Published in final edited form as:

Laryngoscope. 2012 November ; 122(11): 2511–2518. doi:10.1002/lary.23568.

Spatiotemporal analysis of vocal fold vibrations between children and adults

Michael Döllinger, Ph.D.^{a,*}, Denis Dubrovskiy, M.Sc.^a, and Rita Patel, Ph.D., CCC-SLP^b

^aUniversity Hospital Erlangen, Medical School, Department of Phoniatics and Pediatric Audiology, Erlangen, Germany

^bUniversity of Kentucky, Division of Communication Sciences and Disorders, Department of Rehabilitation Sciences, USA

Abstract

Objectives/Hypothesis—Aim of the study is to quantify differences in spatiotemporal features of vibratory motion in typically developing pre-pubertal children and adults with use of high speed digital imaging.

Study Design—Prospective case-control study.

Methods—Vocal fold oscillations of 31 children and 35 adults were analysed. Endoscopic high-speed imaging was performed during sustained phonation at typical pitch and loudness. Quantitative technique of Phonovibrograph was used to compute spatiotemporal features. Spatial features are represented by opening and closing angles along the anterior and posterior parts of the vocal folds, as well as by left-right symmetry ratio. Temporal features are represented by the cycle-to-cycle variability of the spatial features. Group differences (adult females, adult males, and children) were statistically investigated.

Results—Statistical differences were more pronounced in the temporal behavior compared to the spatial behavior. Children demonstrated greater cycle-to-cycle variability in oscillations compared to adults. Most differences between children and adults were found for temporal characteristics along the anterior parts during closing phase. The spatiotemporal features differed more between children and males than between children and females. Both adults and children showed equally high left-right symmetry.

Conclusions—Results suggest a more unstable phonation in children than in adults yielding increased perturbation in periodicity. Children demonstrated longer phase delay in the anterior/posterior and medio-lateral parts during the opening phase compared to adults. The data presented may provide the bases for differentiating normal vibratory characteristics from the disordered in the pediatric population and eventually assist in aiding the clinical utility of high speed imaging.

Keywords

high speed digital imaging; pediatric voice; pediatric vibratory dynamics; Phonovibrograph

*Corresponding author: Michael Döllinger, Ph.D., University Hospital Erlangen, Department of Phoniatics and Pediatric Audiology, Bohlenplatz 21, 91054, Erlangen, Germany, Michael.doellinger@uk-erlangen.de, Contact: 0049-(0)9131-85 33 814, Fax: 0049-(0)9131-85-39 272.

Financial Disclosure: None

Conflict of Interest: None

INTRODUCTION

In the United States, over one million or 6–9%¹ of all children suffer from dysphonia.² Hoarseness can be detrimental to children both psychosocially^{3,4} and academically.⁵ Hence early identification and restoration of optimal vocal health is critical. Early identification is dependent upon accurate diagnosis of disordered vibratory patterns. Fundamental to appropriate differentiation of disordered vibratory patterns, is the knowledge of what constitutes normal vibratory function in the pediatric population.

Laryngeal visualization of vibratory patterns is essential for appropriate assessment and treatment of most voice disorders; however, quantitative assessments of vocal fold vibratory changes associated with growth and development in children are lacking, despite the recognition that the pediatric larynx is not merely a miniature version of the adult larynx. Children have a shorter vocal tract,^{6,7} anteriorly placed larynx, a shorter membranous portion of the vocal folds^{8,9} and less differentiated vocal ligament^{8,10–12} with thicker mucosa until the age of 13¹² to 16⁸ years. Not only are there anatomical differences between children and adults, there are group and gender differences in the use of respiratory and acoustical systems at prepubertal ages. Children have higher fundamental frequency, higher variation in fundamental frequency,¹³ increased signal-to-noise ratio,¹³ 50%–100% greater lung pressure,¹⁴ and higher sound pressure level,¹⁵ than adults. Empirical evidence from measurements of aerodynamic characteristics,^{14–19} acoustic characteristics,^{20–25} and electroglottographic characteristics,^{26,27} emphasize the influences of laryngeal structures on vibratory function as a primary factor affecting children's vocal physiology. How these differential mechanical properties of the body (muscle) and cover (mucosa) of the pediatric vocal folds translate to unique spatiotemporal properties of vibratory motion is not known.

Because children are functionally different from adults, this study proposes to clinically quantify spatiotemporal differences in vibratory dynamics between children and adults as observed on laryngeal imaging. Studies have reported feasibility of stroboscopic examination with the use of flexible²⁸ and rigid endoscope²⁹ in the pediatric population. Pediatric laryngostroboscopy was recommended for comprehensive diagnostic differentiation between benign mucosal lesion and identification of inflammatory processes, especially in children with a prolonged history of dysphonia.³⁰ However, due to variability in attention span and participation in children, it is often difficult to obtain pediatric phonation samples of greater than 2–3 seconds,²⁹ resulting in non-interpretable stroboscopic findings. Unlike stroboscopy, High Speed Digital Imaging (HSDI) offers a means for directly measuring spatiotemporal characteristics of vocal fold vibrations^{31,32} as it records actual cycle-to-cycle variation^{33,34} in vibratory motion with superior temporal resolution of up to 4000 frames per second. This level of detail is especially critical for comprehensive clinical appraisal of children's voice production that has a higher habitual fundamental frequency and shorter phonation duration, during endoscopic imaging, compared to adults.

