


Original Investigation

Establishment of a Normative Cepstral Pediatric Acoustic Database

Scott A. Infusino, BA; Gillian R. Diercks, MD; Derek J. Rogers, MD; Jordan Garcia, BS; Shilpa Ojha, MBChB, MRCS, DO-HNS; Rie Maurer, MA; Glenn Bunting, MS, CCC-SLP; Christopher J. Hartnick, MD, MS

 Supplemental content at jamaotolaryngology.com

IMPORTANCE Few studies have used objective measures to evaluate the development of the normal pediatric voice. Cepstral analysis of continuous speech samples is a reliable method for gathering acoustic data; however, it has not been used to examine the changes that occur with voice development.

OBJECTIVE To establish and characterize acoustic patterns of the normal pediatric voice using cepstral analysis of voice samples from a normal pediatric voice database.

DESIGN, SETTING, AND PARTICIPANTS Cross-sectional study of 218 children aged 4 to 17 years, for whom English was the primary language spoken at home, conducted at a pediatric otolaryngology practice and pediatric practice in a tertiary hospital (April 2012–May 2014).

INTERVENTIONS AND EXPOSURES Sustained vowel utterances and continuous speech samples (4 Consensus Auditory-Perceptual Evaluation of Voice [CAPE-V] and 2 sentences from the rainbow passage) were recorded and analyzed from children with normal voices.

MAIN OUTCOMES AND MEASURES Normal values were collected for the acoustic measures studied (ie, fundamental frequency, cepstral peak fundamental frequency, cepstral peak prominence [CPP], low-to-high spectral ratio [L/H ratio], and cepstral-spectral index of dysphonia in recorded phrases) and compiled into a normative acoustic database.

RESULTS Significant changes in fundamental frequency were observed with a distinct shift in slope at ages 11 and 14 years in boys for sustained vowel (ages 4–11 years, -6.83 Hz/y [$P < .001$]; 11–14 years, -27.62 Hz/y [$P < .001$]; and 14–17 years, -5.68 Hz/y [$P = .001$]), all voiced (ages 4–11 years, -4.19 Hz/y [$P = .002$]; 11–14 years, -29.42 Hz/y [$P < .001$]; and 14–17 years, -4.63 Hz/y [$P < .001$]), glottal attack (ages 4–11 years, -4.51 Hz/y; 11–14 years, -27.23 Hz/y; and 14–17 years, -1.70 Hz/y [$P < .001$ for all]), and rainbow (ages <14 years, -20.68 Hz/y [$P < .001$]; and 14–17 years, -4.50 Hz/y [$P = .001$]) recordings. A decreasing linear trend in fundamental frequency among all recordings (vowel, all voiced, easy onset, glottal attack, plosives, and rainbow) was found in girls (-2.56 Hz/y [$P < .001$], -3.48 Hz/y [$P < .001$], -2.82 Hz/y [$P < .001$], -3.49 Hz/y [$P < .001$], -2.30 Hz/y [$P < .001$], and -2.98 Hz/y [$P = .01$], respectively). A linear increase in CPP was seen with age in boys, with significant changes seen in recordings for vowel (0.10 dB/y [$P = .05$]), all voiced (0.2 dB/y [$P < .001$]), easy onset (0.13 dB/y [$P < .001$]), glottal attack (0.12 dB/y [$P < .001$]), plosives (0.15 dB/y [$P < .001$]), and rainbow (0.17 dB/y [$P = .006$]). A significant linear increase in CPP for girls was only seen in all voiced (0.13 dB/y [$P < .001$]). L/H ratio showed a linear increase with age among all speech samples (vowel, all voiced, easy onset, glottal attack, plosives, and rainbow) in boys (1.14 dB/y [$P < .001$], 0.92 dB/y [$P < .001$], 1.19 dB/y [$P < .001$], 0.79 dB/y [$P < .001$], 0.69 dB/y [$P < .001$], and 0.54 dB/y [$P = .002$], respectively) and girls (0.96 dB/y, 0.60 dB/y, 0.75 dB/y, 0.37 dB/y, 0.44 dB/y, and 0.58 dB/y, respectively [$P \leq .001$ for all]).

