Available online at www.elixirpublishers.com (Elixir International Journal)

Electronics Engineering

Elixir Elec. Engg. 85 (2015) 34616-34623

ISSN: 2229-712X

Evaluation of Velopharyngeal Dysfunction (VPD) Speech Disorder with Special Focus on Normalized Pitch Variation Characteristics

Manasi Dixit[1] and Shaila Apte[2]

[1]Department of Electronics K.I.T.'s College of Engineering, Kolhapur, 416234 Maharashtra, India.

[2]Rajarshi Shahu College of Engineering, Pune, Maharashtra, India.

ARTICLE INFO

Article history:

Received: 16 June 2015;

Received in revised form:

22 August 2015;

Accepted: 26 August 2015;

Keywords

Normal and VPD disordered Speech data,
Pitch variations, Jitter,
Shimmer, HNR,
Speech-Pause Index,
Percentile f_0 Characteristics,
% CQ –Characteristics,
All Tx –Regular Tx Pitch Period Characteristics.

ABSTRACT

Resonance disorders are observed in patients with a history of cleft palate or cleft lips. Evaluation of speech pathology helps for the proper treatment of a resonance disorder. The speech parameters are identified here which characterizes the disorder and correction system is designed to improve the speech quality. The speech signal samples of people of age between five to eighty years are considered for the present study. These speech signals are digitized and enhanced and analyzed for the Jitter, Shimmer, HNR, Pitch variations Tx graphs , Normalized Percentile f_o characteristics and % Close Quotient EGG graphs using MATLAB, PRAAT, SFS and EXCEL platforms.

Introduction

A resonance disorder can occur when the Velopharyngeal mechanism does not function adequately to prevent the transmission of sound into the nasal cavity. Resonance can also be abnormal when there is a blockage in the nasopharynxso that sound transmission is impeded during passage into the nasal cavity for nasal phonemes. Anything that disrupts the normal balance of oral and nasal resonance can result in a resonance disorder .It is caused by a variety of structural abnormalities in the resonating chambers in the oral cavity for speech, or by Velopharyngeal dysfunction [1].

Velopharyngeal valve is closed for oral sounds and it is particularly important for pronunciations of "pressure-sensitive" consonants [2] and all vowels [3]. Velopharyngeal valve is in open position for nasal sounds (m, n, and ng) [3]. The VP Valve is applied for transmission of sound energy and thus the air flow occurs in the oral and nasal cavities during speech [2][4].

Literature Review

Michael A. Berger [5] has discussed about vowel nasalization. When the velum lowers during utterances of vowels, the Velo pharyngeal port is opened. So then the nasal cavity and the main vocal tract get connected acoustically, developing an acoustic quality known as nasality. Nasal coupling results in energy loss of low frequencies, damping of oral cavity formant F_1, and introduction of nasal formants due to the resonances of the nasal cavity and sinuses. However, the acoustic effects of nasalization vary with velar position, with respect to different speakers and vowels.

Ann W. Kummer et.al. have discussed evaluation and treatment of resonance disorders. Size and shape of the resonating cavities that is pharyngeal cavity, oral cavity and nasal cavity determines Resonance [4]. Vowels are resonance sounds. They are produced during articulation by changing the physical characteristics like size of the oral cavity. Smaller resonating cavity results in a higher formant frequency/pitch whereas larger resonating cavity results in a lower formant frequency/pitch.

Velopharyngeal dysfunction (VPD) can be caused by a history of cleft palate, short velum or deep pharynx, Irregular adenoids and Enlarged tonsils [4]. VPD can result into Dysarthria, and Apraxia .Velopharyngeal Mislearning occurs due to abnormal posterior or nasal articulation of certain sounds, phoneme-specific nasal air emission (PSNAE) [6].Rainsonata [2] has developed Flashcards which specify different symptoms regarding Cleft Palate or VPD speech disorder.

James F. Lubker et.al. [7] have discussed five measures of articulatory position made from individual cinefluorographic frames to determine the effects of the divided face mask on the articulators. The nasal air flow depends upon the amount of Velopharyngeal opening and it also depends upon the amount of oral constriction.

Velopharyngeal structures-The system includes the velum, the lateral pharyngeal walls and the posterior pharyngeal wall. Velopharyngeal closure is accomplished by the coordinated movement of all of these structures.

Normal speech -During normal speech, the velum moves in a superior and posterior direction with a type of "knee" action in order to achieve closure against the posterior pharyngeal wall. For normal speech and resonance, Velopharyngeal closure should be complete during the production of oral sounds.

Nasal consonants -For nasal consonants (m, n, ng), the Velopharyngeal port remains open to allow sound transmission into the nasal cavity, which is the primary resonating chamber for these sounds. Very little sound energy resonates in the oral

Tele:
E-mail addresses: manasidixit64@gmail.com

cavity during the production of nasal sounds [2]. This is due to the fact that the acoustic energy begins by traveling in a superior direction toward the nasal cavity, and continues without significant obstruction.

Resonance Disorders- Resonance occurs when there is modification of the sound being generated from the vocal cords .It provides the information regarding the quality of perceived sound during speech. Size and shape of the resonating cavities-pharyngeal cavity, oral cavity and nasal cavity determine the resonance. It can be viewed as a function of the Velopharyngeal valve. Vowels are resonance sounds. They are produced by changing the size and shape of the resonating oral cavity [4]. Effect of Cavity Size on Resonance frequencies is such that smaller resonating cavity results in a higher formant frequency/pitch frequency and larger resonating cavity results in a lower formant frequency/pitch frequency [3]. It is a reciprocal relationship.

VP Insufficiency

VP Insufficiency is due to history of cleft palate, or very short velum with a notch on posterior nasal surface or deep pharynx. Normal VP closure requires a tight seal [2] where as Cleft palate/lips condition prevents a tight seal .It can cause small gap and nasal emission. Enlarged Tonsils can extend into pharynx and thus prevent a tight VP seal.