In order to develop age-specific models of pediatric voice production to facilitate clinical differentiation of normal from disordered pediatric vocal function, it is important to objectively quantify the spatial and temporal differences in laryngeal dynamics of the typically developing pediatric larynx. The quantitative image-based analysis technique called Phonovibrogram (PVG)^{35–37} appears to be ideal for quantifying spatiotemporal features of pediatric vocal fold motion. The PVG is a comprehensive approach of transforming motion into a static image with geometric features, where the vocal fold vibrations can be quantified from the geometric features of the image.³⁶ Objective³⁸ and subjective³² assessment of PVG has been proven to have high accuracy and reproducibility in quantifying and differentiating normal^{39,40} and disordered^{41,42} vocal fold vibrations in adults. Recent work on quantitative⁴³ and subjective⁴⁴ analysis of high speed videos in

adults have revealed mild anterior-posterior, left-right, and phase asymmetries in normal phonation. To the best of our knowledge the aforementioned studies investigating spatiotemporal differences have been confined to the adult population. The specific purpose of this prospective study was to assess whether prepubertal differences in spatiotemporal vibratory motion is present compared to adults. More specifically spatiotemporal differences in opening and closing phases of the vibratory cycle were compared between the left & right vocal fold, and along the anterior and posterior parts of the vocal fold between children and adults.

METHODS

Participants

A total of 40 children, age range (5–11 years) and 35 adults (21–45 years) were recruited at the University of Kentucky, Vocal Physiology and Imaging Laboratory after signing of appropriate IRB approved informed consent and assent forms. Adult and children participants were recruited from IRB approved advertisement fliers placed around the University of Kentucky Campus and at University of Kentucky Children's Hospital. IRB approval was obtained from both the University of Kentucky and University of Erlangen to perform data analysis on de-identified data. Participants were included in the study if they met the following criterion: had negative histories of vocal pathology, not being professional voice users, perceptually judged to have normal voice by a certified speech language pathologist specializing in voice disorders. Participants going through puberty as identified via case history were excluded. Selection criteria for adult participants were similar to those of pediatric group, except that the adult participants had a negative history of smoking.

High Speed Digital Imaging (HSDI)

Sustained phonation on the vowel /i/ was recorded at participant's typical pitch and loudness. The recordings were performed by a digital gray scale KayPentax high speed system model 9710 with a sampling rate of 4000 frames per second with a spatial resolution of 512×256 pixels for a maximum duration of 4.094 seconds. The camera was coupled to a 70° KayPentax endoscope. Simultaneous acoustic signal was recorded at 50kHz. Acoustic recording was used to confirm the presence of steady state phonation and participants' task performance of typical pitch and loudness. A xenon light source of 300 Watts was used for imaging. Participants performed three practice trials of sustained phonation on the vowel /i/ at typical pitch and loudness, prior to recording with HSDI. Only one recording was performed and analyzed per participant.

Data Segmentation and Parameter Extraction

Image segmentation,³⁵ PVG computation³⁷, extraction and computation of parameters were performed by applying a software application (*Glottal Analysis Tools, GAT*), which was developed at the Department for Phoniatics and Pediatric Audiology, Erlangen - Germany. Prior to data segmentation of the high speed videos, the data from all participants were verified for task performance, based on acoustic analysis of sustained phonation on the vowel /i/. From each high speed recording, a sequence of steady state phonation of 50 oscillation cycles (160 ms to 460 ms corresponding to 110 Hz – 310 Hz fundamental frequency) was selected and analysed. The mean fundamental frequency for females was 251 Hz (SD = 31 Hz), male was 145 Hz (SD = 36 Hz), children was 285 Hz (SD = 31 Hz). Within each sequence, the oscillating glottis, the left and right vocal fold edges were segmented and the glottis midline was determined (Figure 1) for generation of the PVG.³⁷

The PVGs (Figure 1) were computed and normalized for each cycle.^{37,41} From the PVG, opening (*O*) and closing (*C*) angles of the left (*L*) and right (*R*) vocal folds, for the 50%

deflection status were automatically computed for anterior (*A*) and posterior (*P*) parts of the vocal fold.³⁷ The PVG angles represent the opening and closing behavior of the anterior and posterior region along the vocal fold margins (Figure 2). The PVG angle parameters were shown to have high accuracy and reproducibility in healthy adult phonation.^{39,40} Therefore, the PVG opening and closing angles were used to quantify the opening and the closing phases of the glottal cycle along the anterior and the posterior parts of the vocal folds.⁴⁰

Briefly, the PVG angles³⁷ were computed as follows: the vocal folds were initially separated at the medio-lateral position along the anterior and posterior parts of the vocal folds. The time point representing the deflection of 50% of the amplitude was selected to represent the closing and opening period for each position along the vocal fold edges (Figure 2). In other words, the resulting contour corresponds to the point in time during which a vocal fold point has reached half of its displacement between minimum and maximum deflection. The deflection status or contour threshold of 50% was selected, as it has been proven to incorporate many characteristic behavior of vibratory motion.⁴¹ This deflection threshold of 50% of the amplitude was applied individually to the anterior and posterior parts of the vocal folds.

Subsequently, a regression line for these positions was computed and the angles for opening and closing at anterior and posterior positions were calculated, yielding 8 angles for each oscillation cycle.³⁹ The angle value of 90° indicates that the glottal parts forming the angles move synchronously with no time delay. For the opening process, angles > 90° suggest that the anterior/posterior glottal parts begin to open earlier than the medio-lateral parts of the vocal fold. Whereas, angles < 90° suggest that the medio-lateral part of the vocal folds begin to open earlier than the anterior/posterior parts. For the closing process, angles > 90° suggest that the anterior/posterior glottal parts begin to close later than the medio-lateral parts of the vocal fold. Whereas, angles < 90° suggest that the medio-lateral part of the vocal folds begin to close later than the anterior/posterior parts (Figure 2).