CONCLUSIONS AND RELEVANCE This represents the first pediatric voice database using frequency-based acoustic measures. Our goal was to characterize the changes that occur in both male and female voices as children age. These findings help illustrate how acoustic measurements change with development and may aid in our understanding of the developing voice, pathologic changes, and response to treatment.

JAMA Otolaryngol Head Neck Surg. 2015;141(4):358–363. doi:10.1001/jamaoto.2014.3545
Published online January 22, 2015.

Author Affiliations: Division of Pediatric Otolaryngology, Massachusetts Eye and Ear Infirmary, Boston (Infusino, Diercks, Rogers, Garcia, Hartnick); Charing Cross Hospital, London, England (Ojha); Center for Clinical Investigation, Brigham and Women's Hospital, Boston, Massachusetts (Maurer); Voice and Speech Laboratory, Massachusetts Eye and Ear Infirmary, Boston (Bunting).

Corresponding Author: Scott A. Infusino, BA, Department of Otolaryngology, Massachusetts Eye and Ear Infirmary, 243 Charles St, Boston, MA 02114 (Sainfusino@gmail.com).

Spoken language is our primary form of communication that facilitates our social development, education, and interpersonal and professional interactions. Childhood and adolescence represent essential periods of development for both vocal quality and speech. Pediatric voice disorders can have a profound impact on a child's social and intellectual development, affecting up to 10% of the population,^{1,2} thus highlighting the importance of better understanding the normal process of vocal development. However, objective data characterizing this progression are lacking. Without an understanding of how acoustic measurements of the normal pediatric voice develop and change over time, our ability to use these objective measurements in evaluating children with voice disorders and their response to treatments is limited.

Perceptual analysis performed by trained speech pathologists and clinicians has traditionally been used to study characteristics of the normal voice and to identify voice disorders. A relatively modern assessment created to judge voice quality and the severity of disorders is the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V).³ To perform a CAPE-V analysis, the patient reads the following 4 sentences:

- Easy onset, "How hard did he hit him?"
- All voiced, "We were away a year ago."
- Glottal attack, "We eat eggs every Easter."
- Plosives, "Peter will keep at the peak."

Despite attempts at objectivity, CAPE-V relies on the trained observer's perceptions and inherently has some degree of subjectivity when used for perceptual analysis. Quality-of-life instruments, such as validated questionnaires, have also been used in adults and children to identify "normal" and "abnormal" voice characteristics and how they affect function.⁴⁻⁶

Unlike analyses performed by a trained observer, acoustic measurements calculate vocal properties from algorithms that use period and frequency and can provide objective information about the qualities of the normal and dysphonic voice. Time-based algorithms performed on sustained vowel utterances provide information about cycle-to-cycle variations in frequency and amplitude, which help objectively characterize subjective perceptions of breathiness, roughness, and strain.⁴

In 2012, Maturo et al⁷ published the first large series of time-based acoustic measurements from normative voice recordings in children. Three hundred thirty-five boys and girls ranging in age from 4 to 18 years were recorded speaking sustained vowel utterances, and then fundamental frequency (F_0), jitter, shimmer percentage, and noise-to-harmonic ratio were calculated. Their findings did not reveal age-based differences in acoustic properties such as jitter and shimmer but did show differences across age for noise-to-harmonic ratio. They also found certain transitional periods of development, marked by critical age points, where significant changes in F_0 began.⁷ This finding gives credibility to the significance of a vocal fold with a developing layered structure rather than a simple cord where function changes in linear relation to cord length.⁸

One potential limitation of the work by Maturo et al⁷ is that acoustic measurements were calculated using 1 voice sample and that sustained vowel utterances obtained from children

at 1 point in time might lack reliability. To address this issue of consistency, Hill et al⁹ obtained sustained vowel utterances from 50 boys and girls aged 4 to 17 years at 2 narrowly spaced points in time. Time-based acoustic measurements between sustained vowel utterances were found to show consistency only in F_0 . Poor reliability was seen with jitter, shimmer, and noise-to-harmonic ratio.⁹ This finding suggests that sustained vowel utterances should not be the only voice sample used to obtain acoustic measurements in children.