Hyper nasality-Hyper nasality is a resonance disorder due to Velopharyngeal inadequacy (VPI). As a result of an inadequate Velopharyngeal valve, sound resonates into the nasal cavity inappropriately, which affects the quality of speech. Hyper nasality is particularly perceptible on vowel sounds because these sounds are voiced and relatively long in duration. Sound resonating in the nasal cavity is most perceptible on vowels. Voiced oral consonants become nasalized m/b, n/d, etc [2]. Causes of Hyper nasality are Velopharyngeal dysfunction, Open palate, thin velum, and Misarticulation.

Effects of Nasal Cavity Resonance

- Nasal Air Emission-Occurs with or without hyper nasality, Air leaks through the valve, occurs on high pressure consonants, particularly voiceless consonants.
- Nasal Emission with Large Opening- It is indicated by no impedance to airflow, Soft, low intensity sound. It can cause Weak or omitted consonants, Short utterance length, Compensatory articulation productions.
- Weak or Omitted Consonants-It is due to loss of oral air pressure, greater the nasal air emission, the weaker the consonants. It affects pressure-sensitive consonants [2].
- Short Utterance Length-Leakage of air pressure causes need to increase respiratory effort and leads to more frequent breaths, Utterance length becomes shortened.
- Compensatory Articulation Productions
- Manner of speech production- It is maintained. Place of speech production is usually nasal, velar, pharyngeal or glottal. VP valve will be open, resulting into nasal emission.
- Compensatory Productions Velar plosives (stops) are replaced by Velar fricatives, Pharyngeal fricatives. Pharyngeal plosives are replaced by Posterior nasal fricatives [61]. Glottal stops are replaced by Nasal sniff.

System Implementation - The implemented system includes following processes.

Procedure

The present work attempts to find few parameters from pathological speech for confirmation of Cleft palate or VPD. Researchers have used lot many parameters. We try to reduce the computational cost and reduce the number of parameters. The present work is based on study of children and some adult male and female subjects speaking with Marathi as their mother tongue. The speech data of normal subjects/children and pathological subjects/children of the same age group between 3 to 10 years is collected. The children were trained to utter similar words before recording. The speech data of normal adult male and female subjects and pathological adult male and female subjects of the same age group between 18 to 65 years is collected

Database

The standard database is not available. The speech data of 11Cleft palate or VPD speakers comprising of above 100 words uttered by each subject are collected. The speech database consists of isolated words, connected words, fast uttered sentences and songs for e.g. School-Prayer, National anthem and Pledge ,Nursery Rhymes ,famous film songs etc. The speech data was recorded using Sony Intelligent Portable Ocular Device (IPOD) and recording facility in COLEA freeware [8] in digital form. The recording was carried out in a pleasant atmosphere and maintaining the children and other subjects in tension-stress free environment. The recorded signal is transformed into '.wav 'file by using GOLDWAVE software. The data was collected at Chetana Vikas Mandir, a special school established to educate Mentally Retarded children as well as children with various disorders. It is located at Kolhapur, India. The data is also collected from the patients under the treatment of speech therapists and ENT specialists in Kolhapur city. We got the database labelled by consulting the doctors as Cleft palate or VPD speech data.

Evaluation of Speech

The present work attempts to confirm the Cleft palate or VPD disorder from the speech signal by extracting only important segmental and supra segmental acoustic indices. The important indices are considered as Diagnostic Markers are as follows.

1. Evaluation of Fundamental frequency fo, Jitter, Shimmer, and HNR for the analysis of harshness and breathiness in the voice to be done in the Training Phase to confirm the speech as of pathological speech category. Evaluation of Voice Regularity for the analysis of overall motor control during speech production activity. This is done with reference to the parameter threshold ranges specified in Table 1
2. Fundamental Frequency Analysis- The fo mean value lies in the high range between 280 Hz to 500 Hz for adult male female speakers. It is because the resonance occurs in nasal cavity instead of oral cavity during speech production and hyper nasality exhibited by the VPD disordered speakers.
3. Percentile fo track variations – the fundamental frequency fo variations with respect to percentile values from 0 to 100 percentile range. The Linear or nonlinear nature of the characteristics is important. The Cleft palate or VPD is confirmed by low gradient index in 0 to 20 fo percentile range and Nonlinear curved naturein 25 to 100 fo percentile range.
4. % Close Quotient (CQ) graph simulates the Laryngograph and indicates the close phase of Glottis pulse signal or vocal folds vibrational cycle. The mean CQ, range of variation of CQ [125] and CQ variations with respect to total time duration of speech sample are important parameters.

All 'Tx' graph indicates histogram of all pitch cycles for the total time duration of speech sample. Regular 'Tx' graph [125] indicates the histogram of the regular pitch periods which vary within +/- 10% with respect to the adjacent pitch periods. The similarity between the two graphs and the mean fo value of the person are important parameters.

The present work attempts to confirm the Cleft palate or VPD disorder from the speech signal by extracting various segmental and supra segmental acoustic indices. The important indices are considered as Diagnostic Markers.

The acoustic indices are evaluated for all isolated word and tool generates 'Ls' signal from the input speech data, which is compatible to Laryngograph signal which is not accessible to us. The 'SFS' tool provides the Qx a histogram of the closed quotient values found in the recording [10].

The closed quotient is found from the Laryngograph signal .It is described as an estimate of the percentage time the vocal folds remained closed in each pitch period.

The SFS tool provides the Dx1 a histogram of all the pitch periods, distributed according to their fundamental frequency, and Dx2 a histogram of all the regular pitch periods distributed according to their fundamental frequency [10].

