Global Journal of
Otolaryngology
ISSN 2474-7556

Juniper Publishers
key to the Researchers

Research Article
Volume 11 Issue 1 – October 2017
DOI: 10.19080/GJO.2017.11.555804

Glob J Otolaryngol

Cepstral Analysis of Voice in Young Children and Adults

Usha Manjunatha*

Kasturba Medical College, India

Submission: September 26, 2017; **Published:** October 12, 2017

***Corresponding author:** Usha Manjunatha, Kasturba Medical College, Attavar, India, Tel: 9611840055; Email: usha_yeshu@yahoo.com

Introduction

Voice is a multidimensional entity and it reveals the speaker's physical and emotional health, personality and identity. Voice production is an aerodynamic process in which the acoustical waves are created by laryngeal modulations of respiratory airflow. These waves are amplified and filtered accordingly by vocal tract resonance. For an economical or optimum vocal output, stability across these respiratory, phonatory, and resonatory subsystems is essential. The voice changes dynamically, minute by minute. But there are long term changes that are associated with growth and decline in life. At the major stages of life, the uses of the voice are different, as are the demands placed upon it. The reasons for those differences are many and include biological maturation and the emotional and social changes that occur in the individual's life. A voice change due to biological maturations is the main concern, as we age; there will be changes in anatomical structures related voice production and physiology of the same. It varies from person to person and across gender. But the trend seen is similar in both the gender.

The voice of the children is similar and when they become older there will be voice differentiation in males and females. At infancy vocal fold length is 6-8mm and at puberty 12- 15mm and in adults 12.5-17mm in females and 17-23mm in males [1]. Other anatomical differences in terms of size and dimension of the larynx attributes to it. The presence of sex hormones estrogens and progesterone in girls and androgens in boys triggers the development of the third layer of the epithelium cells of the vocal folds. Prior to this human larynx have only two layers in the vocal folds. The consequences of puberty reflect on voice and are more obvious in boys than in girls, but they exist in both gender. The fundamental frequency and resonant characteristics also varies accordingly, females tend to have high pitch when they reach puberty and males tend to have low pitch voice. The vocal tract length increases at puberty and there will be additional harmonics. Children will have less number of harmonics, as the age increases, pubertal change takes place; there will be addition of higher harmonics in females and lower harmonics in males. Based on biological maturation, the voice quality changes as the age progresses. The male voice goes through changes usually between the ages of 12 to 15 years and there will be reduction in the vocal pitch anywhere from 1 to 2.5 octaves and the drop is only by a third of an octave in females.

To assess the changes in voice quality and to compare with normal aspects and to determine which is pathological, there are various measures to evaluate voice quality. These measures include subjective and objective measures. Acoustic analysis is one of the objective measures which are easy to administer, non invasive and less time consuming when compared to laryngeal imaging procedures. These methods also help in monitoring treatment outcomes. Acoustic analysis provides information regarding the stability or variability of the vocal fold movement through perturbation measures of amplitude and frequency. It also yields information about harmonics and noise components of the voice, thus, helps in understanding the turbulence at the level of vocal folds. There are many acoustic measures available for measurement of voice as mentioned in the literature. These can be broadly classified into time based measures and frequency based measures. The time based measures such as perturbation measures are dependent on the cycle boundary identification of the acoustic signal and these measures have not met success in their abilities to consistently and reliably quantify the voice. Other set of acoustic measures are frequency based measures/spectral measures. These measures overcome the limitation of traditional based measures. Cepstral analysis is one among the derived measures of spectral measure. Cepstral-based measures profit from the fact that they are computed via frames of signal data rather than cycle boundary identification.

"Cepstrum is described as a discrete Fourier transform of the logarithm power spectrum; i.e. it is a log power of a log power spectrum" [2,3]. One among the Cepstral measures is Cepstral Peak Prominence (CPP). It is the difference in amplitude between the Cepstral peak and the corresponding value on the regression line that is directly below the peak. CPP is, thus, a measure of the degree of harmonic organization, which tells how far the Cepstral peak emanates from the Cepstral "background noise" [2]. Another measure is the smoothened Cepstral Peak Prominence (sCPP) in

which the individualized Cepstra of voice signal are averaged over a given number of frames before extracting the Cepstral peak and calculating the peak prominence [2]. These measures are reliable and reproducible.

