Predicting Mutational Change in the Speaking Voice of Boys

*Michael Fuchs, †Matthias Fröehlich, ‡Bettina Hentschel, ‖Ingo W. Stuermer, ‖Eberhard Kruse, and *Daniel Knauft

*‡*Leipzig, †Erlangen, and ‖Goettingen, Germany*

Summary: The authors investigated whether acoustic speaking voice analyses can be used to predict the beginning of mutation in 21 male members of a professional boys' choir. Over a period of 3 years before mutation, children were examined every 3 months by ear, nose, and throat (ENT) and phoniatric specialists. At the same time, the voice was evaluated acoustically using analysis features of the Goettingen Hoarseness Diagram (GHD). Irregularity component and noise component, jitter, shimmer, mean waveform correlation coefficient, and fundamental frequency were determined from recordings of the speaking voice. Significant changes of acoustic features appeared 7 and 5 months before mutation onset, which indicates that vocal function is already restricted 6 months before mutation onset. This acoustic voice analysis is therefore suitable to support the care of the professional singing voice.

Key Words: Voice mutation—Acoustic analysis—Prediction—Goettingen Hoarseness—Diagram—Adolescence.

INTRODUCTION

Throughout Europe, numerous children and adolescents sing in choirs. Some perform in professional choirs with daily rehearsals and weekly concerts. Our clinical experience shows that young singers who are subject to high demands on their singing voice need a constant surveillance by singing teachers and phoniatricians, especially during phases of vocal instability as, for instance, during mutation.[1] Mutation is accompanied by a decreased vocal efficiency caused by rapid growth of the larynx and the vocal folds due to endocrine control of testosterone and growth hormones. In girls, the vocal folds become 3–4 mm longer and the voicing fundamental frequency (f_0) drops by a third of an octave. In contrast, vocal folds in boys lengthen up to 10 mm

longer, whereas f_0 drops by approximately 1 octave, which results in a higher sensitivity of the vocal apparatus during mutation in boys.

Whereas children with a normal level of vocal strain usually pass through this period without severe problems, it is known that frequent singing during mutation can affect the physiological development of the voice.[2] The changes in growth of the vocal cords in the entire larynx described substantially increase vulnerability. The conditions during mutation are often described with restrictions to the voice because of acute laryngitis.[3] Moreover, Seidner and Wendler[4] point out that the neuromuscular control during the growth phase is markedly restricted, which applies in particular to the singing voice. Young singers are endangered additionally by reduced auditory self-control and a competitive attitude during choir rehearsals. Therefore, singing teachers and phoniatricians prefer to exclude male singers from professional children's choirs at the onset of mutation.

Voice protection

Voice protection presupposes prediction of the onset of voice mutation. Difficulties originate from the great individual variability and the gradual process of mutational voice changes, which make it hard to define a clear cutoff point. Furthermore, onset of puberty and mutation accelerated over the past century, although this phenomenon seems to have slowed down in recent decades.[3,5] If studies on parameters of voice performance in children and adolescents from the beginning of 1900, the 1960s,[3] the 1970s,[6] and the 1980s[7] are considered and compared with today's result, first of all, a continuous advancement of physical development to an ever younger age is found. Since the 1970s, this trend has continued, but very much more slowly than in the comparison period 1900–1970.[5] Early mutation results in the attempt to retain boys' soprano/alto voices until just before mutation onset with the concomitant risk of voice dysfunction.

Phoniatric research must establish first the crucial time when vocal changes occur. This should be defined as exactly as possible, and symptoms of mutation must be distinguished from other phenomena, eg, inflammation of the vocal apparatus or hyperactivity. Second, the time remaining until

mutation onset has to be predicted phoniatrically before its symptoms are manifested.

The methods employed to diagnose mutation in our clinic include video-stroboscopy and well-proven analyses of voice efficiency and quality (eg, voice range profile) that are performed repeatedly during voice development. In previous studies, Habermann[8] and Bonet and Casan[9] stressed the difficulties entailed in distinguishing physiological from pathological dysphonia during mutation with particular reference to similarities in the manifestation of mutation and mild laryngitis. Blood level testosterone, growth velocity, and genital status routinely were used to predict the remaining time up to mutation or to establish whether it has occurred.[1,10] Baken and Orlikoff[11] have already pointed out the need to identify morphological changes of the vocal apparatus in its entirety with a view to evaluating mutational effects on the voice by acoustic analyses.