It is defined as Regular pitch period if it varies in duration by less than +/- 10% with respect to the adjacent pitch periods. Using the acoustic indices the relationship between the Electroglottograpgh (EGG) measures [28] and the physical movements of the vocal folds is expressed as ratios between the temporal measures of open phase of vocal fold movement with closed phase and also between different phases of movement with the full glottal period. Continuous sentence type speech data samples (above 100 words by each subject) uttered by every pathological and normal subject.

The ready to use softwares are used for the development phase .We have used the developmental tool' Praat' [9] to extract intensity, pitch and the formant frequencies f_1 , f_2 ,f_3 and HNR for the speech samples. We have used the developmental tool 'SFS' to extract the parameters voice regularity, Jitter, Shimmer, fo mean mode values.

The Laryngeal Quality Analysis, Glottis Pulse Analysis is performed by using SFS software [10]. SFS reports indicate exactly similar 'all Tx ' and 'regular Tx ' glottis pulse graphs in case of normal subjects, whereas for pathological subjects the 'regular Tx ' pulses produced are very low insufficient in time domain as well as variable in frequency domain .

Classification of Speech on the basis of Segmental and Supra segmental acoustic indices

The following parameters are extracted by using the developmental tool available as an open source software SFS and PRAAT from every speech data to classify whether the speech is Normal or Pathological. The Normal range and Pathological range of the parameter values are as described in Table 1 below.

*VPD : Velopharyngeal dysfunction or Resonance disorders are observed in patients with a history of cleft palate or cleft lips

System Development - The system is developed using two modes training mode and testing mode.

Training Mode- In training mode 50 speech samples are used to train the system. The Laryngograph comprising of % Close Quotient (CQ) with respect to time indicative of the close phase of Glottis pulse signal or vocal folds vibrational cycle and the comparison of histograms of all 'Tx ' which is all pitch periods with respect to regular 'Tx' pitch periods are plotted for these 50 speech samples to confirm cleft palate or VPD characteristics. The observations for Tx are as follows.

1. A pitch period is defined as regular if it varies in duration by no more than 10% from the period before or after it.
2. In case of normal persons, regular Tx graph matches with all Tx graph.
3. For cleft palate or VPD persons, regular Tx contains almost 50 % of frequency range as compared to all Tx.

We have developed a MATLAB routine to evaluate fundamental frequency f0. Then the Normalization routine is developed in Microsoft Excel. The algorithm for evaluation of Normalized fo Variation in Percentile Domain is as follows.

The algorithm for evaluation of fo track is as follows.

4. The percentile value of fo maximum level is considered to be 100 percentile.
5. Hence according to the data values of fo variations as per the speech sample 0-5-10-15-----95-100 percentile values of every speech sample are calculated.
6. The fo frequency variations are plotted with respect to percentile values. This graph provides a very good measure of fundamental frequency analysis to differentiate between normal speech and pathological speech .It also confirms cleft palate or VPD disorder.

The graph is a nonlinear curve from 0 to 90 percentile range and then high gradient in the range 0.9 to 2.5 for cleft palate speakers.

7. For normal speech percentile fo track graphs are linear from 5 to 95 percentile range with very low gradient 0.05 to 0.3

The observations for CQ are as follows.

1. CQ is a time Vs CQ graph. It indicates that the CQ values are in the range 15 % to 75 % for Cleft Palate or VPD speakers. For normal speech, CQ variation is observed in the range15 % to 60 %

Block Schematic

Diagram of the system implemented for evaluation of Cleft palate or VPD disorder is as shown in Figure 1 below.

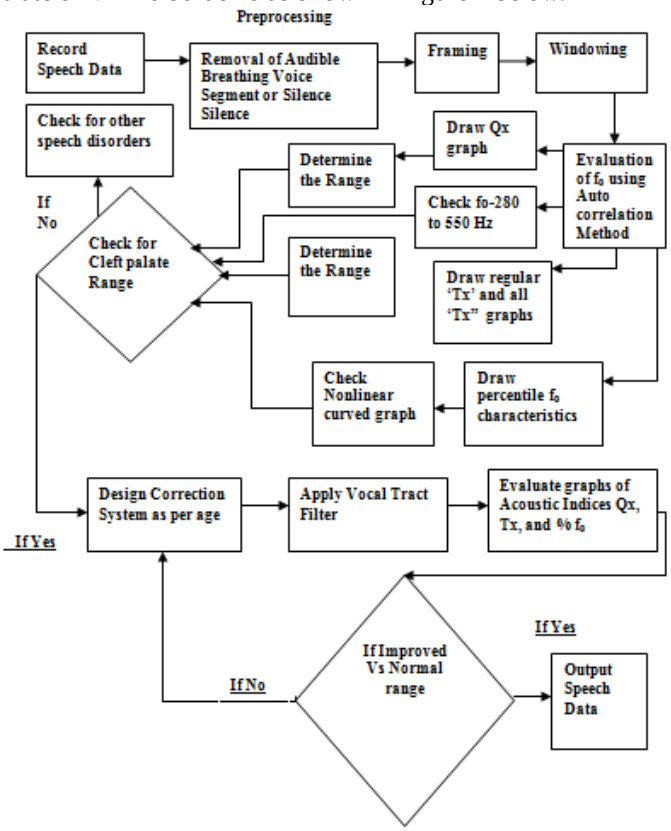

Figure 1. Block Diagram for Evaluation of Cleft Palate or VPD

Graphs of some of the Diagnostic Markers

Figure 2 to 4 show the graphs for all Tx and regular Tx for 3 different pathological persons. It is observed in all 3 cases that the regular Tx is produced for 50 % of frequency range of All Tx graph.

