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

Electronics Engineering

Elixir Elec. Engg. 84 (2015) 33660-33669

ISSN: 2229-712X

# Evaluation of Neurogenic Speech Disorder-Dysarthria 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: 5 June 2015;

Received in revised form:

2 July 2015;

Accepted: 7 July 2015;

Keywords

Normal and Dysarthria 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

In this work the speech disorder namely Dysarthria is discussed. The basics of the disorder are described here. The speech parameters are identified which characterize 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 Speech Pause Index, 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

Dysarthria is described as Motor or Neurogenic Speech Disorder. Dysarthria is an impairment of articulation caused by damage to the nerve centres or tracts immediately involved in direct control of the muscles used in the enunciation and pronunciation of vowels and consonants. Dysarthria is often associated with irregular respiration, phonation and amplitude variations, articulation, in coordination of articulators, restricted movement of articulators, improper resonance and prosody [4]. It is described as a group of impairments which affects the speed, direction, strength, and timing of motor movements and causes improper articulation. It results from paralysis, weakness, or disco ordination of speech muscles. Dysarthria causes difficulties in motor control to produce speech.

Motor speech disorders are characterized by following types.

**i) Motor Planning/Programming Disorders**: It is described as inability to group and sequence the relevant muscles with respect to each other. The disorder comprises of impairments of Planning/Programming leading to disrupted coordination of relevant muscles and muscle groups where as muscle physiology and movement is intact. The Apraxia of Speech (AOS) is Motor Planning /Programming Disorder.

**ii) Motor Execution Disorders:** The disorder is due to deficits in physiology and movement abilities of muscles. It leads to impairments of execution because of disruptions in muscle physiology affected by involuntary movements and reductions in movement abilities while speaking. Dysarthria is Motor Execution Disorder.

**Types of Dysarthria – acquired and developmental**

**Acquired Dysarthria**

It indicates disruption in the execution of speech movements resulting from neuromuscular disturbances to muscle tone, reflexes, and kinematic aspects of movement. Speech sounds slow, slurred, harsh or quiet, or uneven depending on the type of Dysarthria. Dysarthria can be caused by different conditions including a progressive disease or trauma(TBI), paralysis, cerebral palsy, multiple sclerosis, Parkinson's disease, Alzheimer's etc[1]. The particular pattern of speech impairment is determined by the degree to which these muscle groups are affected due to the progressive degenerative disease.

Developmental Dysarthria:-Present at birth. Usually occurs along with known disturbance to neuromotor functioning. It can be caused by pre, peri or post-natal damage to the nervous system.

## Literature Review

When speech is affected, all aspects of speech production may be affected including respiration, phonation, resonation, articulation, and prosody [2]. It is, in part, characterized by pronunciation errors that include deletions, substitutions, insertions, and distortions of phonemes [3]. These errors follow consistent intra-speaker patterns that authors have exploited through acoustic and lexical model adaptation to improve automatic speech recognition (ASR) on dysarthric speech [2]. Kinfe Tadesse Mengistu et.al. have shown that acoustic model adaptation yields reduced average relative word error rate (WER) [4]. Word error rate (WER) for dysarthric speech was shown to be significantly higher than for normal speech.

Tiago H. Falk et. al. [5] have worked on speech disorder due to cerebral palsy called spastic Dysarthria. Other issues related to cerebral palsy are breathing problems, Muscle Rigidity, Velopharyngeal incompetency, uncoordinated articulator movements, Intellectual processing, auditory processing and Language impairments. Dysarthria is not a

Tele:

E-mail addresses: manasidixit64@gmail.com

language disorder but rather a difficulty in motor speech control. intelligibility assessment for speakers with mild or moderate Dysarthria [6].Authors have shown that articulation errors are the major contributing factor to reduced intelligibility in dysarthric speech followed by prosody, voice quality, and nasality. Representative symptoms of imprecise placement of articulators can include prolonged phonemes, unclear distinction between adjacent phonemes, odd speech rates, and rhythmic disturbances, to name a few.

Oscar Saz et.al. have presented [7] the results in the analysis of the acoustic features namely formants and the three supra segmental features viz tone, intensity and duration of the vowel production.

Santiago Omar et.al. discussed [8] about modeling errors in ASR for Dysarthric Speakers. Although automatic speech recognition (ASR) systems have been developed for disordered speech, factors such as low intelligibility and limited phonemic response decrease speech recognition accuracy, resulting into ASR systems performing poorly on dysarthric speakers.