Pubertal voice change can be divided into three time periods: the premutation, the mutation as the specific period of vocal instability, and the postmutation. Diagnosis and prediction of mutation are focused not on interindividual differences, but on intraindividual changes. To our knowledge, so far no one has investigated systematically whether computerized analyses of voice signals can provide the information required.

Therefore, an acoustic analysis based on features taken from the Goettingen hoarseness diagram (GHD) was applied to a longitudinal dataset of continuing speaking voice signals from members of a professional boys' choir. The GHD was designed to distinguish between physiological and pathological voices and to describe the extent of deviation from a normal voice.[12,13] The underlying acoustic features that have been investigated in this study were defined in such a way as to attain a stringent description of voices over the complete range from normal voice to complete aphonia.

Computerized acoustic analyses can be applied to children and adolescents. Increasing efforts have been made to create normative databases of young voices,[14] and acoustic analyses were used to describe voice production under pathological conditions, eg, functional voice disorders,[15] vocal fold nodules,[16] in hearing-impaired children,[17] and postcochlear implantation.[18]

The GHD enables an analysis method and graphic representation of two of the most important independent voice properties. The irregularity of the acoustic signal is assessed by three acoustic parameters: jitter, shimmer, and mean waveform matching the correlation coefficient (MWC); all contribute equally to the irregularity component (IC). Jitter and shimmer are common features in the description of voice irregularities. Nevertheless, the terms do not always refer to the same physical properties. Usually, jitter is taken to refer to a variability in fundamental frequency, whereas shimmer is associated with the variability in amplitude or energy on a period-by-period basis. In the design of the GHD, the highly variable algorithms were defined in such a way that their combination was optimal in terms of several optimisation criteria (eg, stable information content; see Michaelis et al[12]). In the following, the terms ''jitter'' and ''shimmer'' always comply with the particular definitions applied in the GHD.

Besides irregularities, the additive noise content is one of the most important features. It is assessed by the noise component (NC) that is based on the glottal-to-noise-excitation-ratio (GNE). A detailed description is given by Michaelis et al.[12,19] The GHD supplies an accurate description of different voice generation mechanisms or laryngeal phonation conditions.[13] If the general laryngeal condition with regard to phonation is known, the GHD is suitable to define intraindividual voice changes as well as interindividual group differences. It allows the comparison of female and male voices, because none of its acoustic parameters depends on the determination of the fundamental frequency.

In the GHD, the periodic parameters jitter, shimmer, and MWC are calculated on the basis of the waveform matching algorithm.[20,21] This algorithm defines periods by cross-correlation of two adjacent signal segments, and thus it works on such small time scales that the word ''frequency'' cannot be used directly to refer to a repetition of periods over a certain time. However, the periodicity of the waveform matching algorithm can be averaged over a longer stretch of time (''frame'') as a parameter of the fundamental frequency. In this way, the fundamental frequency was calculated for the voice samples in this study.

The GHD is highly sensitive to changes in phonation conditions along with a good reliability, as statistical analyses of data from different recording sessions had shown.[13] The GHD is an established standard diagnostic tool in many European phoniatric departments and does not require expensive custom hardware.

The authors postulate that this acoustic analysis is suitable to provide evidence enabling differential diagnosis of mutation and to predict the beginning of the mutation.

METHODS

The authors examined 21 boys in a longitudinal study between May 1992 and November 1995 until the mutation had started for all subjects. Average age at the first examination was 11.2 years (9.3–12.5); average age at the beginning of mutation was 14.2 years (13.0–15.9). The respective voice development stage (premutation, mutation) was established on the basis of the subject's history and the reports of the phoniatrician as well as the clinical investigation including video-stroboscopy and the voice status. Apart from the symptoms described in video-stroboscopy, typical signs of mutation are a restriction of the pitch range and a change in the sound of the voice with increase of the proportions of roughness. The individual phases of mutation can thus be appraised on the basis of the clinical experience with several consecutive investigations at intervals of about 3 months.

The last examination was carried out during the mutation period. Each examination includes video-stroboscopy and standardized voice analysis, eg, voice range profile, and the recording of voice samples on digital audio tape (Denon DAT-recorder DRT-80, Denon, Pine Brook, NJ; Sennheiser stereo cond. microphone MKE-66, Sennheiser, Wedemark, Germany). The children were asked to read out the first verse of the hymn ''Jesu meine Freude.''