Diagnostic Markers –Speaker 1-Akshay1

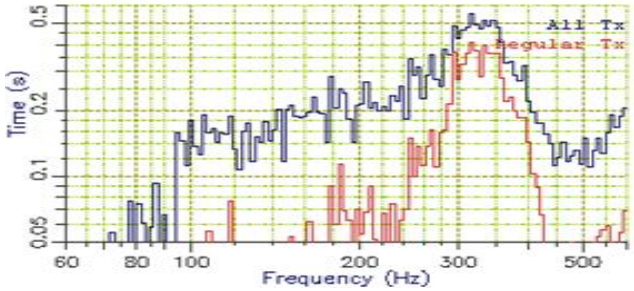

Figure 2. % Tx Glottal Pulses-Speaker1-Akshay1

Note–It is observed that

1. The regular Tx pulses are produced in the 110 Hz to 550 Hz pitch spectrum.
2. The time duration of regular Tx pulses is almost 50% of All Tx graph

Diagnostic Markers –Speaker 2-Vikas 6

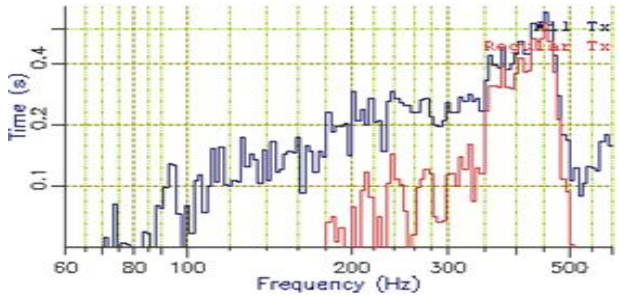

Figure 3. % Tx Glottal Pulses-Speaker2-Vikas6

Note –It is observed that

1. The regular Tx pulses are produced in the 180 Hz to 500 Hz pitch spectrum.
2. The time duration of regular Tx pulses is almost 50% of All Tx graph

Diagnostic MarkersShubham24

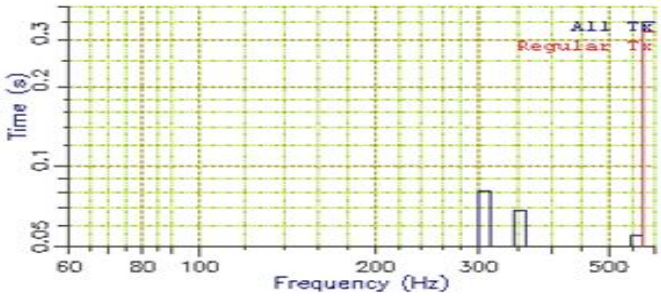

Figure 4 % . Tx Glottal Pulses-Speaker3-Shubham 24

Note-It is observed that

1. The regular Tx pulses are produced in the 520 Hz to 530 Hz pitch spectrum.
2. The time duration of regular Tx pulses is almost 50% of All Tx graph

Figure 5 to 7 show the graphs for %CQ for 3 different pathological persons. It is observed in all 3 cases that the range of % CQ is from 15 % to 75 %.

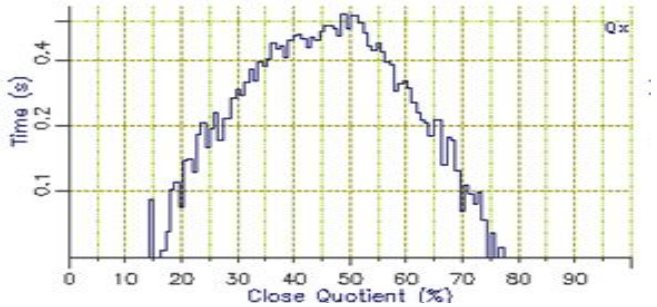

Figure 5. % Close Quotient wrt Time Akshay1

Notes-% Qx graph indicates 14% to 78% wide Qx range with mean value %CQ – 45.9

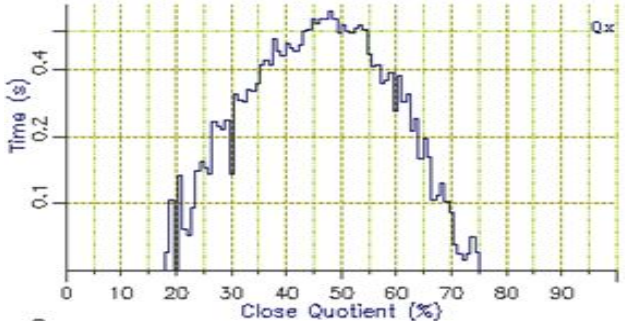

Figure 6. % Close Quotient wrt Time Vikas 6

Note-% Qx graph indicates 18% to 75% wide Qx range with mean value %CQ – 45.9

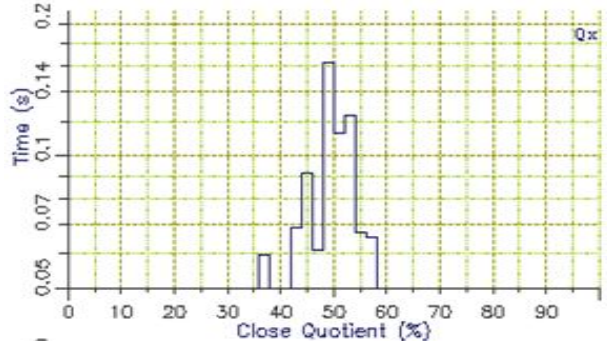

Figure 7. % Close Quotient wrt Time Shubham 24

Note

% Qx graph indicates 35% to 58% wide Qx range with mean value %CQ – 48.4.

Figure 8 to 10 show the graphs for f_0 percentile for 3 different pathological persons. It is observed in all 3 cases that percentile glottal frequency graph indicates nonlinear curved graph from 0 to 90 percentile range. After 90 percentile range, it has high gradient.