Periodic voice signals display well defined harmonic configuration in the spectrum and thus, a more prominent Cepstral peak [2]. The Cepstrum graphically shows the extent to which the dominant rahmonic is individualized. Periodic voice signals display well defined harmonic configuration in the spectrum and thus, a more prominent cepstral peak is obtained [2] and voice signal with disturbed periodicity or increased spectral noise as seen in dysphonic voice is associated with a decrease in amplitude of the cepstral peak i.e., lower harmonic energy [2,3]. The CPP is an ideal acoustic measure which is be able to quantify independently a voice signal. It has a robust voice analysis algorithm that measures the degree of harmonic structure in a voice signal. Thus this measure is reliable; it correlates well with the dysphonia severity and can be reproducible. The CPP can be measured in different ways including the use of CSL-4500 model or *Speech tool program* [2,3]. Speech tool program is advantageous over the other methods [4]. As it uses particular algorithm which is developed by [2,3].

Literature indicates potential clinical applications of Cepstral measures in voice evaluation. Most of these studies reported that Cepstral measures have good correlation with the perceptual evaluation of voice and aids in discriminating normal from dysphonic voice [5-7]. CPP and the CPPS correlate with perception of breathiness, with sCPP being the better predictor. Olson [8]. Unlike perturbation measures and Noise to Harmonic Ratio (NHR), CPP and sCPP measures do not rely on the accurate identification of the fundamental frequency; they are based on a peak-to-average calculation. For this reason, measures of CPP and CPPS tend to be more consistent than other measures of periodicity [9]. Studies have also shown that CPP is more reliable indicator of dysphonia than any other approaches because CPP doesn't depend on the accuracy of fundamental frequency (f0) extraction which is difficult to establish in severely disordered voices. It is more reliable measure to analyze both phonation and connected speech [4]. Studies have shown that measurement of CPP derived from the acoustic spectrum correlates best with auditory perceptual classification of dysphonia [10-12].

Ample studies have been done in the recent past on the usefulness of cepstral analysis in differentiating the normal voices from dysphonic voices and in comparison with perceptual scales. But only few western studies have been done using the cepstral analysis to see the age and gender effect. Garrett [13]. Conducted a study in an attempt to establish normative data for cepstral measures (CPP and CPP F0) and Low-High Spectral Ratio (L/H ratio) in adults. The author considered 30 males and 30 females in two age groups that are 20-30 years and 30-40 years. CPP was measured for phonation (of vowel /a/ and /i/) and continuous speech (2nd and 3rd sentences of Rainbow passage). The authors reported that the mean CPP for /a/ was 11.74 dB (SD=1.81) and CPP for continuous speech (with vocalic detection) to be 6.60 (SD=1.16).

An attempt to estimate the presence or absence of age effect, gender effect and interaction between age and gender for the mentioned parameters was done. The results revealed there was a significant gender effect for all the dependent variables on both the tasks. It was found that males had higher cepstral and spectral scores than females indicating a better voice quality in them. Also, the values for vowels (both /a/ and /i/) were greater compared to connected speech samples. There was no significant effect of age on the dependent variables for the vowels /a/ and /i/. The study had certain limitations such as; the young female subjects had a relatively poor voice quality on perceptual evaluation, multiple recording re-takes in order to maintain adequate vocal intensity to meet the requirements of the Analysis of Dysphonia in Speech and Voice (ADSV) software.

In a similar line Sujitha [14]. investigated the age and gender effect for CPP and sCPP in adults. The participants included hundred adults in the age range of 20-40 years. Later they were further subdivided into two groups with age interval of ten years (20-30 years and 30-40 years). In which each group included 50 individuals with equal number of males and females. Voice samples included phonation of vowel /a/,/i/and /u / for a duration of 5 seconds and reading of 300 word Kannada reading passage' Savithri 2007 and the Bengaluru passage. The phonated samples were subjected to Cepstral analysis using *Speech tool program*. The results revealed that, for phonation of vowel /a/, there was a gender effect for sCPP and no gender effect for CPP in 20 to 30 and 30 to 40 age groups and also reported no significant age effect on Cepstral parameter. The limitation of the study was that it included narrow age range, thus revealing minimal effects on the Cepstral measures.

Need for the Study

Young children and adults differ anatomically and physiologically in terms of laryngeal and vocal tract functioning. With the rapid advancement in digital technology and means of signal processing, the usage of Cepstarl analysis of voice across age and gender is questionable since no studies have been done comparing the children and adult groups. And hence the need arises. Also there is limited evidence on acoustic characteristics of voice (Cepstral) that investigate the effect of biological maturation among different groups. The study could also provide as a tool to check progress and provides objective gradation of voice. In the current study it is hypothesized that the cepstral measures may vary among the different age and gender. Hence the study is taken up to investigate the cepstral based voice measures in the mentioned groups.