Cepstral analysis uses frequency-based algorithms to examine continuous speech samples. These algorithms are used to create a cepstrum, the broken-down components of a speech sample represented as a distribution of pitches and their corresponding intensities (measured in decibels). A great deal of information can be uncovered from a cepstrum, such as frequency and variability, much like time-based acoustic analyses. Unlike sustained vowel utterances, however, cepstral analyses of continuous speech samples may provide a more representative assessment of a patient's voice.¹⁰⁻¹⁵ In 2013, Diercks et al¹⁶ demonstrated that standardized continuous voice recordings obtained from normal children at 2 points in time were consistent for measures of fundamental frequency and variability.

Given the advantages of frequency-based acoustic measurements on standardized continuous speech samples, the goal of this research was to establish the first cepstral analysis database in children aged 4 to 17 years. With this database, we hope to provide a more accurate representation of the developing pediatric voice compared with prior studies. Without first characterizing the normal pediatric voice, acoustic measurements cannot be used to quantify degrees of dysphonia and treatment response in children with vocal dysfunction.

Methods

This study was carried out between April 2012 and May 2014 with the approval of the institutional review board of the Massachusetts Eye and Ear infirmary. Patient participation in this study was dependent on the agreement of both the parent and child, with written informed consent obtained. Patients between the ages of 4 to 17 years were recruited in an outpatient pediatric otolaryngology clinic as well as a general pediatric clinic. Patients with a hearing loss greater than 20 dB, history of airway procedures or voice disorders, developmental delay, or cognitive delay were excluded from the study. All patients used English as their primary language spoken at home.

Children were seated in a quiet room with an adjustable microphone (Shure Beta 53; Shure Inc) placed to the right oral commissure. Recordings were collected with a Dell Optiplex 960 personal computer (Dell Inc) with an Intel Core Duo 2 CPU (3.1 GHz, 1.94 GB of RAM) (Microsoft Windows XP Professional Version 2002; Microsoft Corp). Multidimensional Voice Program (model 5101) and Analysis of Dysphonia in Speech and Voice (model 5109) for the Computerized Speech Laboratory (model 4500) (KayPENTAX) were used to analyze voice recordings. The MDVP was used to evaluate F_0 for sustained

Table 1. Definitions of Measures Calculated by ADSV Software

Statistic (Unit of Measure)	Description
CP (dB)	A high-intensity amplitude in the cepstrum that corresponds to the dominant F_0 . ³
CPP (dB)	Measures the difference between F_0 intensity and the average intensity of the signal. A higher CPP should be seen with normal voices. ¹⁶
L/H ratio (dB) ^a	The mean ratio of signal energy below 4000 Hz to the energy above 4000 Hz for the selected voice data frames. ¹⁷
Mean CPP F_0 (Hz) ^a	The mean frequency of the cepstral peaks corresponding to the range 60 Hz to 300 Hz for the selected voiced data frames. ¹⁷
CSID ^a	A multivariate calculation that gives an estimation of vocal abnormality. Higher scores correlate with dysphonia. ¹⁷

Abbreviations: ADSV, Analysis of Dysphonia in Speech and Voice (KayPENTAX); CP, cepstral peak; CPP, cepstral peak prominence; CPP F_0 , cepstral peak fundamental frequency; CSID, cepstral-spectral index of dysphonia; F_0 , fundamental frequency; L/H ratio, low-to-high spectral ratio.

^a L/H ratio, mean CPP F_0 , and CSID were taken from the KayPENTAX ADSV software manual.¹⁷

vowel utterances. The ADSV was used to evaluate cepstral peak fundamental frequency (CPP F_0), cepstral peak prominence (CPP), CPP standard deviation (SD), low-to-high spectral ratio (L/H ratio), L/H ratio SD, and cepstral-spectral index of dysphonia of sustained vowel utterances (CSID) in CAPE-V sentences and the second and third sentences of the rainbow passage (“The rainbow is the division of white light into many beautiful colors. These take the shape of a long round arch with its path high above and its 2 ends apparently beyond the horizon”). The definitions of these values are given in Table 1.