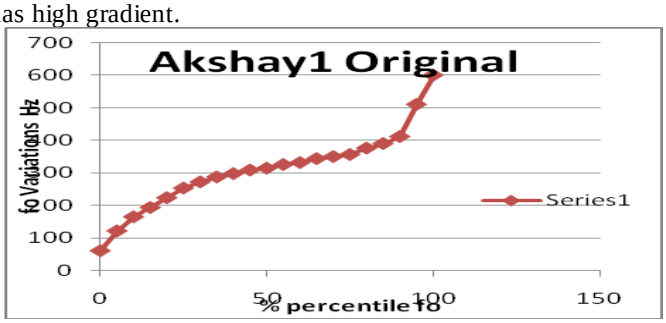

Figure 8. Percentile fo track variations for Akshay1

Note- Percentile glottal frequency graph indicates nonlinear curved graph from 0 to 90 percentile range and 90 to 100 percentile range – 1.88 high gradient index

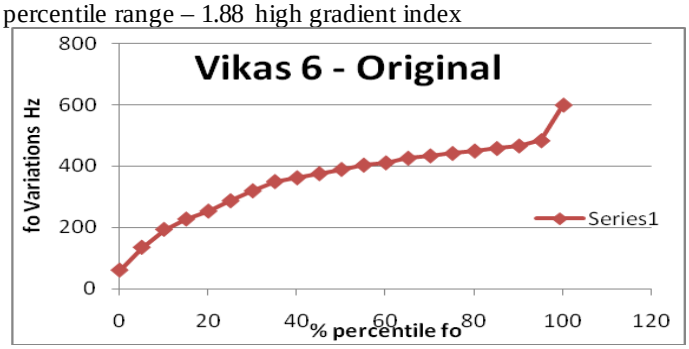

Figure 9. Percentile fo track variations Vikas 6

Note- Percentile fo track indicates nonlinear curved graph from 0 to 95 percentile range and 95 to 100 percentile range – 2.33 high gradient index

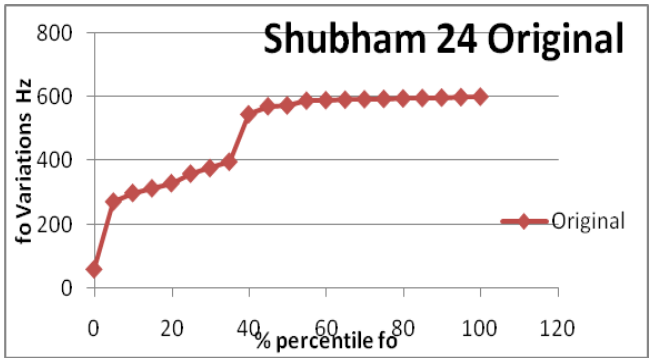

Figure 10. Percentile fo track variations Shubham 24

Note- Percentile fo track indicates nonlinear graph from 5 to 35 percentile range- 0.42 high gradient and 40 to 100 percentile range – 0.09 low gradient index

Testing Mode- In testing mode remaining 50 speech samples are used for confirmation of Cleft palate or VPD. The testing mode checks the Laryngograph characteristics, CQ graph and fo track to confirm Cleft palate or VPD. The following observations are made.

1. Fundamental Frequency fo mean is in the range 280 Hz to 500 Hz due to resonance in nasal cavity instead of oral cavity and hyper nasality in cleft palate disordered speech.
2. It is observed that Laryngograph comprising of regular Tx pulses contains less than 50 % of the frequency spectrum as compared to all Tx. Regular Tx pulses are present in 140 Hz to 550 Hz range.
3. Time Vs closed quotient graph indicates closed quotient range more than 50 % wide with mean %CQ within 42% to 48%.
4. Percentile fo track has a Nonlinear curved nature from 0 to 90 percentile range and a high gradient (slope) in 90 to 100 percentile range –0.5 to 2.8

System Developed for correction

We have used MATLAB based developmental tool E-System compatible with SFS developmental software for trying methods for correcting Cleft palate or VPD speech. The developmental tool COLEA and Adobe Audition are used for preprocessing the speech samples. In preprocessing the silence zone or the audible breathing voice segments are removed. Using E-System software following processing blocks can be designed.

- Amplifier /Attenuator – Design specifications are gain and Bandwidth
- Low Pass Filter- Design specifications are Cut off Frequency
- High Pass Filter- Design specifications are Cut off Frequency
- Band Pass Filter- Design specifications are lower and upper Cut off Frequency
- Band Stop Filter- Design specifications are lower and upper Cut off Frequency
- Vocal Tract Filter- Design specifications are f_1,f_2 and f_3 formant Frequencies.
- Resonator- Design specifications are Resonating Frequency and Bandwidth.

The system applied for correction is developed with the help of following filters.

Band Pass Filter- The lower cut off frequency is in the range 10 Hz to 100 Hz and the upper cut off frequency selected should be such that the second formant frequency f_2 should lie in the pass band. Hence it is selected as 1500 Hz, 2000 Hz or 2500 Hz as per male, female or children based on pitch frequency range.

Resonator - The resonating center frequency selected should be such that the second formant frequency f_2 should lie in the pass band. Hence it is selected as 1500 Hz,2000 Hz or 2500 Hz as per male, female or children based on pitch frequency range.

Vocal Tract Filter – It is realized as a cascaded combination of three resonators acting as per three formant frequencies. The standard adult male formant frequencies are 500 Hz, 1500Hz and 2500 Hz. The first formant frequency is amplified by 20 dB. The second formant frequency is amplified by 10 dB and the third formant frequency is maintained at 0 dB. Hence this filter boosts up the input speech signal spectrum as per the formant frequencies. In case of pathological speech the amplitudes of upper formants **are** degraded .Hence the VTF is the better solution to lift up the second formant spectrum.

System Applied for Correction of Cleft palate or VPD Disorder

The performance of Vocal Tract Filter was found to be better in comparison with Band Pass Filter and Resonator Filter during the Training mode. Hence the Vocal Tract Filter is applied for correction of Cleft palate or VPD speech. The Vocal Tract Filter is designed with formant frequencies f_1, f_2, f_3 of individual speakers. The standard adult male formant frequencies are 500 Hz, 1500Hz, and 2500 Hz.