Aim of the study

The study aimed at investigating Cepstral measures in Young children and adults.

Objectives of the study

The objectives of the study were

a. To investigates the effect of age on CPP and sCPP and

b. To investigate the effect of gender on CPP and sCPP.

 How to cite this article: Usha M.Cepstral Analysis of Voice in Young Children and Adults. Glob J Oto 2017; 11(1): 555804. DOI: 10.19080/GJO.2017.11.555804.

Method

Participants

Two groups of participants were considered for the study. Group 1 consisted of 51 (21 Males and 30 females) participants in the age range of 6-12 years (who have not entered their puberty) with mean age of 9 years for Males and 8years for females as per Erickson's stage of development wherein this age range is considered as childhood group. Group 2 consisted of 43 (22 Males and 21 females) participants in the age range 20-40 years (with adult stable voice) with mean age of 27.3 for Males and 25.5 years for females, and according to Erickson's stage of development this range comes under young adults. Individuals with perceptually normal voice as judged by a Speech language pathologist were included for the study. Inclusion criterion was taken such that the participants were free from upper respiratory tract infections, asthma, allergies, sensory problems such as hearing loss, and motor speech disorders such as dysarthria or apraxia of speech. It was also ensured that none of the participants were actively involved in vocal loading within a day prior to the recording.

Procedure

The participants were informed regarding the purpose of the study and procedures involved. An informed written consent was obtained from all the adult participants and parents of school children before the initiation of recording. All the recordings were performed in a quiet room in a solo sitting for all the participants. The participants were made to sit comfortably on a chair with their back straight and were instructed to phonate vowels /a/ for minimum of five seconds each at their habitual pitch and loudness. There were three trials given with 5 minute of break every time. At end of the each trial participants gave opinion on their voice as having soft, loud or habitual loudness and pitch. The task was repeated whenever the participant revealed their voice as either too loud or too soft or different from their habitual voice. Voice recording was done with laptop (DELL, Inspiron), where the voices were directly recorded in Speech tool program using head mounted microphone (Logitech H110 Headphone Microphone) with 44 kHz sampling rate and with a constant mouth to microphone distance of 5cm. Thus recorded samples were saved in the hard disk using save option in speech tool program for further analysis.

Acoustic analysis

The samples recruited for the study were accessed by the Speech Language Pathologist who was blinded and reported the samples as having normal voice quality using CAPEV. Samples reported as having normal voice quality were edited to retain middle and stable portion of vowel /a/ for duration of 5 sec. *Speech tool program* [2]. Version 1.65 was used to analyze the Cepstral parameters like Cepstral Peak Prominence (CPP) and smoothened Cepstral Peak Prominence (sCPP) for phonation. This program is freely downloadable software which is available from the site http;//homepages.wmich.edu/_hillenbr/. The trimmed samples of phonation of /a/ for 5 sec were retrieved from the PRAAT software (Version 5.2.36) and were opened using Speech tool program and the CPP and sCPP were calculated for each sample. By clicking analysis icon, the Speech tool program automatically calculates CPP, sCPP and means f0 values. These values were tabulated and statistical analysis was performed using Statistical Package for Social Sciences (SPSS) software 20.0 version to meet objectives.

Results and Discussion

Shapiro Wilks test of normality was done to the values of CPP and sCPP for phonation task for both the groups. Data followed normal distribution with p value >0.05.

CPP and sCPP across both Groups

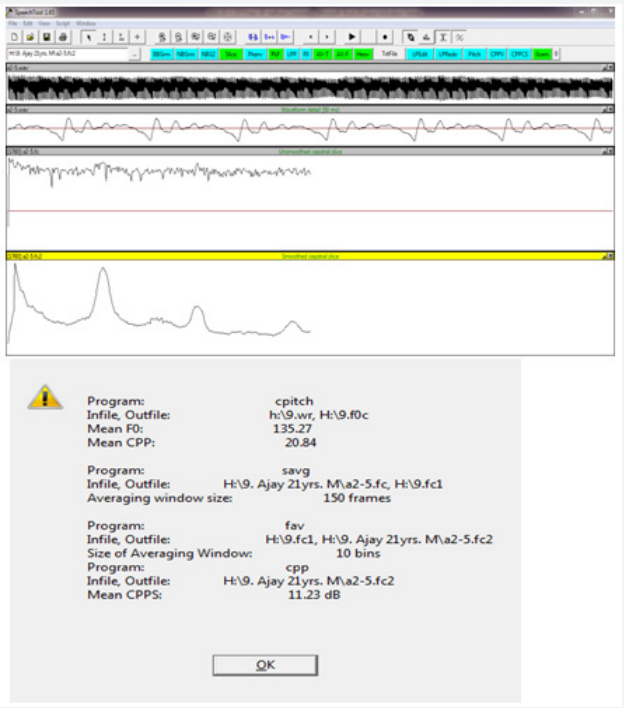

Figure 1: Illustration of unsmoothed and smoothened Cepstrum for vowel /a/ phonation in speech tool software and obtained CPP and sCPP values for the same.