Lingyun Gu et. Al. [9] discussed assessment of disordered dysarthric speech using automatic methods based on quantitative measures, objective measure such as dynamic time warping (DTW). M. Hafidz et.al. [10] have designed and implemented the speech recognition system for Cerebral Palsy using MATLB (GUI).. W.A. Simmt et.al.[11] have proposed possible measures for analysis of speech. Therapists currently exploit speech rate and the ratio between durations of utterances and silences. The second measure, a rate / gaps display to present this information automatically, is discussed here.

Dysarthria has particular effects on the prosody of the speech [11]. Prosody is the study of characteristics of speech beyond the basic words and spectra, e.g. stress, intonation, temporal variations etc. Speech Language Therapists (SLT) currently make use of rate/gap observations. Their measurement is often done by a manual timing of passages of speech. The measurement of the ongoing ratio between durations of utterances and silences is also valued by the SLTs Bon K. Sy et.al. [12] have perceived Dysarthric speech as articulatory error patterns compared with normal speech. These articulatory error patterns result in unintelligible or partially intelligible speech.

**Identification of Motor Speech Disorders**

**Assessment of Motor Speech Disorders**

The Assessment Process:-Professionals consider how the disorder affects the individual's life to determine the impairment and the course for treatment. Assessment of motor speech disorders should include measures of speech and non speech oral motor skills and should isolate particular motor subsystems to determine impairment.

**Measurement Methods**

**Perceptual measures** – This method includes perceptual judgments of speech intelligibility, accuracy, and rate of speech production.

**Acoustic measures** – This method includes spectrogram, formant analysis, Segmental Analysis and Supra segmental Analysis for detection of pathological speech and Classification of speech disorder as AOS or Dysarthria etc.

**Comprehensive Motor Speech Evaluation**

It involves motor control tasks that consider speech and non speech motor activities. Evaluation includes assessment of each of the subsystems separately– respiration, phonation, resonation, articulation and also prosody [10]. Most of the speakers involved in this studies presented individual error patterns, and variability is observed in the range of Diagnostic Markers.

**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.

**System Implementation**

The implemented system includes following processes.

**Procedure**

The present work attempts to find few parameters from pathological speech for confirmation of Dysarthria of speech disorder. 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 speech uttered by children and some adult male and female subjects speaking in 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 20 to 50 years is collected.

**Database**

The standard database is available but it is very costly. The speech data of 12 Dysarthric 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 [13] freeware 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 Dysarthria of speech disordered data.

.**Evaluation of Speech**

The present work attempts to confirm the Dysarthria of speech 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 $f_o$, 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. This is done with reference to the parameter threshold ranges specified in Table 1. Fundamental Frequency Analysis- The $f_o$ mean value lies in the high range between 180 Hz to 440 Hz for adult male female speakers and children

2. Glottal frequency ($f_o$) variations – the glottal frequency $f_o$ variations with respect to % percentile $f_o$ values from 0% to 100%. The Linear or nonlinear nature of the characteristics is important. Dysarthria is confirmed by Staircase Pattern of variable low – high-- low- high gradient of the $f_o$ characteristics in 0% to 100% $f_o$ percentile range.

3. % Close Quotient (CQ) graph simulates the Laryngograph and indicates the close phase of Glottis pulse signal or vocal folds vibration cycle. The mean CQ, range of variation of CQ and CQ variations with respect to total time duration of speech sample are important parameters.

4. All 'Tx' graph indicates histogram of all pitch cycles for the total time duration of speech sample. Regular 'Tx' graph 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 $f_o$ value of the person are important parameters.

**System Block Diagram**

Block Diagram for Evaluation of Diagnostic Markers in Case of Dysarthria is shown in Figure 2.