Generally, automatic voice analysis systems rely on recordings of isolated vowels. Speech fluidity was used in the current study because isolated vowels differ from the ''normal'' vocal use, and because subtle phonation changes before the beginning of the mutation had to be detected and quantified.
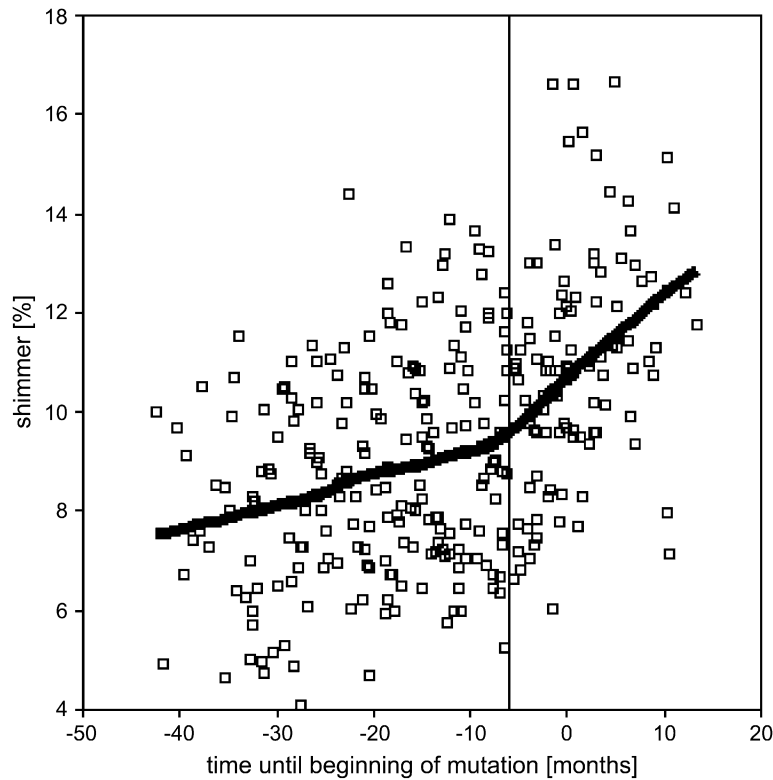
FIGURE 1. Shimmer. Scatter plot of all data courses (n = 21) from the first
examination to the beginning of the mutation (*t* = 0), with a significant change
at 6 months before the beginning of the mutation (arrow).

A preprocessing segmentation algorithm is applied
that automatically selects the signal parts that cor-
respond to voiced phonemes. These were later ana-
lyzed acoustically (for details, see Lessing et al
[22,23]). The spectral envelope was calculated by lin-
ear prediction, transformed to the Bark scale,[24] and
normalized for each time frame that did not belong
to a speech pause. A neural network consisting of

19 input cells (one for each Bark channel), 6 hidden
cells, and 1 output cell with a sigmoid characteristic
was applied to classify the current frame into
a voiced or an unvoiced phoneme.

During the calculation of the GHD, the acoustic
parameters jitter, shimmer, MWC, and GNE were
calculated for each 40-ms signal frame that had been
classified as voiced in the preprocessing segmentation

TABLE 1. *Slope of Regression Lines and Their 95% Confidence Range for Each Parameter Derived From the
Examinations of 21 Boys Either Investigated ≤6 or >6 Months Before the Beginning of the Mutation*