When recording sustained vowel utterances, children were asked to maintain the vowel /a/ at a comfortable pitch and loudness level for 4 seconds. For CAPE-V and rainbow sentence recordings, children were asked to use their normal speaking voice. The rainbow sentences could not be reliably spoken by younger participants and only were included for children aged 10 to 17 years. Samples from 220 children (113 male and 107 female) were recorded.

We used descriptive statistics, obtaining mean (SD), median, and range, for both MDVP and ADSV data across age. Data were summarized as both sexes combined and stratified by sex. For each set of data, means with error bars (SDs) were initially plotted for visual inspections of slope for both sexes combined and stratified by sex. Locally weighted scatterplot smoothing (LOESS) was then applied to sex-stratified data to determine if the values changed linearly over time or if it showed break points of change in slope. If the values changed constantly over time, a simple linear regression was fitted. If LOESS fits showed any break points of change in slope, then piecewise regression was performed.

Results

A total of 220 children (113 male and 107 female) participated in this study (Table 2). All participants had normal voices that were either self-reported or confirmed by their

parents and were screened for any disqualifying factors prior to recording.

Piecewise regressions showed distinct periods of change, marked by critical points, in male frequency measures across age, with both the MDVP and ADSV data (Figure). The rate of change of the MDVP and ADSV data for each distinct period, marked by critical points, is given in Table 3. In the male population, MDVP F_0 and ADSV CPP F_0 from sustained vowel, all voiced, and glottal attack all showed statistically significant critical points at ages 11 and 14 years. These transitional periods were not observed in frequency data from our female participants. Instead, a more linear relationship was seen with the mean frequency significantly decreasing across age (Table 4 and Figure).

For CPP in the male population, linear regression showed a significant positive slope for all voiced (approximate change per increasing year, 0.2 dB; $P < .001$), easy onset (0.13 dB; $P < .001$), glottal attack (0.12 dB; $P < .001$), plosives (0.15 dB; $P < .001$), and the rainbow sentence (0.17 dB; $P < .01$). Only all voiced had a linear regression that showed significant positive slope in CPP for girls (approximate change per increasing year, 0.13 dB; $P < .001$). Low-to-high spectral ratio showed a small, yet significant linear increase in all ADSV recordings against age for both sexes (Table 4).

The measured ADSV variables for each sex and age group are given in eTables 1 and 2 in the Supplement.

Discussion

To our knowledge, this is the largest study evaluating healthy voice development in children using cepstral analysis. Prior databases have been collected using time-based acoustic measures; however, sustained vowel utterances do not provide reliable data regarding variability in acoustic signal. Continuous speech samples can be evaluated objectively through cepstral analysis and provide a more representative speech sample for analysis.

Fundamental frequency—the rate of vocal fold vibration—is a common measurement in voice analyses. In 2012, Maturo et al⁷ published the largest database examining normal pediatric voice development using time-based algorithms. This work demonstrated specific periods of transition in F_0 during development in both male and female participants. In male participants, F_0 began to shift at age 12 years and reached maturation at age 16 years. In female participants, F_0 began to transition at age 11 years and reached maturation at age 14 years.⁷

While our results from both time-based and cepstral-based algorithms support the findings of Maturo et al⁷ of 2 transition periods for F_0 in male participants, our transitional periods were marked by critical points primarily at 11 and 14 years. In addition, our fundamental frequency data for female participants showed a decreasing linear correlation with age with no critical points. This discrepancy is surprising because both studies were performed primarily in the same clinic at the Massachusetts Eye and Ear Infirmary with very similar sample populations. A possible explanation for these disparities could stem from the different researchers collecting the voice data