After applying the VTF to Cleft palate or VPD the following observations are made.

5. It is observed that Laryngograph comprising of regular Tx pulses graph shows improvement and contains more than 50 % of the frequency spectrum as compared to all Tx
6. Time Vs closed quotient graph indicates closed quotient range more than 50 % wide with mean %CQ within 42% to 48%.
7. Percentile fo track has become almost linear with a low gradient (slope)- In 10 to 85 percentile range – 0.17 to 0.24 and a high gradient (slope) in 85 to 100 percentile range –0.55 to 1.94

Improvement in the diagnostic markers due to application of VTF Filter is indicated below with the help of Regular Tx graph, % Close Quotient (CQ) graph and percentile fo track graphs through Figure 11 to Figure 15 respectively.

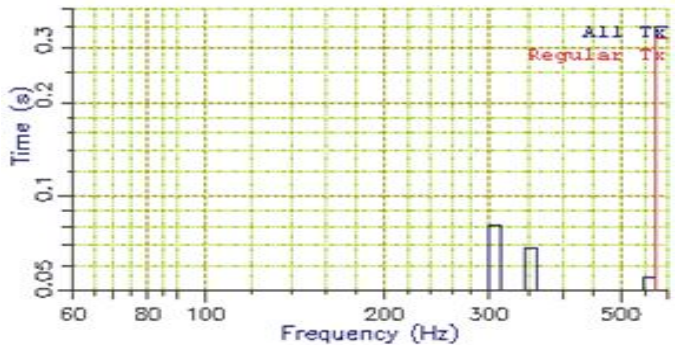

Figure 11. % Tx Glottal Pulses-Speaker3-Shubham 24-Original

Diagnostic Markers Shubham24 after VTF

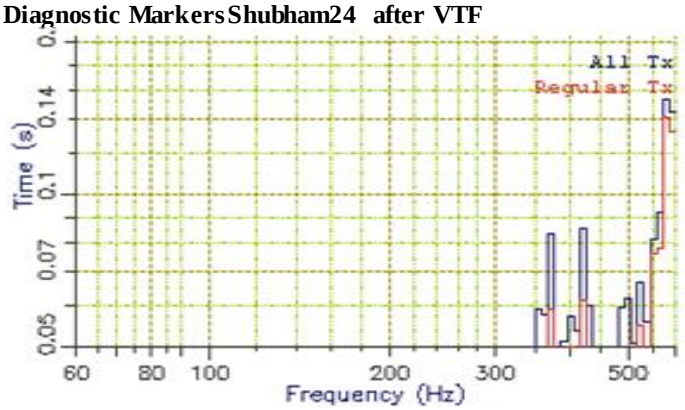

Figure 12. % Regular Glottal Pulses VTF processed

Table 1. Range of Segmental and Supra segmental Acoustic Indices for Classification of Normal Speech and Pathological Speech

Parameter	Normal Speech Range			Pathological Speech Range
Fundamental Frequency-Pitch	Children	Adult Male	Adult Female	for VPD* 300 -550 Hz otherwise same as Normal Speech
	208-440 Hz	85-196 Hz	155-334Hz	
Jitter (mean)	Range 0.0 % to 18%			Range 14% to 45%
Shimmer (mean)	Range 0.0 % to 5 %			Range 0.0% to15%
HNR (mean)	Range 12 dB to 45 dB			Range 5dB to 11 dB
Voice Regularity	Range 50% to 95 %			Range 5 % to 45%
Audible Breathing Voice Segment	Not Present			Always Present

Table 2. Range of diagnostic markers for 50 % of Cleft palate or VPD Disordered speech sample and 50 % of normal speech

Diagnostic marker	Range of values for Cleft palate or VPD Disordered speech	Range of values for normal speech
Time Vs % CQ	35 % to 75 %	10 % to 60 %
Time Vs frequency TX graph	Regular Tx has less than 10 % frequency range as compared to all Tx 150 to 600 with intermittent pulses	Regular Tx matches with all Tx for more than 90 %.
Frequency Vs Percentile fo Graph	Nonlinear-Curved Characteristics for almost 90 percentile range	Linear Characteristics with very low gradient for almost 90 percentile range

Table 3. % confirmation for 50 % of Cleft palate or VPD Disordered speech and normal speech samples

Parameter used	% samples confirmed for Cleft palate or VPD Disordered speech	% samples confirmed for normal speech
Time Vs frequency TX graph	98 %	100 %
Percentile Fo track variation	100 %	100 %
All 3 parameters % CQ, Tx, F0	97 %	100%

Table 4. Results of correction for Cleft palate or VPD Disordered speech after applying correction using VTF

Diagnostic marker	Range of values for Cleft palate or VPD speakers	Range of values for Cleft palate or VPD after applying correction using VT Filter
Time Vs % CQ	35 % to 75 %	30 % to 60 %
Time Vs Frequency TX graph	Regular Tx has less than 10 % frequency range with intermittent pulses for small time period as compared to All Tx graph in 150 Hz to 600 Hz range.	Regular Tx graph exists for almost 40% frequency range and present for 50% time period as compared with All Tx graph.
Percentile Glottal Frequency fo variation	Nonlinear-Curved Characteristics for almost 90 percentile range	Approximately Linear smooth Characteristics with 0.3 and 0.6 gradient level for almost 90 percentile range. Improved Characteristics

Regular Tx pulses graph shows improvement and contains more than 50 % of the frequency spectrum as compared to all Tx(Refer to Figure 4 for original & regular Tx. The regular Tx pulses are produced in the 520 Hz to 530 Hz pitch spectrum in original where as in 330 Hz to 550 Hz pitch spectrum in VTF processed graph.