The PVG angles are abbreviated as follows: e.g. angle for left (*L*) vocal fold during closing (*C*) in anterior area (*A*) is abbreviated as *LCA* (Figure 2). The eight PVG angles (*LOA*, *LOP*, *LCA*, *LCP*, *ROA*, *ROP*, *RCA*, *RCP*) were averaged over the sequence of 50 oscillatory cycles and their standard deviations (*LOA_s*, *LOP_s*, *LCA_s*, *LCP_s*, *ROA_s*, *ROP_s*, *RCA_s*, *RCP_s*) were computed for each participant. The angles reflect spatial (i.e. positioning of the vocal folds) as well as temporal (i.e. changes of vocal fold positioning during oscillation) characteristics of the vocal folds. The standard deviations reflect only temporal behavior and can be viewed as perturbation measures or as regularity parameters of the oscillations. This yields a total of 16 parameters for each recording.

Additionally, left-right ratios reflecting the degree of left-right symmetry between the vocal folds were computed. The left-right ratios of *Ratio_OA*, *Ratio_OP*, *Ratio_CA*, *Ratio_CP*, were computed by dividing the left four angles by the corresponding right four angles (Figure 2) for each oscillation cycle.

$$Ratio_OA = \frac{LOA}{ROA}, Ratio_OP = \frac{LOP}{ROP}, Ratio_CA = \frac{LCA}{RCA}, Ratio_CP = \frac{LCP}{RCP}.$$

The ratios were averaged over the sequence of 50 cycles, yielding four mean ratios (*Ratio_OA*, *Ratio_OP*, *Ratio_CA*, *Ratio_CP*) and four standard deviation values (*Ratio_OAs*, *Ratio_OPs*, *Ratio_CAs*, *Ratio_CPs*) for each participant. Ratio values equal to 1 correspond to perfect symmetry between the left and the right vocal fold. The further away the ratio is from 1, the more asymmetry is present between left and right vocal fold

vibrations. The standard deviations of the ratios can again be considered as perturbation measure and reflect the stability or periodicity of the oscillations.

In summary, a total of 24 PVG parameters (12 spatial features and 12 temporal features indicated by “s”) were computed and statistically analyzed for detecting spatiotemporal differences between children and adults. Hereby, the spatial features are represented by the opening and closing angles along the anterior and posterior parts of the vocal folds (*LOA*, *LOP*, *LCA*, *LCP*, *ROA*, *ROP*, *RCA*, *RCP*), as well as by left-right symmetry ratio (*Ratio_OA*, *Ratio_OP*, *Ratio_CA*, *Ratio_CP*). Temporal features are represented by the cycle-to-cycle variability of the spatial features (*LOA_s*, *LOP_s*, *LCA_s*, *LCP_s*, *ROA_s*, *ROP_s*, *RCA_s*, *RCP_s*, *Ratio_OAs*, *Ratio_OPs*, *Ratio_CAs*, *Ratio_CPs*).

Statistical analysis

Group differences between the three groups (adult females versus adult males, females versus children, and male versus children) were analysed for all the 24 parameters. Shapiro-Wilk tests were used to test for normal distribution for all 24 spatiotemporal parameters. For normally distributed parameters, ANOVA was performed following least-square difference (LSD) post-hoc tests. For non-normal distributed parameters, Kruskal-Wallis tests were performed. The significant levels of the Kruskal-Wallis tests were Bonferroni corrected. Subsequently, post-hoc tests (Mann-Whitney U) were performed with significant level $p = 0.05$, since a Bonferroni correction was already performed before. Post-hoc tests were only performed if ANOVA or Kruskal-Wallis tests showed significant levels. Results were considered significant for, $p = 0.05$, if not Bonferroni corrected. Statistical testing was performed using IBM SPSS Statistics 19.0.

Statistics on single parameters only enable to analyse group differences, since there is always a certain amount of interference. Hence we performed a multi-parameter approach to investigate if a combination of several parameters enables to assign single subjects to their specific group males, females or children. We applied an unsupervised Clustering approach (method: X-means clustering, measure: cosine similarity) with 10-times stratified cross validation. For the clustering procedure, the parameters were z-transformed for standardisation. High correct classification results would indicate that combined parameters reflect the group characteristics and allow for separating or classifying the different groups. Unsupervised clustering was performed using the open source software “Rapidminer”.

RESULTS

Due to heightened gag reflex data from 9 children could not be obtained. High speed recordings of a total of 66 participants, including 31 (5–11 years) children and 35 adults (21–45 years) were analysed (Table 1).

Computed absolute values from the 8 PVG angles as well as their 8 standard deviations (stability) are summarized in Figure 3 and Figure 4. Already evident are two observations: (1) The opening angles in children are mostly higher than the closing angles (Figure 3), reflecting a longer opening than closing duration of the vocal folds. (2) The standard deviations of the angles increase from males to females to children (Figure 4). The left-right ratio parameters are summarized in Table 2, reflecting high symmetry in all three groups.

Spatial features (i.e. PVG angles and left-right ratios) and temporal features (parameters marked by “_s”) were compared across three groups: children, male, and, female groups. Shapiro-Wilk tests showed that 9 parameters were Gaussian distributed (*LOA*, *LOP*, *ROA*, *ROA_s*, *RCA*, *RCA_s*, *RCP*, *Ratio_OP*, *Ratio_CP*) and 15 were Non-Gaussian (*LOA_s*, *LOP_s*, *LCA*, *LCA_s*, *LCP*, *LCP_s*, *ROP*, *ROP_s*, *RCP_s*, *Ratio_OP*, *Ratio_OAs*,

Ratio_OPs, Ratio_CA, Ratio_CAs, Ratio_CPs). For the 15 Non-Gaussian distributed parameters, a Kruskal-Wallis test was performed following a Bonferroni correction. Hence the significant level here was decreased to $p=0.05/15=0.0033$. The results of the performed statistics are given in Table 3. The last column in Table 3 shows the statistical results for ANOVA and Kruskal-Wallis tests. Within the previous three columns, the p-values for the post-hoc tests are given which were only performed if previous analysis showed statistical significance.