Table 2. Distribution of Recorded Patients

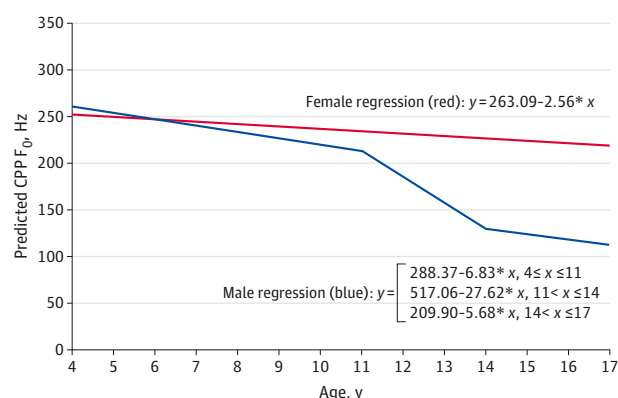
Participants	Age, y													
	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Male, No.	5	8	10	10	10	8	10	9	4	10	10	7	4	8
Female, No.	5	4	10	10	7	9	8	9	8	7	10	7	8	5

or the larger sample size used in the study by Maturo et al.⁷ While further work should be done to solve this inconsistency, it is important to note that our results also include additional ADSV data. The transitional periods in F_0 we found in male participants aged 11 to 14 years were seen with MDVP data as well as 4 of 6 ADSV CPP F_0 measures (vowel, all voiced, glottal attack, and rainbow). In female participants, no transition periods in F_0 were seen with the MDVP data or any of the 6 additional ADSV frequency measures we collected. It is possible that with our female population, critical points in F_0 and CPP F_0 were too subtle to detect and a larger sample size is needed.

In 2013, Diercks et al¹⁶ demonstrated that cepstral analysis of continuous speech recordings offers a more reliable method for measuring voice perturbation and variability. Frequency-based acoustic measurements of signal variability, including CPP, L/H ratio, L/H ratio SD, and CSID score (multivariate index measure) are all reliable calculations for measuring the pediatric voice.¹⁶ Cepstral peak prominence is the difference between F_0 intensity and the average intensity of the voice sample.⁹ Previous literature has reported that a low CPP correlates strongly with perceived ratings of dysphonia and breathiness.^{12,14,18,19} Our data showed a small but significant increasing trend in CPP in male participants (Table 4). One possible explanation is that as age increases, phonation is better controlled during speech, which leads to a higher CPP. In female participants, a significant increasing trend was only seen with all voiced (Table 4). Larger sample sizes in future studies may better resolve potential trends in vowel, easy onset, glottal, plosives, and rainbow with age. L/H ratio is a measure of the distribution of low to high frequency energy in a voice sample, and studies have associated low values of this statistic with dysphonia.^{3,16} We found a significant linear increase in L/H ratio in both sexes for all phrases. As F_0 decreases, the addition of lower frequencies acquired by an individual could potentially raise L/H values.

We initially believed that we could compare existing normative CSID data to children's values in our study to exclude any candidates with voice abnormalities. Diercks et al¹⁶ found that the CSID score generated by the ADSV software was a reliable measure when examined across 2 points in time. Publications examining CSID have shown the ability to distinguish between normal and dysphonic voices and that the score is consistent with perceptual ratings from speech pathologists.^{20,21} However, the CSID scoring system for the ADSV software was developed based on adult data, and normative values have only been researched and published in adult populations. Our work suggests that the values calculated from mature voices are not representative of the pediatric population. While all of the children included in the study had normal voices, younger children tended to display higher CSID values that would otherwise be labeled as dysphonic in adults. Our pediatric CSID values

Figure. Piecewise Linear Regression Across Age Fit to Male and Female ADSV Data, Sustained Vowel Utterances



ADSV indicates Analysis of Dysphonia in Speech and Voice (KayPENTAX); and CPP F_0 , cepstral peak fundamental frequency.

for easy onset, all voiced, glottal attack, and plosives all decreased with age, eventually falling within the normative range seen in adult populations. For both sexes, CSID values for sustained vowel were still high by age 17 years. However, this is expected because sustained vowels have a higher variability than continuous speech samples.¹⁶ Recordings from the rainbow passage remained relatively constant, likely since they were used for ages 10 years and older. To define expected CSID scores for children across age and sex, further work needs to be done to characterize the normal ranges and values for each age group.