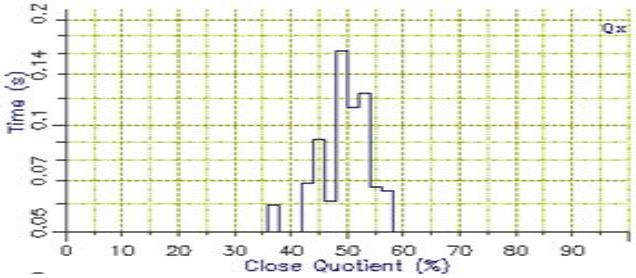

Figure 13. % Close Quotient wrt Time Shubham 24-Original

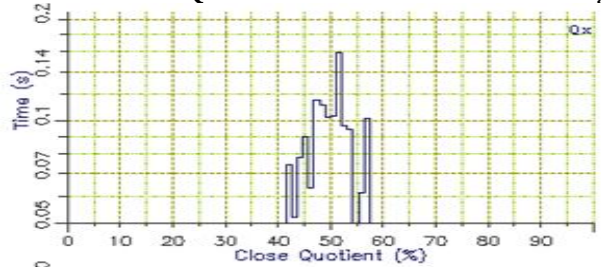

Figure 14. % Close Quotient wrt Time VTF Processed

% CQ varies between 43% to 57% after VTF processing. Refer to Figure 7 for original graph before processing.% Qx graph indicates 35% to 58% wide range in original where as 42% to 58% wide range in VTF processed.

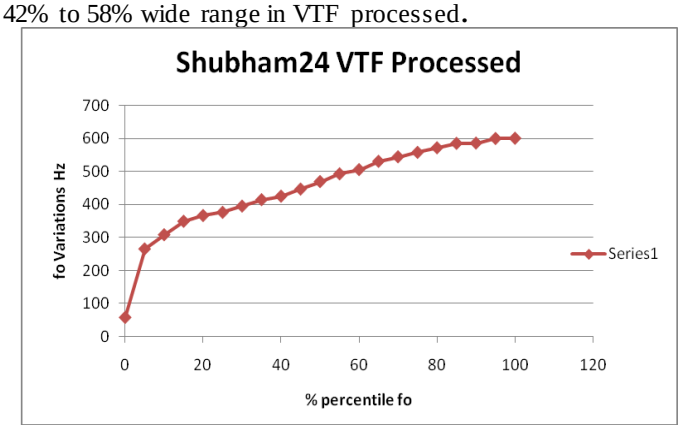

Figure . 15 Percentile f_0 track variations Shubham 24 after VTF processing

Note- Original percentile fo track indicates nonlinear graph from 5 to 35 percentile range- 0.42 high gradient and 40 to 100 percentile range – 0.09 low gradient index and VTF processed percentile fo track indicates almost linear smooth graph with 0.37 gradient index.

Figure 16 indicates the result of comparison after processing using VTF, resonator and BPF.

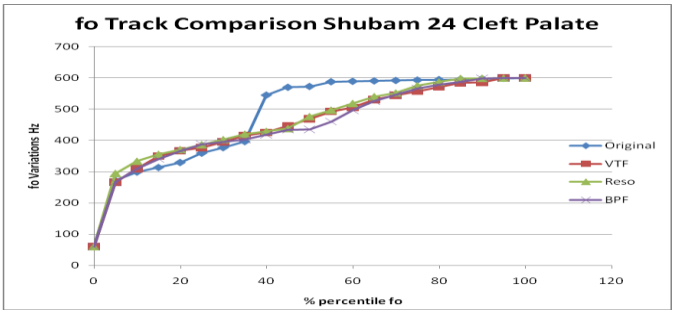

Figure 16. Comparative Response of the Correction System for Cleft palate

Note-It is evident from the Figure 16 graphs that the VTF processed graph is smoother. Hence it is preferred as Correction System**.**

Segmental and Supra segmental Acoustic indices Analysis

The analysis of segmental and supra segmental acoustic indices was carried out for particular isolated words and continuous speech data. The isolated word data above 100 words uttered by each of 25 normal subjects and8Cleft palate/Lips disabled pathological subjects were analyzed and reference /threshold level was considered for each isolated word Various Misarticulation cases were studied and analyzed in case of pathological subjects. The spectrograms were studied for the purpose of Formant analysis. In case of fast uttered words or continuous sentences complex pronunciation errors occur and the speech intelligibility is very poor.

Considering the observations for % CQ variation, Tx variation and fo track variation observations the system is designed for confirmation of Cleft palate or VPD Disorder. Results for training system are shown in Table 2.

Results for testing of remaining 50 % samples are shown in Table 3 as % samples confirmed for Cleft palate or VPD Disordered speech and normal speech.

Observations

The existing speech enhancement algorithms like spectral subtraction do not help in enhancement of pathological speech . The pathological speech due to Cleft Palate suffers from following conditions.

•Breathing voice segments are audible in speech because the subjects are under stress when they speak .When the speakers are supposed to take pause in between utterances of two successive words generally the breathing voice segment is heard.

•fo mean is in the range 300 Hz to 500 Hz .Hence the loudness is very high.

•The speakers have to put more efforts and the utterances of different words are not appropriate. The labial consonants are replaced by nasal consonants.

•High Jitter levels in the range 25 % to 40 %

•Due to low HNR levels below the pathological threshold of 12 dB the speech indicates harshness.

Our Contribution to present work

The Cleft Palate lips or VPD disordered speech database is not available.

a. We got the database labeled by the doctors.

b. We have evaluated and analyzed the speech of the Cleft Palate lips or VPD disordered people with the help of few segmental an supra segmental acoustic indices like fo mean, Percentile glottal frequency fo track Characteristics, Laryngeal Quality represented by % Close Quotient characteristics , All 'Tx'-Regular 'Tx' time-frequency Histogram and % voice Regularity.

c. Evaluation and confirmation of Cleft Palate lips or VPD disordered speech using the present theory is done for the first time by us and it is not done by any one before.

