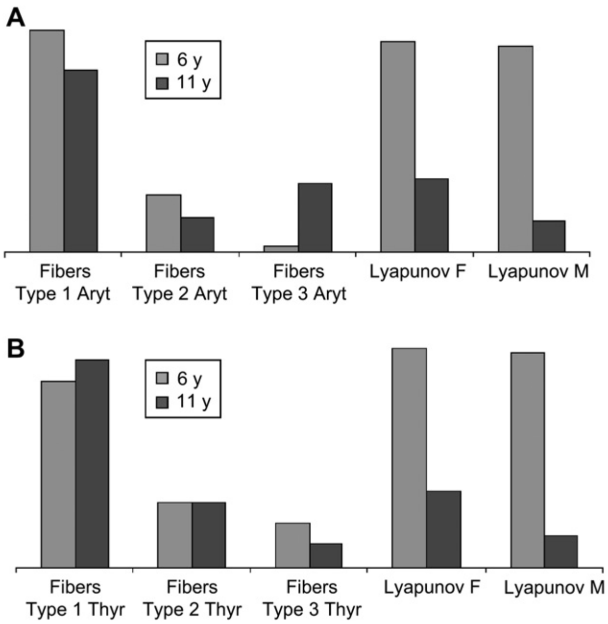

FIGURE 3. A. Compared evolution of collagen fibers and LLE with age on arytenoid cartilages (non-dimensional scale). **B**. Compared evolution of collagen fibers and LLE with age on thyroid cartilage (non-dimensional scale).

the degrees of freedom, according to Bernstein's theory.[28]

CONCLUSION

A database of objective parameters of a strictly pediatric population has been built. It further seems that LLE can be useful for assessing physiological modifications of the voice, and not only in pathological cases. Special attention has to be given to parameterization (notably, to embedding space and delay), which requires great care for reliable results to be obtained.

Acknowledgments: We thank especially the Lacordaire Private School (Marseille, France) for allowing us to record the pupils and leaving a soundproof room at our disposal, and Reza Rahbar (MD, DMD, FACS) from the Massachussets Children's Hospital of Boston for his help in writing this article.

REFERENCES

1. Bennett S. A 3-year longitudinal study of school-aged children's fundamental frequencies. *J Speech Hear Res*. 1983;26:137–141.
2. Glaze L, Bless D, Milenkovic P, Susser R. Acoustic characteristics of children's voice. *J Voice*. 1988;2:312–319.
3. Baken R. Irregularity of vocal period and amplitude: a first approach to the fractal analysis of voice. *J Voice*. 1990;4:185–197.
4. Giovanni A, Ouaknine M, Guelfucci B, Yu P, Zanaret M, Triglia JM. Nonlinear behavior of vocal fold vibration: the role of coupling between the vocal folds. *J Voice*. 1999;13:465–476.
5. Berge P, Pomeau Y, Vidal C. *Order Within Chaos: Towards a Deterministic Approach to Turbulence*. New York, NY: Wiley; 1986.
6. Giovanni A, Ouaknine M, Triglia JM. Determination of largest Lyapunov exponent of vocal signal: application to unilateral laryngeal paralysis. *J Voice*. 1999;13:1–14.
7. Yu P, Ouaknine M, Revis J, Giovanni A. Objective voice analysis for dysphonic patients: a multiparametric protocol including acoustic and aerodynamic measurements. *J Voice*. 2001;15:529–542.
8. Yu P, Ouaknine M, Giovanni A. Signification clinique du coefficient de Lyapunov dans l'évaluation objective des dysphonies. *Rev Laryngol Otol Rhinol (Bord)*. 2000;121:301–305.
9. Eckel HE, Koebke J, Sittel C, Sprinzl GM, Pototschnig C, Stennert E. Morphology of the human larynx during the first five years life studied on whole organ serial sections. *Ann Otol Rhinol Laryngol*. 1999;108:232–238.
10. Hammond TH, Gray SD, Butler JE. Age- and gender-related collagen distribution in human vocal folds. *Ann Otol Rhinol Laryngol*. 2000;109:913–920.
11. Sato K, Hirano M, Nakashima T. Fine structure of the human newborn and infant vocal fold mucosae. *Ann Otol Rhinol Laryngol*. 2001;110:417–424.
12. Cohen SR, Perelman N, Nimni ME, Mahnovski V, Cheung DT. Whole organ evaluation of collagen in the developing human larynx structures and adjoining anatomic structures (hyoid and trachea). *Ann Otol Rhinol Laryngol*. 1993;102:655–659.
13. Grassberger P, Procaccia I. Characterisation of strange attractors. *Phys Rev Lett*. 1983;50:346–349.
14. Wolf A, Swift JB, Swinney L, Vastano JA. Determining Lyapunov exponents from a time series. *Physica D*. 1985;16:285–317.
15. Boek W, Wieneke GH, Dejonckere PH. Clinical relevance of the fractal dimension of F0 perturbations computed by the box-counting method. *J Voice*. 1997;11:437–442.
16. Aparicio-Acosta FM. On some chaos techniques and the modelling of nonlinear time series. *Signal Processing*. 1996;55:269–283.
17. Glaze LE, Bless DM, Susser RD. Acoustic analysis of vowel and loudness differences in children's voice. *J Voice*. 1990;4:37–44.
18. Bohme G, Stuchlik G. Voice profiles and standard voice profile of untrained children. *J Voice*. 1995;9:304–307.
19. Campisi P, Tewfik TL, Manoukian JJ, Schloss MD, Pelland-Blais E, Sadeghi N. Computer-assisted voice

analysis: establishing a pediatric database. *Arch Otolaryngol Head Neck Surg*. 2002;128:156–160.

20. Sederholm E, McAllister A, Dalkvist J, Sundberg J. Aetiologic factors associated with hoarseness in ten-year-old children. *Folia Phoniatr Logop*. 1995;47:262–278.
21. McAllister A, Sederholm E, Sundberg J. Perceptual and acoustic analysis of vocal registers in 10-year-old children. *Logoped Phoniatr Vocol*. 2000;25:63–71.
22. Woisard V, Percodani J, Serrano E, Pessey JJ. La voix de l'enfant, évolution morphologique du larynx et ses conséquences acoustiques. *Rev Laryngol Otol Laryngol*. 1996; 117:313–317.
23. Kent RD. Anatomical and neuromuscular maturation of speech mechanism: evidence from acoustic studies. *J Speech Hear Res*. 1976;19:421–447.
24. Heylen L, Wuyts FL, Mertens F, De Bodt M, Pattyn J, Croux C, Van de Neyning PH. Evaluation of the vocal performance of children using a voice range profile index. *J Speech Lang Hear Res*. 1998;41:232–238.
25. Jiang JJ, Zhang Y, Stern J. Modeling of chaotic vibrations in symmetric vocal folds. *J Acoust Soc Am*. 2001;110: 2120–2128.
26. Titze IR. Physiologic and acoustic differences between male and female voices. *J Acoustic Soc Am*. 1989;66:60–74.
27. Hacki T, Heitmuller S. Development of the child's voice: premutation, mutation. *Int J Pediatr Otorhinolaryngol*. 1999;49(Suppl 1):S141–S144.
28. Bernstein NA. *Regulation and Control of Movements*. London: Pergamon Press; 1967.