Female versus male group

Greater statistically significant gender differences between adult male and adult females were observed for temporal features (*RCA_s, LOA_s, LCA_s, LCP_s, RCP_s, Ratio_CPs*) compared to the spatial features (*LOP, RCA*).

The closing phase showed greater statistical differences in the temporal variability along the anterior and posterior parts of the vocal folds (*RCA_s, LCA_s, LCP_s, RCP_s, Ratio_CPs*) compared to the opening phase (*LOA_s*). Females demonstrated greater values for these temporal parameters, i.e. for females, the variability within cycle-to-cycle along the anterior and posterior parts of the vocal folds during the closing phase were higher compared to males. Spatial difference in the opening of the posterior part (*LOP*) and closing of the anterior part (*RCA*) was statistically significant between male and females. In females the posterior part begins to open earlier and the anterior part begins to close later compared to the medio-lateral part. This is exhibited by angles higher than 90° (Figure 3). In contrast, men exhibit more shutter like vibrations, exhibited by smaller angles closer to 90° (Figure 3). Variation of left-right symmetry for the closing phase (*Ratio_CPs*) is less in males compared to the females. Overall spatiotemporal symmetry parameters are rather equal between females and males (Table 2).

Children versus adult female group

Statistically significant differences were obtained for the spatial features of *LOA* and *RCA* along the anterior part. In both children and females, the anterior part opens earlier compared to the medial part. However, in children there is a greater delay between the anterior and medial parts of the vocal folds compared to the female group. This is also subjectively seen by the larger opening angles (*LOA, ROA, ROP*) in children (Figure 3). However, this was only found to be statistically significant for *LOA*. During the closing phase the anterior and posterior parts of the vocal fold close slightly later in females compared to the medial part. In children there is less difference in the closing of anterior/posterior and medial parts (values closer to 90°) (Figure 3). No statistical significant differences were determined in spatial symmetry characteristics

The four statistically significant temporal features (*RCA_s, LCP_s, Ratio_CAs, Ratio_CPs*) exhibit greater cycle to cycle variability in children during closing, but not during opening.

Children versus adult male group

Greater number of statistically significant differences were observed in both spatial and temporal features between children versus adult males (10 parameters) compared to adult females (6 parameters) (Table 3). Statistically significant differences were observed for the spatial features of *LOA, LOP*, and *RCA*. In both children and adult males the anterior / posterior parts of the vocal fold opens earlier than the medial part, however in children there is a greater delay between the anterior / posterior and medial parts of the vocal folds (Figure 3). In males, the closing especially in the anterior part is rather shutter like (angles close to 90°) and differs more along the anterior part than along the posterior part compared to children

Statistically significant differences were obtained for the following 7 temporal features: *RCA_s*, *LCA_s*, *LCP_s*, *RCP_s*, *LOA_s*, *Ratio_CPs* and *Ratio_CAs*, suggesting much greater cycle-to-cycle variability in children. The differences are especially pronounced for the closing phase. Also an increased cycle to cycle variability in symmetry during closing is exhibited in children due to statistical significant differences for *Ratio_CPs* and *Ratio_CAs*.

Classification

For classification purposes, only the 10 parameters that showed statistical significance within ANOVA or Kruskal-Wallis were considered (Table 3). The other parameters were omitted, since there were obviously no differences between the groups. Unsupervised cluster analysis was performed. The results are given in Table 4. In the last column, the *class recall* is given: percentage of the correct predicted subjects in the specific class. In the bottom row, the *class precision* values are given: percentage of the correct predicted subjects within all predicted subjects to the specific class.

The *class recall* is very high (<70%) for all three groups. No males were assigned towards the children group. Two females were assigned to the male and to the children group. Only two children were assigned to the males, but 7 children were assigned towards the female group. Regarding the *class precision*, it can be stated, that the predicted children group shows best values (92%); only 2 females were assigned towards the children and no males were assigned. The female *class precision* (60%) is worst, since several males and children were assigned to this class. The *class precision* of the male class is in between children and females. This indicates that the vocal fold vibrations within children are closer to vocal fold vibrations of females than to males and that female vocal fold vibrations are somehow in between children and male vocal fold vibrations.

DISCUSSION

In this work we present the first study on objective characterization of spatiotemporal PVG differences of vocal fold vibration on HSDI in typically developing children compared to adults. Unique functional differences in both spatial and temporal features of vibratory motion were evident between adult and pediatric vocal fold dynamics.

The overall differences in spatiotemporal characteristics were more distinct for adult males vs. children. To the best of our knowledge, this is the first paper providing empirical evidence from *in vivo* assessment of vocal fold vibrations of greater differences between adult males and children. Differences in both spatial and temporal aspects were observed between adult males and children; however differences between adult female and children were reflected predominantly in the temporal aspect of vocal fold vibrations. The spatial differences between adult male and children were predominantly significant for the closing phase of the anterior glottis (*LCA*, *RCA*).

Spatial characteristics

A novel finding of this study is that children had especially high angle values for opening along the posterior (*LOP*, *ROP*) glottis. This finding is attributed to the high occurrence (73%) of the posterior glottal gap during modal phonation in typically developing children, in the study.