|	Slope (95% CI)	
	≤6 Months Before Mutation Onset	>6 Months Before Mutation Onset
Irregularity component [skt/month]	0.0111 (0.0052; 0.0169)	0.0360 (0.0209; 0.0511)
Jitter [%/month]	0.0090 (−0.0001; 0.0181)	0.1000 (0.0621; 0.1380)
Shimmer [%/month]	0.0641 (0.0294; 0.0989)	0.2772 (0.1689; 0.3855)
Mean waveform correlation coefficient [1/month]	−0.0008 (−0.0012; −0.0003)	−0.0040 (−0.0055; −0.0026)
Noise component [skt/month]	−0.0006 (−0.0067; 0.0055)	−0.0330 (−0.0506; 0.0154)
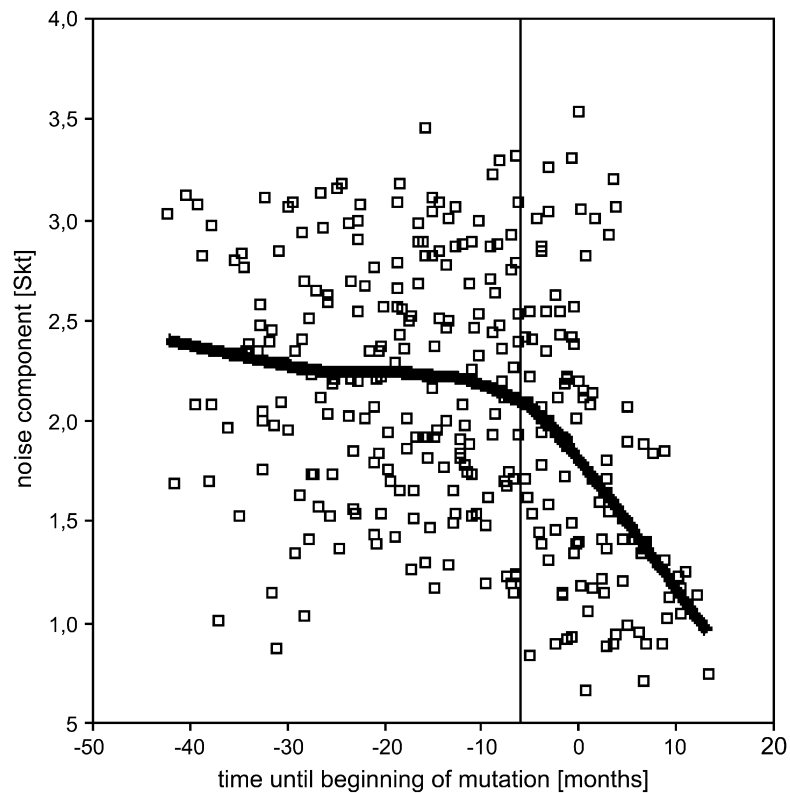
FIGURE 2. Noise component. Scatter plot of all data courses (n = 21) from the first examination to the beginning of the mutation (*t* = 0), with a significant change at 6 months before the beginning of the mutation (arrow).

(provided that the complete continuous voiced segment lasted at least 15 s). Based on these values, the IC and NC values were calculated. The data distribution that resulted from all analyzed signal segments of one recording session was characterized by the mean and the standard deviation for each coordinate and for each individual acoustic measure. In the GHD, the means and standard deviations are used to parameterize the data distribution.[13,22,23]

All parameters were represented in scatter plots. Together with the singing teacher, the authors defined the onset of the mutation (*t* = 0) under clinical and heuristic aspects as the cutoff point for obvious changes of voice quality and efficiency over time, independent of the boy's age. The authors used curve fitting to identify significant changes in the curves of each parameter retrospectively. Linear regression models were applied to repeated measurements for the time before and after the cut-point. All statistical analyses were performed using *SPSS* version 9.0 (SPSS Inc., Chicago, IL) and *S-Plus* (Insightful, Durham, NC).