One potential limitation of this study was that patients were recruited in a tertiary care pediatric otolaryngology practice and general pediatric practice and may not accurately represent the general pediatric population. We did, however, attempt to include only children with normal voices through appropriate participant screening. In addition, we did not screen for external factors, such as sleep deprivation, caffeine intake, and hydration, which can also affect accuracy of acoustic measurements. Some age groups in this study are not represented equally, and because the patient population is stratified by sex and age, each group has a relatively small sample size. Therefore, some trends in the data, such as clear transitional periods in female F_0 development, may emerge given a larger sample size. While Diercks et al¹⁶ were able to show that continuous recordings of children aged 4 to 18 years were reliable, they did not look at each individual age group. It is possible that some of the age groups used in this study cannot reliably produce continuous speech.

A great deal of work has been done examining the acoustic properties of the human voice in both normal and dysphonic populations. While there is good literature on adult normative

Table 3. Transitional Periods in Fundamental Frequency for Male Participants, With MDVP and ADSV Measures

Statistic	Software	Critical Points	Ages	Slopes	P Value
Sustained vowel, F ₀	MDVP	11 and 14	4 to 11	-6.93	<.001
			11 to 14	-31.88	<.001
			14 to 17	-5.87	<.001
Sustained vowel, CPP F ₀	ADSV	11 and 14	4 to 11	-6.83	<.001
			11 to 14	-27.62	<.001
			14 to 17	-5.68	.001
All voiced, CPP F ₀	ADSV	11 and 14	4 to 11	-4.19	.002
			11 to 14	-29.42	<.001
			14 to 17	-4.63	<.001
Easy onset, CPP F ₀	ADSV	10 and 13	4 to 10	-2.63	.08
			10 to 13	-22.78	<.001
			13 to 17	-7.30	<.001
Glottal attack, CPP F ₀	ADSV	11 and 14	4 to 11	-4.51	<.001
			11 to 14	-27.23	<.001
			14 to 17	-1.70	<.001
Plosives, CPP F ₀	ADSV	11 and 13	4 to 11	-3.45	.01
			11 to 13	-35.26	<.001
			13 to 17	-3.35	<.001
Rainbow, CPP F ₀	ADSV	14	10 to 14	-20.68	<.001
			14 to 17	-4.50	.001

Abbreviations: ADSV, Analysis of Dysphonia in Speech and Voice (KayPENTAX); CPP F₀, cepstral peak fundamental frequency; MDVP, Multidimensional Voice Program (KayPENTAX).

Table 4. Change in Fundamental Frequency (MDVP and ADSV), L/H Ratio, and CPP for Male and Female Participants

Recorded Phrase	Measure	Male		Female	
		Approximate Change per Increasing Year	P Value	Approximate Change per Increasing Year	P Value
Sustained vowel	MDVP F ₀ , Hz	...		-3.83	<.001
	ADSV CPP F ₀ , Hz	...		-2.56	<.001
	CPP, dB	0.10	.05	0.05	.38
	L/H ratio, dB	1.14	<.001	0.96	<.001
All voiced	ADSV CPP F ₀ , Hz	...		-3.48	<.001
	CPP, dB	0.2	<.001	0.13	<.001
	L/H ratio, dB	0.92	<.001	0.60	<.001
Easy onset	ADSV CPP F ₀ , Hz	...		-2.82	<.001
	CPP, dB	0.13	<.001	0.01	.73
	L/H ratio, dB	1.19	<.001	0.75	<.001
Glottal attack	ADSV CPP F ₀ , Hz	...		-3.49	<.001
	CPP, dB	0.12	<.001	0.03	.20
	L/H ratio, dB	0.79	<.001	0.37	<.001
Plosives	ADSV CPP F ₀ , Hz	...		-2.30	<.001
	CPP, dB	0.15	<.001	0.04	.08
	L/H ratio, dB	0.69	<.001	0.44	<.001
Rainbow	ADSV CPP F ₀ , Hz	...		-2.98	.01
	CPP, dB	0.17	.006	0.06	.16
	L/H ratio, dB	0.54	.002	0.58	.001

Abbreviations: ADSV, Analysis of Dysphonia in Speech and Voice (KayPENTAX); CPP, cepstral peak prominence; CPP F₀, cepstral peak fundamental frequency; F₀, fundamental frequency; L/H ratio, low-to-high spectral ratio; MDVP, Multidimensional Voice Program (KayPENTAX); ellipses, data are reported in Table 3.

acoustic values, it is not applicable to pediatric populations.⁴ Several studies have used acoustic measures to examine dysphonia in children,^{11,22-26} yet there is currently no database that looks at these measures in normal developing voices. To better understand dysphonic voices in children and the developing voice in general, it is important to properly characterize the nor-

mal pediatric voice. Future work with cepstral analysis needs to be performed using larger sample sizes to more accurately define the acoustic properties of the pediatric voice.