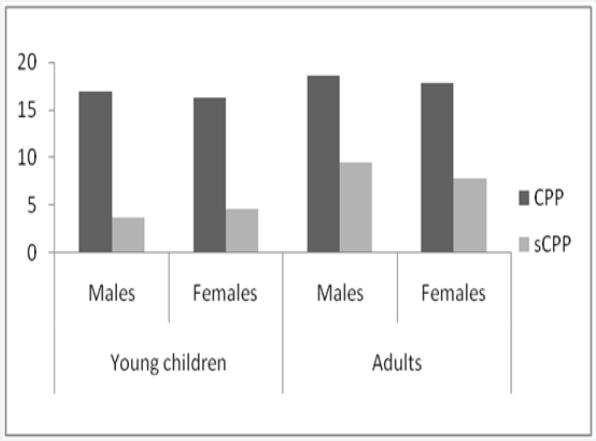

Figure 2: CPP and sCPP for phonation across groups.

How to cite this article: Usha M.Cepstral Analysis of Voice in Young Children and Adults. Glob J Oto 2017; 11(1): 555804. DOI: 10.19080/GJO.2017.11.555804.

Descriptive statistics was done to obtain the mean and standard deviations for both the parameters across both the groups and is given in Table 1 and Figures 1 & 2.

For both the parameters, CPP and sCPP, least mean is seen for young children group and maximum for adult group as listed in table 1. Therefore the mean value can differentiate the voice quality of children and adult groups. The mean CPP in adult group is similar to the studies done by Heman-Ackah [15]. The results also indicated that the overall sCPP values are lower than CPP values. In both groups the lower sCPP values could be due to the effect smoothening, which averages and reduces the artifacts in cepstral peaks. Similar results were also reported by Brinca et al. [7].

Table1: Mean and Standard Deviation (SD) of CPP and sCPP of phonation across the 2 groups.

Phonation						
		CPP (dB)			sCPP (dB)	
		N	Mean	SD	Mean	SD
Young children	M	21	16.87	2.13	3.67	0.79
Adults	F	30	16.23	2.80	4.57	0.92
	M	22	18.63	2.58	9.55	1.31
	F	21	17.77	2.44	7.86	1.20

Mean values of males and females in young children group is almost overlapping as it is due to voice quality and fundamental frequency of children would be similar before puberty, it will get well differentiated during puberty and it becomes stable during adulthood; similar findings was seen in the present study. On the other hand in adult group, males had slight higher values compared to females. The decrement of measures in females can be attributed to the factor that females usually have softer habitual voice, hence reducing the CPP. Also, it could be because of the posterior phonatory gap which relatively increases the noise component in female voice compared to male voice [6].

Effect of age and gender on CPP and sCPP

To study the effect of age and gender on CPP and sCPP parametric tests were performed. One way ANOVA was performed independently. For the parameter CPP there was main effect of age on CPP [$F(1, 94) = 10.778$, $p < 0.05$] and there was no main effect of gender with $p>0.05$.There was interaction effect CPP *age with $p<0.05$ and no interaction effect for CPP*gender with $p>0.05$ and there was overall interaction seen CPP*age*gender with $p<0.05$. In a similar way for sCPP, there was main effect of age [$F(1, 94) = 307.757$, $p < 0.05$] and there was no significant difference across gender with $p>0.05$ and there was interaction effect seen. Whereas there was interaction effect seen for age with $p<0.05$ and overall interaction effect (sCPP*age*gender) was present with $p<0.05$.

Overall results of the study indicated no significant effect of gender on Cepstral parameters, this can be supported by the mean values both the groups where males and females have performed similar ways, whereas there was significant main effect of age on cepstral parameters with $p<0.05$. For this, though no supporting studies are available for this specific age groups (8 to 10 and 20 to 40 age groups), a study done by Garret [13] in 20 to 30 and 40 to 50 years age groups reported there was no significant age effect on the vowels /a/ and /i/ on CPP measures; and study done by Sujitha [14]. Reported a gender effect for sCPP and no gender effect for Cpp in 20 to 30 and 30 to 40 age groups and also reported no significant age effect on Cepstral parameter.