RESULTS

Voice change occurred in all subjects before November 1995. On average, 15 examinations were carried out in each boy (range 5–18, total 341). The standard GHD illustration in elliptical form did not show any obvious correlation with the voice development status. Jitter, shimmer, MWC, and fundamental frequency were therefore analyzed separately. The figures show the course of the individual parameters over time before mutation and in relation to the beginning of mutation. Irrespective of the age of the individual test subjects, the time "0" of the x-axis was defined as the beginning of mutation. In this way, the curve characteristics in the period before mutation could be observed. As a result of the curve fitting, each
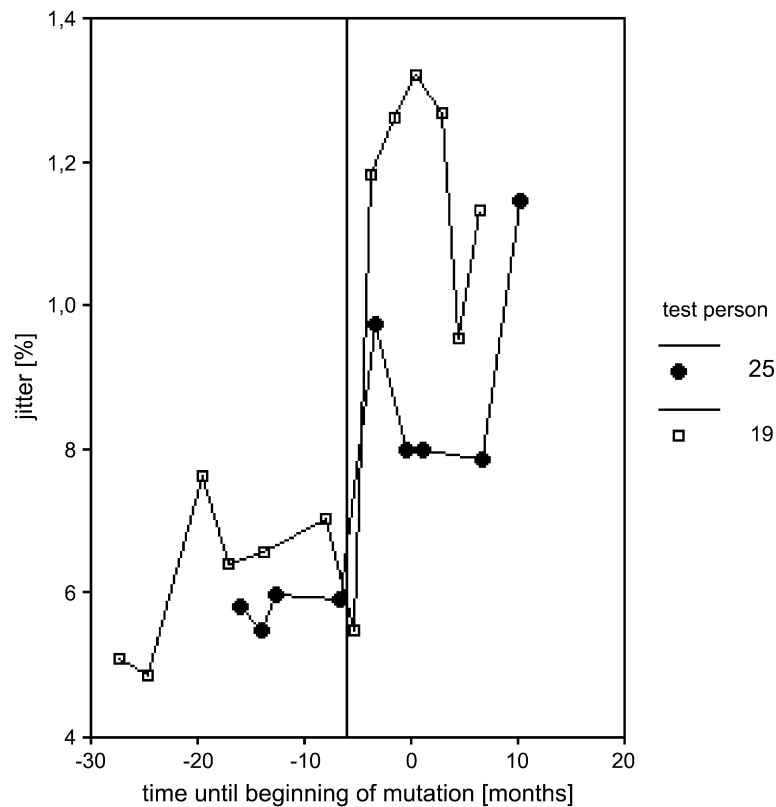
FIGURE 3. Individual curve of jitter of test persons #19 and #25.

parameter showed a conspicuous change at a cutoff point 6 months before the beginning of the mutation ($t = 0$) (for shimmer and noise component, see Figures 1 and 2). As the curves before and after this cutoff point were characterized by approximate linear courses, the authors performed regression analyses for the respective datasets from the time before and the time after this cutoff point. The slopes of the regression lines with 95% confidence ranges are presented in Table 1. The confidence ranges of the two time intervals did not overlap. The authors can therefore conclude that the slopes differ significantly. In other words, each parameter changes its course significantly 6 months before it is possible to diagnose the mutation by an auditive assessment or a clinical examination.

For the individual prediction of the time remaining until the onset of mutation, a fundamental frequency of 226.2 Hz (95% CI: 187.8; 264.4), an irregularity component of 6.3 Skt (95% CI: 5.5; 7.1), a jitter of 1.6 % (95% CI: 0.3; 2.8), a shimmer of 11.7 % (95% CI: 6.9; 16.5), a noise component

of 2.3 Skt (95% CI: 1.4; 3.2), and a mean waveform correlation of 0.9 (95% CI: 0.8; 1.0) are to be expected. After this cutoff point, irregularity component, jitter, and shimmer show a sharp rise, whereas noise component, mean waveform correlation coefficient, and fundamental frequency decrease.

Whereas the curves in Figures 1 and 2 show average values, individual analyses bear out the significant changes at 6 months before the beginning of mutation (eg, individual curves of two test persons; Figures 3 and 4). Thus, an individual prediction of the time remaining until mutation is possible by the identification of the curve "jump." The confidence range is exceeded 6 months before mutation (mentioned above), which facilitates interpretation of the curves.