Although we have shown both distinct and subtle changes of the developing voice using cepstral analysis, the underlying causes of these changes have yet to be identified. Histo-

logical work has shown that there is a clear change in vocal fold architecture from birth to adulthood.⁸ Other factors such as vocal fold length and hormonal influences are also linked to acoustic development.^{7,27} It is likely a combination of several of these characteristics that lead to the mature, adult voice.

Conclusions

To our knowledge, this is the first study to look at normative pediatric voice development using cepstral-based algorithms to analyze continuous speech. With our voice recordings, we have been able to map the development of the pediatric voice for both male and female populations. Our data corroborates

previous findings that male voice development has a transitional period marked by 2 critical points in frequency change. While many studies have examined acoustic measurements of the abnormal voice based on time- and frequency-based algorithms, more work is needed to define the acoustic profiles of the normal pediatric voice. Voice maturation is a complex process, which involves changes in vocal fold length and tissue architecture. Correlating data about the acoustic profile of normal children with changes in vocal fold structure during development may help us to understand more about how the voice matures. Moving forward, we hope these data can facilitate additional acoustic voice investigations and eventually aid in diagnosing voice disorders and measuring response to treatment.

ARTICLE INFORMATION

Submitted for Publication: August 5, 2014; final revision received October 15, 2014; accepted December 2, 2014.

Published Online: January 22, 2015.
doi:10.1001/jamaoto.2014.3545.

Author Contributions: Mr Infusino had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Infusino, Diercks, Bunting, Hartnick.

Acquisition, analysis, or interpretation of data: Infusino, Diercks, Rogers, Garcia, Ojha, Maurer, Bunting.

Drafting of the manuscript: Infusino, Garcia, Ojha.
Critical revision of the manuscript for important intellectual content: Infusino, Diercks, Rogers, Garcia, Maurer, Bunting, Hartnick.

Statistical analysis: Maurer.

Administrative, technical, or material support: Infusino, Rogers, Ojha, Bunting, Hartnick.

Study supervision: Diercks, Rogers, Bunting, Hartnick.

Conflict of Interest Disclosures: None reported.