Voice changes with respect to acoustical and perceptual level at the early stage are always attributed to the changes in the anatomical structure. Children possess unstable growing structures before the puberty and at the time of puberty other aspects like hormonal changes come into picture along with the anatomical changes. At the end of puberty a stable stage will appear where the acoustical parameters of the voice remain stable until there is deterioration due to senescence. Therefore CPP and sCPP can be easily differentiated in children from adult groups. Because of this reason there was significant age effect seen in the present study.

Conclusion

The results of the present study would augment in understanding the Cepstral based voice characteristics in young children and adult groups, the variations among them and the changes in the voice characteristics with biological maturation. The result of the study showed that there was overall main effect of age was seen for CPP and sCPP and there was no gender effect on CPP and sCPP. This indicates that Cepstral measures are affected by factors such as age. However, the previous research results have indicated that vocal loudness/ intensity and vowel type have a significant effect on measures of the CPP. Monitoring loudness instrumentally is not considered for the study and only vowel /a/ was taken in the present study. Therefore future studies could consider and control this factor and also recommends for future studies which are warranted with wider range of ages so as to verify the effect of the pediatric, adult and aging voice on Cepstral measures and hence developing age specific reference data if necessitates.

References

1. Ballenger WL, Ballenger HC (1969) Diseases of the Nose Throat and Ear.
2. Hillenbrand J, Ronald A, Cleveland RA, Erickson R L (1994) Acoustic Correlates of Breathy Vocal Quality. Journal of Speech and Hearing Research 37(4): 769-778.
3. Hillenbrand J, Houde RA (1996) Acoustic correlates of breathy vocal quality dysphonic voices and continuous speech. Journal of Speech and Hearing Research 39(2): 311-321.

 How to cite this article: Usha M.Cepstral Analysis of Voice in Young Children and Adults. Glob J Oto 2017; 11(1): 555804. DOI: 10.19080/GJO.2017.11.555804.

4. Heman Ackah YD, Michael DD, Goding GSJr (2002) The relationship between cepstral peak prominence and selected parameters of dysphonia. J Voice 16(1): 20-27.

5. Kumar BR, Bhat JS, Prasad N (2010) Cepstral Analysis of Voice in Persons with Vocal Nodules. J Voice 24(6): 651-653.

6. Kumar BR, Bhat JS, Fahim S, Raju R (2011) Cepstral analysis of voice in unilateral adductor vocal fold palsy. J Voice 25(3): 326-329.

7. Brinca LF, Batista APF, Tavares AI, Gonçalves IC, Moreno ML (2014) Use of cepstral analyses for differentiating normal from dysphonic voices A comparative study of connected speech versus sustained vowel in european portuguese female speakers. Journal of Voice 28(3): 282-286.

8. Olson ELD, Goding SG, Deirdre D, Michael DD (1998) Acoustic and perceptual evaluation of laryngeal reinnervation by ansa cervicalis transfer. The Laryngoscope 108(12): 1767-1772.

9. Heman Ackah YD, Heuer RJ, Michael DD, Ostrwoski R, Horman M, et al. (2003) Cepstral Peak Prominence A more reliable measure of Dysphonia. Annals of Otology Rhinology Laryngology 112: 323-333.

10. Awan SN, Roy N (2005) Acoustic prediction of voice type in women with functional dysphonia. Journal of Voice 19(2): 268-282.

11. Maryn Y, Roy N, De Bodt, M, Van Cauwenberge P, Corthals P (2009) Acoustic measurement of overall voice quality a meta-analysis. The Journal of the Acoustical Society of America 126(5): 2619-2634.

12. Moers C, Möbius B, Rosanowski F, Nöth E, Eysholdt U, Haderlein T (2012) Vowel and text based cepstral analysis of chronic hoarseness. J Voice 26(4): 416-424.

13. Garrett R (2013) Cepstral and Spectral-Based Acoustic Measures of Normal Voices. Thesis University of Wisconsin Milwaukee.

14. Sujitha PS (2015) Cepstral analysis of voice in phononormic adults Unpublished Masters' Dissertation Submitted to University of Mysore.

15. Heman Ackah YD (2004) Reliability of Calculating the Cepstral Peak without Linear Regression Analysis. J Voice 18: 203-208.

DOI: 10.19080/GJO.2017.11.555804