CONCLUSIONS

The results confirm our hypothesis that acoustic analyses of the speaking voice are suitable to
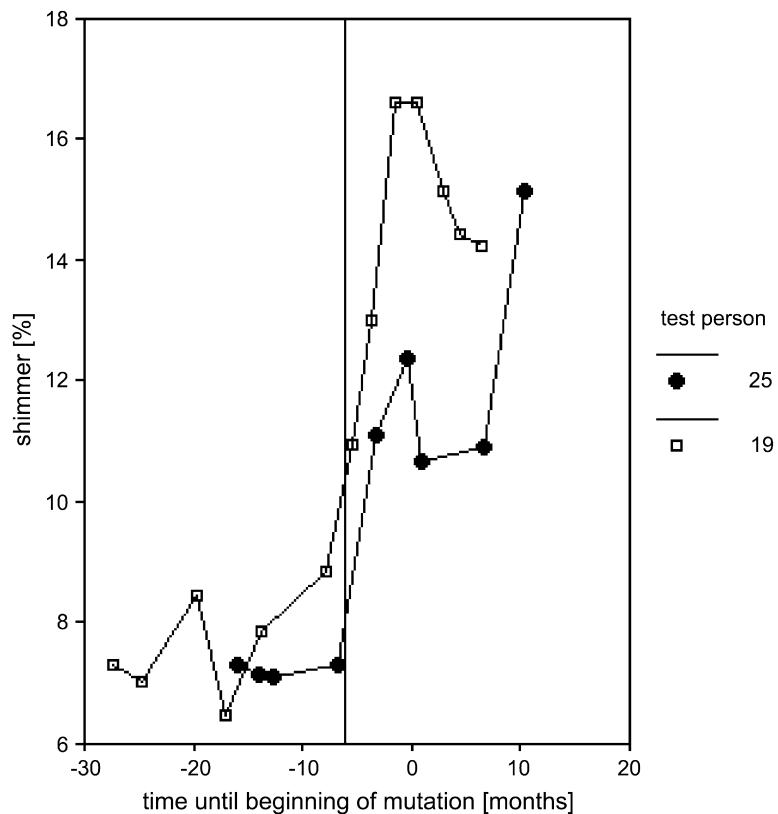
FIGURE 4. Individual curve of shimmer of test persons #19 and #25.

support the differential diagnosis of mutation and to predict its onset. The 3-month schedule applied in our study allows changes of the voice quality during voice development to be detected before they are perceptible on the basis of auditive assessment or clinical methods. These analyses can support established methods to predict the remaining time that a boy will be able to sing in soprano/alto, and to spare the singer's daily vocal strain before onset of mutation. In this way, the method helps the conductor, the singing teacher, and the phoniatrician to provide make plans for every individual boy's voice in a choir.[10]

The mean fundamental frequency decreases already 6 months before the mutation during the period of premutation. This correlates with physiological development, laryngeal growth, and a subsequent decrease of mean fundamental frequency. The voice range changes periodically without obvious effect on voice quality and efficiency.[1]

Members of children's and youth choirs usually change from soprano to alto during this phase at the latest.

On the assumption that voice quality deteriorates in the period after mutation, the irregularity component, noise component, jitter, and shimmer must be expected to increase, whereas the mean waveform correlation coefficient decreases at the beginning of this period. All parameters showed a significant change in their courses as early as 6 months before the mutation. This means all parameters change at the same time before a loss of the voice quality is perceptible. All GHD-based parameters are independent of the fundamental frequency,[12] so that these observations cannot be attributed to the decrease in fundamental frequency that occurs simultaneously. For a reliable prediction of the time remaining until mutation commences, several examinations are necessary during the premutation period and even the time before because the change
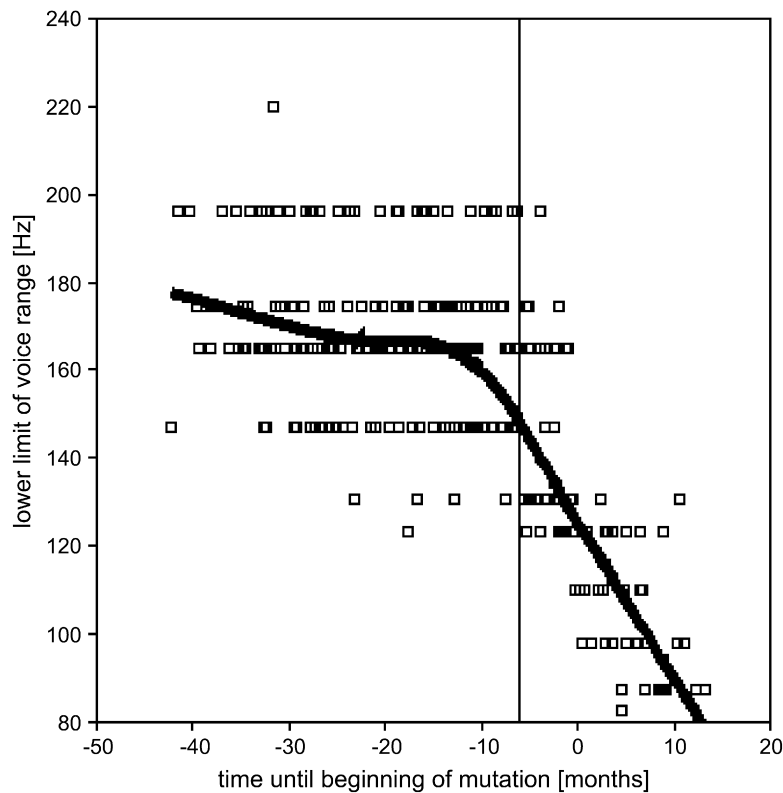
FIGURE 5. Scatter plot of all data courses (n = 21) from the first examination to the beginning of the mutation (*t* = 0): lower limit of voice range.

of the course is discernible only by a documentation and comparison of at least three separate examinations.