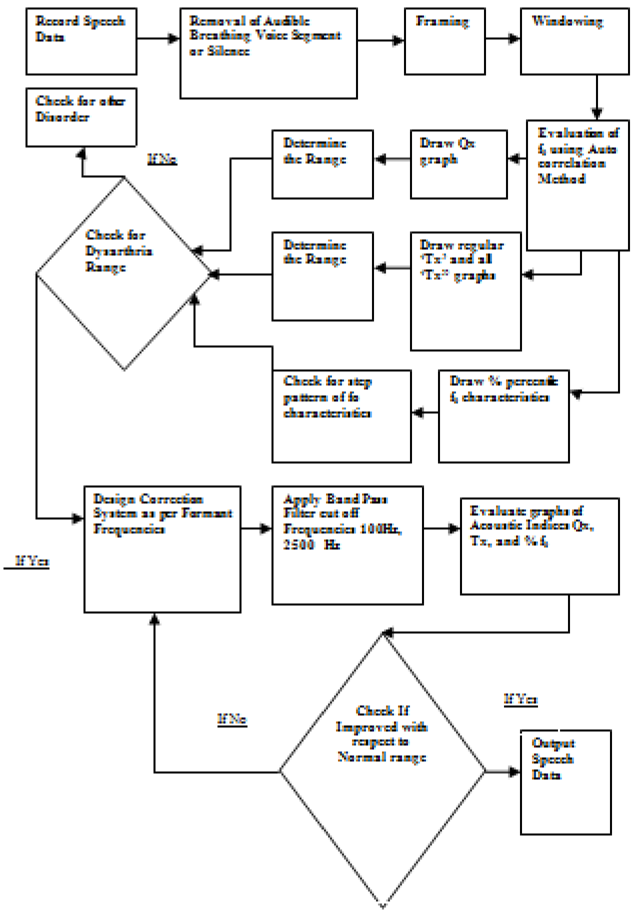


**Figure 1. Block Diagram for Evaluation of Dysarthria**

The present work attempts to confirm the Dysarthria of Speech disorder from the speech signal by extracting various segmental and supra segmental acoustic indices. The acoustic indices are evaluated for all isolated word and 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' [14] 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 like voice regularity, Jitter, Shimmer, $f_0$ mean mode values. The SFS 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 [15]. 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 [15]. It is defined as Regular pitch period if it varies in duration by less than +/- 10% with respect to the adjacent pitch periods [15]. Using the acoustic indices the relationship between the Electroglottograpgh (EGG) measures [16] 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. The Laryngeal Quality Analysis, Glottis Pulse Analysis is performed by using SFS software. SFS reports indicate exactly similar 'all Tx' and 'regular Tx' glottis pulse graphs in case of normal subjects where as for pathological subjects the 'regular Tx' pulses produced are very low, insufficient in time domain as well as variable in frequency domain.

**System Development**

The system is developed operates in 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 periods indicative of the close phase of Glottis pulse signal or vocal folds vibration 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 Dysarthric characteristics. The observations for Tx are as follows.

1. Regular Tx is defined as Regular pitch period [15] if it varies in duration by less than +/- 10% with respect to the adjacent pitch periods In case of normal persons, regular Tx graph matches with all Tx graph.

2. For Dysarthric persons, regular Tx graph contains almost 50 % of frequency range as compared to all Tx.

We have developed a MATLAB routine to evaluate fundamental frequency $f_0$ using Autocorrelation Technique. Percentile Glottal signal frequency graphs are drawn in Percentile Domain after adopting Normalization Procedure. The Microsoft Excel Software is also used.

The algorithm for evaluation of $f_o$ track is as follows.

1. The % percentile value of $f_0$ maximum level is considered to be 100% percentile.

2. Hence according to the data values of $f_0$ variations as per the speech sample 0%-5%-10%-15%-----95%-100% percentile values of every speech sample are calculated.

3. The $f_0$ 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 Dysarthria. It was observed that the % percentile $f_0$ track graphs are having a high- low- high –low gradient step pattern of 3 to 6 steps in 0% to 100 % percentile $f_o$ range.

For normal speech % percentile $f_o$ track graphs are linear from 5% to 95 % range with very low gradient 0.03 to 0.3.

The observations for Closed Quotient (CQ) are as follows.

1. CQ is a time Vs CQ graph. It indicates that the CQ values in the range 15 % to 75 % are present for Dysarthric speakers.

For normal speech, CQ variation is observed in the range 15 % to 60 % only.

**Graphs of the Diagnostic Markers**

Figures 2 to 4 show the variations of diagnostic markers for disordered speaker 1 named Kashinath. Figure 2 shows the %Tx graph for disordered speaker 1. It can be observed that the regular Tx pulses are produced in the 250 Hz to 280 Hz pitch spectrum exactly matching with All Tx waveform but it is not at all produced in other frequency range. The time duration of regular Tx pulses is almost 80% of All Tx graph.

**Diagnostic Markers- speaker 1 - Kashinath_1b**

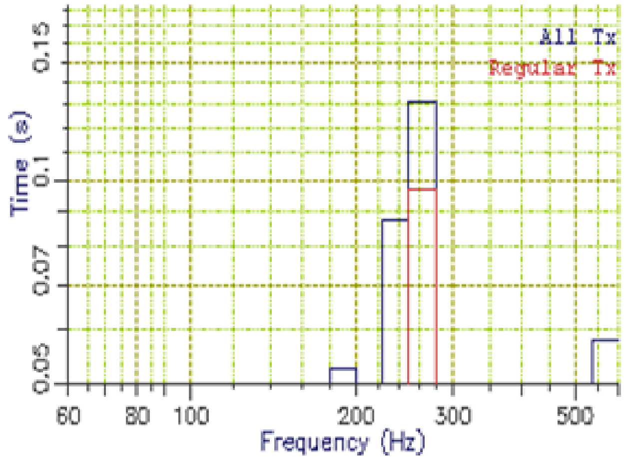


**Figure 2. % Tx Glottal Pulses-Speaker1-Kashinath_1b**

Figure 3 shows %CQ graph for the same speaker. It can be seen that % Qx graph indicates 34 % to 57 % wide CQ range with mean value %CQ = 45.2.