REFERENCES

- Baker BM, Blackwell PB. Identification and remediation of pediatric fluency and voice disorders. *J Pediatr Health Care*. 2004;18(2):87-94.
- Carding PN, Roulstone S, Northstone K; ALSPAC Study Team. The prevalence of childhood dysphonia: a cross-sectional study. *J Voice*. 2006;20(4):623-630.
- Awan SN. *Analysis of Dysphonia in Speech and Voice (ADSV): An Application Guide*. Montvale, NJ: KayPENTAX; 2011.
- Baken RJ, Orlikoff R. *Clinical Measurement of Speech and Voice*. 2nd ed. San Diego, CA: Singular; 2000.
- Boseley ME, Cunningham MJ, Volk MS, Hartnick CJ. Validation of the Pediatric Voice-Related Quality-of-Life survey. *Arch Otolaryngol Head Neck Surg*. 2006;132(7):717-720.
- Merati AL, Keppel K, Braun NM, Blumin JH, Kerschner JE. Pediatric voice-related quality of life: findings in healthy children and in common laryngeal disorders. *Ann Otol Rhinol Laryngol*. 2008;117(4):259-262.
- Maturo S, Hill C, Bunting G, Ballif C, Maurer R, Hartnick C. Establishment of a normative pediatric acoustic database. *Arch Otolaryngol Head Neck Surg*. 2012;138(10):956-961.
- Hartnick CJ, Rehbar R, Prasad V. Development and maturation of the pediatric human vocal fold lamina propria. *Laryngoscope*. 2005;115(1):4-15.
- Hill CA, Ojha S, Maturo S, Maurer R, Bunting G, Hartnick CJ. Consistency of voice frequency and perturbation measures in children. *Otolaryngol Head Neck Surg*. 2013;148(4):637-641.
- Awan SN, Roy N. Outcomes measurement in voice disorders: application of an acoustic index of dysphonia severity. *J Speech Lang Hear Res*. 2009;52(2):482-499.
- Halberstam B. Acoustic and perceptual parameters relating to connected speech are more reliable measures of hoarseness than parameters relating to sustained vowels. *ORL J Otorhinolaryngol Relat Spec*. 2004;66(2):70-73.
- Heman-Ackah YD, Heuer RJ, Michael DD, et al. Cepstral peak prominence: a more reliable measure of dysphonia. *Ann Otol Rhinol Laryngol*. 2003;112(4):324-333.
- Lafren JB, Lazarus CL, Amin MR. Pitch deviation analysis of pathological voice in connected speech. *Ann Otol Rhinol Laryngol*. 2008;117(2):90-97.
- Maryn Y, Corthals P, Van Cauwenberge P, Roy N, De Bodt M. Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels. *J Voice*. 2010;24(5):540-555.
- Watts CR, Awan SN. Use of spectral/cepstral analyses for differentiating normal from hypofunctional voices in sustained vowel and continuous speech contexts. *J Speech Lang Hear Res*. 2011;54(6):1525-1537.
- Diercks GR, Ojha S, Infusino S, Maurer R, Hartnick CJ. Consistency of voice frequency and perturbation measures in children using cepstral analyses: a movement toward increased recording stability. *JAMA Otolaryngol Head Neck Surg*. 2013;139(8):811-816.
- KayPENTAX. *Analysis of Dysphonia in Speech and Voice Model 5109: Software Instruction Manual*. Vol Issue A. Montvale, NJ: KayPENTAX; 2011.
- Hillenbrand J, Cleveland RA, Erickson RL. Acoustic correlates of breathy vocal quality. *J Speech Hear Res*. 1994;37(4):769-778.
- Wolfe VI, Martin DP, Palmer CI. Perception of dysphonic voice quality by naïve listeners. *J Speech Lang Hear Res*. 2000;43(3):697-705.
- Awan SN, Roy N, Jetté ME, Meltzner GS, Hillman RE. Quantifying dysphonia severity using a spectral/cepstral-based acoustic index: comparisons with auditory-perceptual judgements from the CAPE-V. *Clin Linguist Phon*. 2010;24(9):742-758.
- Lowell SY, Colton RH, Kelley RT, Hahn YC. Spectral- and cepstral-based measures during continuous speech: capacity to distinguish dysphonia and consistency within a speaker. *J Voice*. 2011;25(5):e223-e232.
- Dejonckere PH, Lebacqz J. Acoustic, perceptual, aerodynamic and anatomical correlations in voice pathology. *ORL J Otorhinolaryngol Relat Spec*. 1996;58(6):326-332.
- Gramuglia AC, Tavares EL, Rodrigues SA, Martins RH. Perceptual and acoustic parameters of vocal nodules in children. *Int J Pediatr Otorhinolaryngol*. 2014;78(2):312-316.
- Lopes LW, Barbosa Lima IL, Alves Almeida LN, Cavalcante DP, de Almeida AA. Severity of voice disorders in children: correlations between perceptual and acoustic data. *J Voice*. 2012;26(6):819.e7-819.e12.
- Reynolds V, Buckland A, Bailey J, et al. Objective assessment of pediatric voice disorders with the acoustic voice quality index. *J Voice*. 2012;26(5):672.e1-672.e7.
- Simões-Zenari M, Nemr K, Behlau M. Voice disorders in children and its relationship with auditory, acoustic and vocal behavior parameters. *Int J Pediatr Otorhinolaryngol*. 2012;76(6):896-900.
- Harries ML, Walker JM, Williams DM, Hawkins S, Hughes IA. Changes in the male voice at puberty. *Arch Dis Child*. 1997;77(5):445-447.