Although this article was focused on speaking voice, the authors also found noticeable changes of singing voice, eg, a reduction of the upper and lower limit of voice range within the voice range profile in corresponding figures (Figure 5). On the other hand, there was no significant change in the voice range.

At the moment, a predictive statement and a differential diagnosis are only possible retrospectively. In this study, the data pool was too small to create normative data for boys. The authors suggest that acoustic analyses should be included in regular examinations of boys' voices.

The noise component shows a significant decrease starting 6 months before the onset of mutation instead of an expected increase that may be explained as follows: On the one hand, it seems plausible that compensation mechanisms may occur with regard to the changes in vocalization because the vocally trained child aims unconsciously for a sound that is free of breathy and throaty parts. It can be assumed that a gradual deterioration of the voice quality can thus be compensated to some extent to conceal the approaching mutation.[1] On the other hand, one cannot rule out an effect of the algorithm that calculates the glottal-to-noise excitation ratio due to the small basis of the current data. Even though the glottal-to-noise is in principle independent of the fundamental frequency, a systematic change caused by a decreasing fundamental frequency may lead to a slight systematic trend.

Irrespective of the cause of the decrease in the noise component, the data show that all measures may be used to assess initial morphological changes accurately, as well as the resulting limitation of vocal efficiency and, possibly, vocal compensation.

The results confirm our hypothesis that acoustic analyses using the GHD are suitable to provide evidence for differential diagnosis of the mutation

and for prediction of its beginning. The information that nonaudible changes of voice occur 6 months before the beginning of mutation in the GHD permits reliable evaluation of the time remaining until mutation as well as a differential diagnostic statement on the presence of the mutation. This supplements and enhances the phoniatric examination and other established methods that predict the time of the mutation. This method can be combined, eg, with determination of the level of testosterone in the blood that enables relatively exact prediction of the time remaining up to mutation onset.[1] The determination of the growth rate is also appropriate to appraise the voice development before beginning of mutation when the lowest point of the curve of the growth rate is determined at the beginning of puberty.[10] Simultaneous determination of these parameters and the interpretation of the results together with the acoustic analyses described in this article are of great assistance in the appraisal of voice development by the phoniatrician and the voice teacher.[25,26] Further studies are required to show whether the statistical reliability can be enhanced by combining the individual techniques.

The main advantage of the new method consists in its noninvasive character. Furthermore, it does not impose any strain on the children. GHD does not require any expensive custom hardware and can be used in daily clinical practice. However, a prerequisite for a proper interpretation is that the acoustic analyses are applied to a longitudinal dataset.[1] The concept of an interdisciplinary management to safeguard professional voices in childhood and adolescence includes a periodic phoniatric examination in periods of vocal instability. The new method can be integrated into these investigations. The authors recommend examinations at intervals of about 3 months, which should include phoniatric examination including video-stroboscopy, voice status including voice range profile, physical examination including genital status and growth rate, measurement of the capillary blood level of testosterone, and acoustic analyses using GHD. All methods can be combined depending on the particular question under examination and the results of the preceding investigations.

Further studies are required to establish whether other changes of the singing voice can be verified with the GHD. Reference values for GHD in childhood and adolescence are required.