Additionally, the opening angles in children are higher than the closing angles, indicating that the opening is longer compared to males and females. The finding of longer opening correlates with the finding of higher airflow open quotient produced by children during the prepubertal years. However the current finding provides direct physiological measurement of longer opening compared to the indirect measure of airflow open quotient which relates

to the duration of the opening.¹⁵ Children in this study had closing angles slightly closer to 90°, suggesting that the closing is more shutter-like with a decreased longitudinal delay between the anterior and the medio-lateral parts of the vocal folds compared to the opening phase. This difference between spatial opening and closing characteristics was also seen in adults (Figure 3), however less pronounced. Overall, angles for males were closer to 90° (i.e. shutter like vibrations) compared to females and children, who exhibited higher phase delays between anterior/posterior and medio-lateral parts, as seen by larger angles (Figure 3). Spatial differences between children and adults may be indicative of differential mechanisms involved in opening and closing of the vocal folds in children compared to the adults and could be attributed to the immature tissue characteristics.^{8–12}

The computed left-right mean symmetry ratios in adults as well as in children were always 1.0, representing “perfect” average symmetry between left and right spatial characteristics. However, the standard deviations in adults, for the left-right asymmetry were up to 20% confirming results for adults from earlier studies.^{44,45} The asymmetry ratio for children was also up to 20%, a novel finding of this study. The 20% asymmetry between the left and the right vocal fold is unlikely to be influenced by factors related to vocal fold deformation or endoscopic distortion,⁴⁵ as, similar factors were present for both children and adults. Hence the computed asymmetry of up to 20% could be attributed to differences in the histological make-up⁸ and of the central control of left and right vocal fold.⁴⁶ In spite of increased variability in spatiotemporal movement of vocal fold vibrations in children, the left-right symmetry ratio is similar to adults in children as young as 5 years of age; providing indirect evidence of adult-like cortical control of vocal fold motion.

Temporal characteristics

Differences in temporal features between children and adults were much more pronounced and statistically significant (Table 3) than spatial features. Children in this study demonstrated increased cycle-to-cycle temporal variability (Figure 4), providing physiological evidence of uncontrolled phonation at the level of the vocal folds. The reason for increased temporal variability in oscillatory dynamics in children is not known yet, but may be related to known anatomical differences towards the changing microstructure of the vocal fold layered structure,^{12,47} the size of the vocal folds,⁸ and the size of the vocal tract,⁹ in the prepubertal children. Increased temporal variability could also be due to the variability in amplitude, fundamental frequency,¹³ and aerodynamic function¹⁵ during the pre-pubertal years. However, this is speculative at this point. Further investigations with large number of participants, examining the growth trends across the various spatiotemporal vibratory features could establish the nature of the trend / profile of vocal development that could be useful for understanding the process of change responsible for the development of voice and voice disorders in children. It is interesting that children as young as 5 years of age try to maintain the shutter-like action of closing of adults, but in doing so have considerable cycle-to-cycle temporal variability especially during closing. This variability during closing is especially emphasized by the statistical differences for temporal characteristics based on symmetry (*Ratio_CAs*, *Ratio_CPs*) as well as angles (*LCA_s*, *LCP_s*, *RCA_s*, *RCP_s*).

Classification

Classification approaches in voice research have been mainly applied to differentiate pathologies from normal subjects.^{41,32,42} However, they were also applied to investigate if parameters are sufficient to differentiate between vocal fold dynamics like pre- and post vocal loading.⁴⁸ The goal of the here performed unsupervised clustering approach was to investigate if the considered set of parameters are sufficient to differentiate between children and adults. The clustering approach confirmed the characteristic dynamical behavioural properties within the selected parameters due to high correct classification (76% in average).

However, considering additional parameters (e.g. open quotient, close quotient, etc.) in future classification approaches, even better results could be achieved. The results of the clustering analysis also confirm that female vocal fold vibration characteristics are in between of those of children and males.

Limitations

Additional research for the application of the proposed features must be conducted with a larger number of subjects and multiple trials for development of reliable and standardized application of spatiotemporal measurements for clinical assessment of pediatric voice disorders. In future, several other aspects of oscillation like amplitude periodicity, open quotient, speed quotient, etc. will have to be analysed to determine the nature of pediatric vocal fold vibrations and the clinical utility of HSDI in assessment of pediatric voice disorders. In future it would be desirable to combine flexible fiberoptic with high speed in children as is currently done with stroboscopy.²⁸ Finally, future studies should explore this technique to investigate when puberty initiates and how the oscillatory dynamics change. This can be performed by PVG analysis and also by adapting biomechanical multi-mass models by comparing the optimized parameters like masses and stiffness.⁴⁵ For future classification approaches, it will be necessary to include more parameters and to analyse which of the parameters are best suited for distinguishing vocal fold vibrational characteristics for different groups, e.g. healthy versus certain pathologies.

CONCLUSION

This study provides the first quantitative evidence of functional differences in spatiotemporal vibratory characteristics in normal phonation between children and adults. Children have greater temporal cycle-to-cycle variability during normal phonation compared to adults, especially during closing. This study provides preliminary evidence that, in children the glottis opens in a zipper-like pattern and closes in a more shutter-like pattern without longitudinal delay. Like adults, children have similar left versus right asymmetry of up to 20%. Vocal fold oscillations of children are more similar to females than to males. The data presented here may provide the bases for differentiating normal vibratory characteristics from the disordered in the pediatric population. An improved understanding of the spatiotemporal properties of vocal fold vibrations could provide a foundation for development of empirically based age-specific clinical models for early identification of children with voice disorders, and assessment of treatment outcomes by providing clinical measurement of typical and disordered vocal function in children.

Acknowledgments

Financial Support: DFG (grant no. FOR894/2), NIH/NIDCD (grant no. R03DC011360-01)

Prof. Döllinger's contribution was made possible by Deutsche Forschungsgemeinschaft (DFG) grant no. FOR 894/2 "Strömungsphysikalische Grundlagen der Menschlichen Stimmgebung". Dr. Patel's contributions to this investigation were supported by NIH/NIDCD grant no. R03DC011360-01. We thank Dr. Richard Kryscio, Biostatistician at the University of Kentucky for his advice on the statistical analysis.