Concluding Remarks

The Cleft palate disorder is identified by evaluation of speech of 25 normal and 10 cleft palate pathological subjects .Out of 10 cleft palate subjects 8 were confirmed to be cleft palate/lips disabled .The cleft palate/lips disorder can be evaluated on the basis of following segmental and supra segmental acoustic indices.

• The Cleft palate disorder is always indicated by hyper nasality exhibited in the disabled speech data .The regular Tx pulses are produced in 120 Hz to 600 Hz spectrum mostly on higher side of the spectrum. This is evident because the nasal cavity is more active than oral cavity and often voice resonates in nasal cavity.

• It is observed that if the same speaker utters the same isolated word data then also the acoustic indices like Jitter ,%CQ, voice regularity and spread of Formant frequencies show at least 20% deviation for the speech samples in case of Cleft palate disabled speakers.

• Percentile fo Track Characteristics change the piecewise linear or nonlinear nature with variable gradients Nonlinear nature with high gradient.

• When the breathing voice segments are attenuated or replaced by zero segments, the intelligibility of the disabled speech is observed to be improved.

We have designed Vocal Tract Filter and tried to improve the acoustic Indices of the Cleft palate disordered speech. The improvements in acoustic indices are as follows.

1. Percentile Glottal frequency fo characteristics indicate smooth approximately linear response instead of nonlinear curved response.

2. % CQ curve indicates smoothening effect

3. % All Tx and % regular Tx graphs are more spread with respect to time duration and frequency spectrum of fo variations.

Acknowledgments

The authors would like to thank the participants of research work- all the elderly people with various disabilities, speech disorders and without speech disorders for their extended cooperation and support for the completion of the work. The authors would also like to thank the family members and caretakers of all the participants for their patience and support while acquiring the speech data. The authors would also like to thank Dr. Mrs. Shilpa Hujurbajar Director of "Sanvad Speech and Hearing Therapy Centre, Kolhapur" for authenticating our work

References

[1] Ann W. Kummer, Linda Lee "Evaluation and Treatment of Resonance Disorders" Language, Speech, and Hearing Services in Schools Vol. 27 July 1996 0161-1461/96/2703-0271 © American Speech-Language-Hearing Association

[2] Rainsonata" Phonology-Cleft Palate" Flashcards March 2014

[3] Ann W. Kummer, PhD, CCC-SLP "Resonance Disorders and Velopharyngeal Dysfunction: Evaluation and Treatment" March 2010 Cincinnati Children's Hospital Medical Center

[4] Ann W. Kummer, PhD, CCC-SLP "Speech Therapy for Cleft Palate or Velopharyngeal Dysfunction (VPD)" 2008 Cincinnati Children's Hospital Medical Center

[5] Michael A. Berger "Measurement of vowel nasalization by multi-dimensional acoustic analysis" Department of Linguistics University of Rochester, New York 2007

[6]"Human Communication Disorders" Topic 7 Speech Disorders Motor Speech Disorders CSD 2230

[7] James F. Lubker, Kenneth Moll"Silultaneous Oral-Nasal Air Flow Measurements and Cinefluorographic Observations During Speech Production" 1963 Convention of the American Speech and Hearing Association,Chicago,Illinois
[8] Philip Loizou**,"** COLEA: A MATLAB software tool for speech analysis" Department of Electrical Engineering University of Texas at Dallas Copyright © 1998-99
[9]Pascal van Lieshout "PRAAT Short Tutorial" October 2003 University of Toronto, Graduate Department of Speech-Language Pathology, Faculty of Medicine
[10] SFS Manual Queen's College, London
[11] Zoran C´ irovic´, Milan Milosavljevic "Multimodal Speaker Verification Based on Electroglottograph Signal and Glottal Activity Detection" Hindawi Publishing Corporation EURASIP Journal on Advances in Signal Processing Volume 2010, Article ID 930376, 8 pages doi:10.1155/2010/930376

Biography

[1]Mrs. Manasi Dixit is working as Associate Professor in KIT's College of Engineering, Kolhapur .Her teaching experience is 30 years. Her main fields of interest are Digital Signal Processing, Speech Processing, Image Processing and Microwave Engineering. 28 PG students have completed their research work and have been awarded M.E.(E&TC) degree under her guidance. She has published 20 papers in reputed international journals. She has worked in the capacity of the SENATE member Shivaji University, Kolhapur and BOS-Board of Studies Member for Electronics and Telecommunication Engineering, Shivaji University, Kolhapur.

[2] Prof. Dr. Shaila Dinkar Apte is currently working as a professor on PG side in Rajarshi Shahu College of Engineering, Pune and as reviewer for the International Journal of Speech Technology by Springer Publication, International Journal of Digital Signal Processing, Elsevier Publication. She is currently guiding 8 Ph.D. candidates. Five candidates have completed their Ph.D. under her guidance. About 54 candidates completed their M.E. dissertations under her guidance. Almost all dissertations are in the area of signal processing. She has a vast teaching experience of 37 years in electronics engineering, and enjoys great popularity amongst students. She has been teaching Digital Signal Processing and Advanced Digital Signal Processing since last 18 years. Her previous designations include being an Assistant Professor in Walchand College of Engineering, Sangli, for 27 years; a member of board of studies for Shivaji University and a principle investigator for a research project sponsored by ARDE, New Delhi. She has published 32 papers in reputed international journals, more than 40 papers in international conferences and about 15 papers in national conferences. She has a patent published to her credit related to generation of mother wavelet from speech signal. Her books titled "Digital Signal Processing" and "Speech and Audio Processing" are published by Wiley India. A second reprint of second edition of the first book is in the market. The book titled "Advanced Digital Signal Processing" is published in June 2013 by Wiley publishers.