REFERENCES

1. Fuchs M, Behrendt W, Keller E, Kratzsch J. Methods for prediction of the beginning of mutation in boys voices: investigations in singers of the Thomaner Choir Leipzig. *Folia Phoniatr Logop*. 1999;51:261–271.
2. Biesalski P, Frank F. *Phoniatrie–Pädaudiologie. Vol. 1*. Stuttgart, Germany: Thieme; 1994:297–298.
3. Frank F. Mutation in consideration of acceleration. *Heilpädagogik*. 1969;12:18–20.
4. Seidner W, Wendler J. *The Singer's Voice*. Berlin: Hentschelverlag Kunst und Gesellschaft; 1982:175–177.
5. Schilling A, Karthaus K. Acceleration and mutation. *HNO*. 1961;9:271–275.
6. Frank F, Sparber M. New evidence of voice ranges in children. *Folia Phoniatr Logop*. 1970;22:397–402.
7. Doll J, Frank F, Sparber M. New evidence of vocal development in 12- to 16-years old boys. *Sprache Stimme Gehör*. 1986;10:143–146.
8. Habermann G. *Voice and Speech–An Introduction in its Physiology and Hygienics*. Stuttgart, Germany: Thieme; 1986:146–147,247,256.
9. Bonet M, Casan P. Evaluation of dysphonia in a children's choir. *Folia Phoniatr Logop*. 1994;46:28.
10. Fuchs M, Behrendt W, Kratzsch J, Keller E. Forecast of voice mutation at singers of professional boy's choirs with parameters of growth and puberty, insulin-like growth factor I and testosterone—investigation on singers of the Thomanerchor Leipzig. *Horm Res.* 1997;48:133.
11. Baken RJ, Orlikoff RF. Voice measurement: is more better? *Log Phon Vocol*. 1997;22:147–151.
12. Michaelis D, Fröhlich M, Strube HW. Selection and combination of acoustic features for the description of pathologic voices. *J Acoust Soc Am*. 1998;103:1628–1639.
13. Fröhlich M, Michaelis D, Strube HW, Kruse E. Acoustic voice analysis by means of the hoarseness diagram. *J Speech Lang Hear Res*. 2000;43:706–720.
14. 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.
15. Niedzielska G. Acoustic analysis in the diagnostic of voice disorders in children. *Int J Ped Otorhinolaryngol*. 2001;57:189–193.
16. Niedzielska G, Glijer E, Niedzielski A. Acoustic analysis of voice in children with noduli vocales. *Int J Ped Otorhinolaryngol*. 2001;60:119–122.
17. Monsen RB. Acoustic qualities of phonation in young hearing-impaired children. *J Speech Hear Res*. 1979;22:270–288.

18. Szkiekowska A, Maniecka-Aleksandrowicz B, Dolecki J. Voice rehabilitation in children post cochlear implantation. *Central East Eur J Otorhinolaryngol Head Neck Surg*. 2000;5:23–27.

19. Michaelis D, Gramss T, Strube HW. Glottal-to-noise excitation ratio—a new measure for describing pathological voices. *Acta Acustica*. 1997;83:700–706.

20. Milenkovic P. Least mean square measures of voice perturbation. *J Speech Hear Res*. 1987;30:529–538.

21. Titze IR, Liang H. Comparison of F_0 extraction methods for high-precision voice perturbation measurements. *J Speech Hear Res*. 1993;36:1120–1133.

22. Lessing J, Fröhlich M, Michaelis D, Strube HW, Kruse E. A neural network based method to assess voice quality from continuous speech for different voice disorders. Advances in quantitative laryngosocopy, voice and speech research. In: Braunschweig T, Hanson J, Schellhorn-Neise P, Witte H, eds. *Proceedings of the 4th International Workshop*. University of Jena; 2000:124–131.

23. Lessing J, Fröhlich M, Michaelis D, Strube HW, Kruse E. Use of neuronal networks for assessment of voice quality in pathological voices. In: Gross M, ed, *Aktuelle Phoniatrisch-Pädaudiologische Aspekte 1998/99*. Vol. 6. Heidelberg, Germany: Median; 1999:39–43.

24. Zwicker E, Fastl H, Frater H. *Psychoacoustics: Facts and Models*. Vol. 22. Berlin: Springer; 2000.

25. Behrendt W, Fuchs M. Mutation of singing voice in adolescence. In: Gundermann H, ed. *Die Ausdruckswelt der Stimme*. 1st ed. Heidelberg, Germany: Hüther; 1998:122–127.

26. Fuchs M, Behrendt W. Diagnostical value of voice range profile for phoniatrical care of professional singing voice during mutation. In: Gross M, ed. *Aktuelle Phoniatrisch-Pädaudiologische Aspekte1999/2000*. Vol. 7. 1st ed. Heidelberg, Germany: Median; 2000:133–138.