REFERENCES

1. Wilson, DK. Voice problems in children. 3rd. ed.. Baltimore: Williams and Wilkins; 1987.
2. McMurray S. Disorders of phonation in children. The Pediatrics Clinics of North America. 2003; 50:363–380.
3. Lass NJ, Ruscello DM, Bradshaw KH, Blankenship BL. Adolescents' perceptions of normal and voice-disordered children. J Commun Disord. 1991 Aug; 24(4):267–274. [PubMed: 1791216]

4. Connor NP, Cohen SB, Theis SM, Thibeault SL, Heatley DG, Bless DM. Attitudes of children with dysphonia. *J Voice*. 2008 Mar; 22(2):197–209. [PubMed: 17512168]
5. Ruddy BH, Sapienza CM. Treating voice disorders in the school-based setting: working within the framework of IDEA. *Language, Speech, and Hearing Services in Schools*. 2004 Oct; 35(4):327–332.
6. Hacki T, Heitmüller S. Development of the child's voice: premutation, mutation. *Int J Pediatr Otorhinolaryngol*. 1999 Oct 5; 49(Suppl 1):S141–S144. [PubMed: 10577793]
7. Vorperian H, Wang S, Chung MK, et al. Relational growth of oral and pharyngeal structures with vocal-tract length during the first two decades of life as visualized from MRI and CT studies: University of Wisconsin Madison. 2008
8. Hirano M, Kurita S, Nakashima T. Growth, Development, and Aging of Human Vocal Fold. In *Vocal Physiology contemporary research and clinical issues*, Bless D.M.; Abbs J.H. 1983:22–43.
9. Kahane JC. Growth of the human prepubertal and pubertal larynx. *J Speech Hear Res*. 1982 Sep; 25(3):446–455. [PubMed: 7176619]
10. Sato K, Hirano M, Nakashima T. Fine structure of the human newborn and infant vocal fold mucosae. *Ann Otol Rhinol Laryngol*. 2001 May; 110(5 Pt 1):417–424. [PubMed: 11372924]
11. Eckel HE, Koebke J, Sittel C, Sprinzl GM, Pototschnig C, Stennert E. Morphology of the human larynx during the first five years of life studied on whole organ serial sections. *Ann Otol Rhinol Laryngol*. 1999 Mar; 108(3):232–238. [PubMed: 10086614]
12. Hartnick CJ, Rehbar R, Prasad V. Development and maturation of the pediatric human vocal fold lamina propria. *Laryngoscope*. 2005 Jan; 115(1):4–15. [PubMed: 15630357]
13. Stathopoulos ET, Huber JE, Sussman JE. Changes in acoustic characteristics of the voice across the life span: measures from individuals 4–93 years of age. *J Speech Lang Hear Res*. 2011 Aug; 54(4):1011–1021. [PubMed: 21173391]
14. Hoit JD, Hixon TJ, Watson PJ, Morgan WJ. Speech breathing in children and adolescents. *J Speech Hear Res*. 1990 Mar; 33(1):51–69. [PubMed: 2314085]
15. Stathopoulos ET, Sapienza CM. Developmental changes in laryngeal and respiratory function with variations in sound pressure level. *J Speech Lang Hear Res*. 1997 Jun; 40(3):595–614. [PubMed: 9210117]
16. Stathopoulos ET, Sapienza C. Respiratory and laryngeal measures of children during vocal intensity variation. *J Acoust Soc Am*. 1993 Nov; 94(5):2531–2543. [PubMed: 8270731]
17. Sapienza CM, Stathopoulos ET. Respiratory and laryngeal measures of children and women with bilateral vocal fold nodules. *J Speech Hear Res*. 1994 Dec; 37(6):1229–1243. [PubMed: 7877283]
18. Sapienza CM, Stathopoulos ET. Comparison of maximum flow declination rate: children versus adults. *J Voice*. 1994 Sep; 8(3):240–247. [PubMed: 7987426]
19. Tang J, Stathopoulos ET. Vocal efficiency as a function of vocal intensity: a study of children, women, and men. *J Acoust Soc Am*. 1995 Mar; 97(3):1885–1892. [PubMed: 7699170]
20. Kent RD. Anatomical and neuromuscular maturation of the speech mechanism: evidence from acoustic studies. *J Speech Hear Res*. 1976 Sep; 19(3):421–447. [PubMed: 979206]
21. Glaze LE, Bless DM, Susser RD. Acoustic Analysis of Vowel and Loudness Differences in Children's Voice. *J Voice*. 1990; 4(1):37–44.
22. Sussman JE, Sapienza C. Articulatory, developmental, and gender effects on measures of fundamental frequency and jitter. *J Voice*. 1994 Jun; 8(2):145–156. [PubMed: 8061770]
23. 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 Feb; 128(2):156–160. [PubMed: 11843724]
24. Hasek CS, Singh S, Murry T. Acoustic attributes of preadolescent voices. *J Acoust Soc Am*. 1980 Nov; 68(5):1262–1265. [PubMed: 7440849]
25. Nicollas R, Garrel R, Ouaknine M, Giovanni A, Nazarian B, Triglia JM. Normal Voice in Children Between 6 and 12 Years of Age: Database and Nonlinear Analysis. *Journal of Voice*. 2008 Nov; 22(6):671–675. [PubMed: 17509823]
26. Robb MP, Simmons JO. Gender comparisons of children's vocal fold contact behavior. *The Journal of the Acoustical Society of America*. 1990 Sep; 88(3):1318–1322. [PubMed: 2229666]

27. Cheyne HA, Nuss RC, Hillman RE. Electroglottography in the pediatric population. *Arch Otolaryngol Head Neck Surg*. 1999 Oct; 125(10):1105–1108. [PubMed: 10522502]
28. Hartnick CJ, Zeitels SM. Pediatric video laryngo-stroboscopy. *Int J Pediatr Otorhinolaryngol*. 2005 Feb; 69(2):215–219. [PubMed: 15656955]
29. Wolf M, Primov-Fever A, Amir O, Jedwab D. The feasibility of rigid stroboscopy in children. *Int J Pediatr Otorhinolaryngol*. 2005 Aug; 69(8):1077–1079. [PubMed: 16005350]
30. Mortensen M, Schaberg M, Woo P. Diagnostic contributions of videolaryngostroboscopy in the pediatric population. *Archives of otolaryngology-- head & neck surgery*. 2010 Jan; 136(1):75–79. [PubMed: 20083783]
31. Zhang Y, Jiang JJ. Spatiotemporal chaos in excised larynx vibrations. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2005 Sep.72(3 Pt 2):035201. [PubMed: 16241503]
32. Inwald EC, Döllinger M, Schuster M, Eysholdt U, Bohr C. Multiparametric Analysis of Vocal Fold Vibrations in Healthy and Disordered Voices in High-Speed Imaging. *J Voice*. Aug 19.
33. Döllinger M. The next step in voice assessment: High-Speed digital endoscopy and objective evaluation. *Current Bioinformatics*. 2009; 4(2):101–111.
34. Patel R, Dailey S, Bless D. Comparison of high-speed digital imaging with stroboscopy for laryngeal imaging of glottal disorders. *Ann Otol Rhinol Laryngol*. 2008 Jun; 117(6):413–424. [PubMed: 18646437]
35. Lohscheller J, Toy H, Rosanowski F, Eysholdt U, Döllinger M. Clinically evaluated procedure for the reconstruction of vocal fold vibrations from endoscopic digital highspeed videos. *Medical image analysis*. 2007 Aug; 11(4):400–413. [PubMed: 17544839]
36. Lohscheller J, Eysholdt U. Phonovibrogram visualization of entire vocal fold dynamics. *Laryngoscope*. 2008 Apr; 118(4):753–758. [PubMed: 18216742]
37. Lohscheller J, Eysholdt U, Toy H, Döllinger M. Phonovibrography: mapping highspeed movies of vocal fold vibrations into 2-D diagrams for visualizing and analyzing the underlying laryngeal dynamics. *IEEE Trans Med Imaging*. 2008 Mar; 27(3):300–309. [PubMed: 18334426]
38. Kunduk M, Döllinger M, McWhorter AJ, Lohscheller J. Assessment of the variability of vocal fold dynamics within and between recordings with high-speed imaging and by phonovibrogram. *Laryngoscope*. 2010 May; 120(5):981–987. [PubMed: 20422695]
39. Döllinger M, Lohscheller J, McWhorter A, Kunduk M. Variability of normal vocal fold dynamics for different vocal loading in one healthy subject investigated by phonovibrograms. *J Voice*. 2009 Mar; 23(2):175–181. [PubMed: 18313896]
40. Havla J, Döllinger M, Eysholdt U, Lohscheller J. [Reproducibility and reliability of phonovibrograms. Quantification of healthy vocal fold vibrations]. *HNO*. 2009 Feb; 57(2):160–168. [PubMed: 19183918]
41. Voigt D, Döllinger M, Braunschweig T, Yang A, Eysholdt U, Lohscheller J. Classification of functional voice disorders based on phonovibrograms. *Artif Intell Med*. 2010 May; 49(1):51–59. [PubMed: 20138486]
42. Voigt D, Döllinger M, Yang A, Eysholdt U, Lohscheller J. Automatic diagnosis of vocal fold paresis by employing phonovibrogram features and machine learning methods. *Comput Methods Programs Biomed*. 2010 Sep; 99(3):275–288. [PubMed: 20138386]
43. Mehta DD, Deliyski DD, Quatieri TF, Hillman RE. Automated measurement of vocal fold vibratory asymmetry from high-speed videoendoscopy recordings. *J Speech Lang Hear Res*. 2011 Feb; 54(1):47–54. [PubMed: 20699347]
44. Bonilha HS, Deliyski DD, Gerlach TT. Phase asymmetries in normophonic speakers: visual judgments and objective findings. *Am J Speech Lang Pathol*. 2008 Nov; 17(4):367–376. [PubMed: 18840697]
45. Wurzbacher T, Schwarz R, Döllinger M, Hoppe U, Eysholdt U, Lohscheller J. Modelbased classification of nonstationary vocal fold vibrations. *J Acoust Soc Am*. 2006 Aug; 120(2):1012–1027. [PubMed: 16938988]
46. Ludlow, CL.; Sedory, SE.; Fujita, M., editors; Gauffin, J.; Hammarberg, B., editors. Neurophysiological control of vocal fold adduction and abduction for phonation onset and offset during speedh. San Diego: Singular Publishing Group Inc.; 1991. eds. Vocal fold physiology: acoustic, perceptual, and physiological aspects of voice mechanisms

47. Boseley ME, Hartnick CJ. Development of the human true vocal fold: depth of cell layers and quantifying cell types within the lamina propria. *Ann Otol Rhinol Laryngol*. 2006 Oct; 115(10): 784–788. [PubMed: 17076102]
48. Lohscheller J, Döllinger M, McWhorter AJ, Kunduk M. Preliminary study on the quantitative analysis of vocal loading effects on vocal fold dynamics using phonovibrograms. *The Annals of otology, rhinology, and laryngology*. 2008 Jul; 117(7):484–493.

Figure 1.

(A) Segmentation of the glottis of a single high speed frame. The right vocal fold edge is indicated red, the left is indicated blue. (B) The computation of the PVG is shown. (C) PVG image of three oscillation cycles. Superimposed in the last cycle, three corresponding HSDI images are shown.

Figure 2.

Generation of angles: PVG without angles, 50% deflections status along the vocal fold edges (left and right), regression lines through these, and the eight resulting angles *LOA* (*Left-Opening-Anterior*), *LCA*, *LOP*, *LCP*, *ROA*, *ROP*, *RCA*, *RCP*.

Figure 3. Mean values of the PVG angles (*LOA, ROA, LOP, ROP, LCA, RCA, LCP, RCP*) representing *spatial* features of opening (O), closing (C) dynamics of the left (L) and right (R) vocal fold along the anterior (A) and posterior (P) segments, across the three participant groups.

Figure 4. Standard deviations of the PVG angles (*LOA_s*, *ROA_s*, *LOP_s*, *ROP_s*, *LCA_s*, *RCA_s*, *LCP_s*, *RCP_s*) representing *temporal* features of opening (O), closing (C) dynamics of the left (L) and right (R) vocal fold along the anterior (A) and posterior (P) segments, across the three participant groups.

Table 1

Participant groups and age summary of mean and standard deviations in years.

	Males	Females	Children
Number	16	19	31
Mean $\pm$ SD (year)	29 $\pm$ 8	28 $\pm$ 7	8 $\pm$ 2

Table 2

Summary of the mean values and standard deviations of the spatiotemporal symmetry parameters. Characters “*O*” and “*C*” represent the opening and closing dynamics. Characters “*A*” and “*P*” represent the anterior and posterior portion. “_s” indicates the temporal parameters. All groups exhibit high left-right symmetric vibrations.

Mean $\pm$ SD	Female	Male	Children
Left – right symmetry ratios			
<i>Ratio_OA</i>	1.0 ± 0.2	1.0 ± 0.1	1.0 ± 0.1
<i>Ratio_OAs</i>	0.14 ± 0.09	0.13 ± 0.09	0.15 ± 0.07
<i>Ratio_OP</i>	1.0 ± 0.2	1.0 ± 0.2	1.0 ± 0.2
<i>Ratio_Ops</i>	0.13 ± 0.14	0.13 ± 0.08	0.11 ± 0.07
<i>Ratio_CA</i>	1.0 ± 0.1	1.0 ± 0.1	1.0 ± 0.1
<i>Ratio_CAs</i>	0.08 ± 0.03	0.08 ± 0.04	0.13 ± 0.05
<i>Ratio_CP</i>	1.0 ± 0.1	1.0 ± 0.1	1.0 ± 0.1
<i>Ratio_CPs</i>	0.10 ± 0.09	0.04 ± 0.01	0.12 ± 0.07

Table 3

Test for analysing statistical significance between the three groups of females, males, and children are performed by ANOVA (normal distributed data) and Kruskal-Wallis (not normal distributed data). Post-hoc tests are performed for the 10 parameters which showed significant vales for ANOVA with $p < 0.05$ or Kruskal-Wallis with $p < 0.0033$ due to Bonferroni correction, grey highlighted. Significant values within post-hoc tests are in bold: Eight parameters exhibit differences between females vs. males, 6 parameters show significant differences between females vs. children and 10 parameters show significant differences between males vs. children.

	Female (n = 19) vs. males (n = 16)	Female (n = 19) vs. children (n = 31)	Male (n = 16) vs. children (n = 31)	
Normal distributed data: Post hoc p-Values by LSD ($p < 0.05$)				ANOVA ($p < 0.05$)
<i>LOA</i>	0.961	0.039	0.045	0.049
<i>LOP</i>	0.001	0.657	0.001	0.001
<i>ROA</i>	-	-	-	0.062
<i>ROA_s</i>	-	-	-	0.681
<i>RCA</i>	0.000	0.002	0.004	0.000
<i>RCA_s</i>	0.008	0.018	0.000	0.000
<i>RCP</i>	-	-	-	0.425
<i>Ratio_OP</i>	-	-	-	0.296
<i>Ratio_CP</i>	-	-	-	0.193
Not-normal distributed data: Post-hoc p-Values by Mann-Whitney U-tests ($p < 0.05$)				Kruskal-Wallis ($p < 0.0033$)
<i>LOA_s</i>	0.004	0.074	0.000	0.000
<i>LOP_s</i>				0.022
<i>LCA</i>				0.008
<i>LCA_s</i>	0.013	0.091	0.000	0.000
<i>LCP</i>				0.936
<i>LCP_s</i>	0.003	0.006	0.000	0.000
<i>ROP</i>				0.022
<i>ROP_s</i>				0.781
<i>RCP_s</i>	0.008	0.059	0.000	0.000
<i>Ratio_OA</i>				0.783
<i>Ratio_OAs</i>				0.285
<i>Ratio_OPs</i>				0.535
<i>Ratio_CA</i>				0.045
<i>Ratio_CAs</i>	0.466	0.001	0.002	0.000
<i>Ratio_CPs</i>	0.000	0.005	0.000	0.000

Table 4

Mean values for the 10 times stratified cross validation of the unsupervised clustering. *Class recall* is the percentage of the correct predicted subjects. For the male group, this means that 81% (13 males out of 16 males) are correctly predicted to the male class. *Class precision* means the percentage of the correct predicted subjects within all predicted subjects to the specific class. For males, this means, that within the predicted class males (subjects: 13+2+2) actually 76% (13 subjects) were males. The overall mean accuracy is given with $76\% \pm 12\%$.

True Groups	Predicted Groups			Class recall
	Male	Female	Children	
Male	13	3	0	81%
Female	2	15	2	79%
Children	2	7	22	71%
Class Precision	76%	60%	92%	