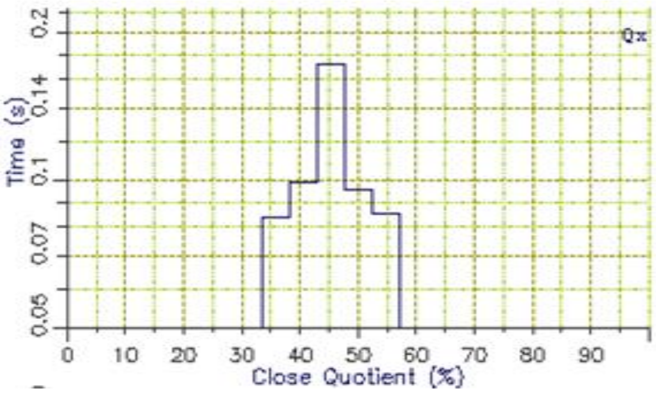


**Figure 3. % Close Quotient wrt Time Kashinath_1b**

Figure 4 shows the percentile $f_o$ track variation for the same speaker. It indicates that percentile $f_o$ track indicates gradient index varying in a step pattern. The slopes for different steps are as follows. 0 to 10 percentile range – 1.62 –High, 15 to 70 percentile range-0.07 Low , 75 to 95 percentile range-1.44 – High, 95 to 100 percentile range -0.06-Low

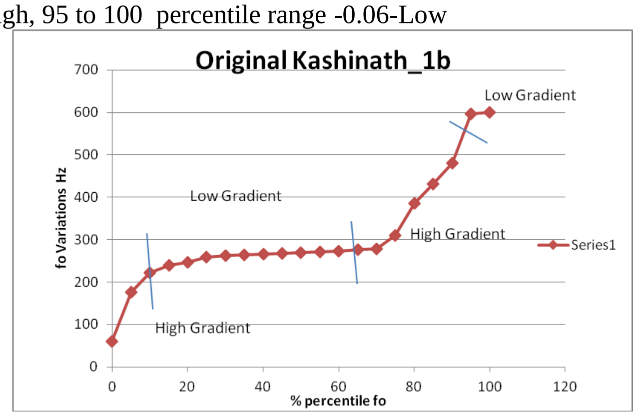


**Figure 4 . Percentile $f_o$ track variations Kashinath_1b**

Figures 5 to 7 show the graphs for diagnostic marker for disordered speaker 2 named Kiran. Figure 5 shows the %Tx graph for disordered speaker 2. The regular Tx pulses are produced in the 140 Hz to 170 Hz pitch spectrum exactly matching with All Tx waveform but it is not at all produced in other frequency range. The time duration of regular Tx pulses is almost 90% of All Tx graph when they coexist.

**Diagnostic Markers- Kiran_1a**

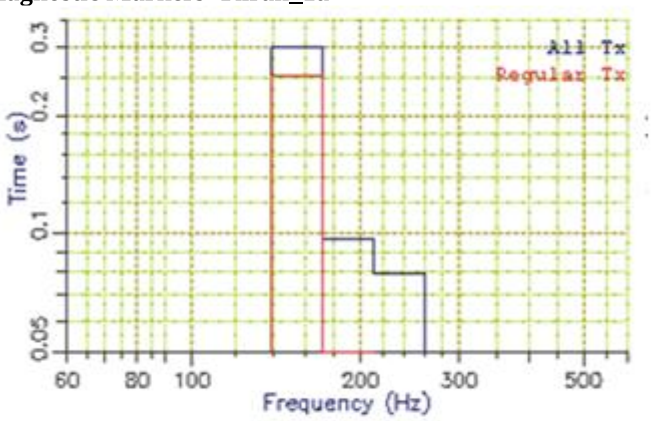


**Figure 5. % Tx Glottal Pulses-Speaker1-Kiran_1a**

Figure 6 shows the % Qx graph for the same speaker. % Qx graph indicates 27 % to 73 % wide Qx range with mean value %CQ – 47.6.

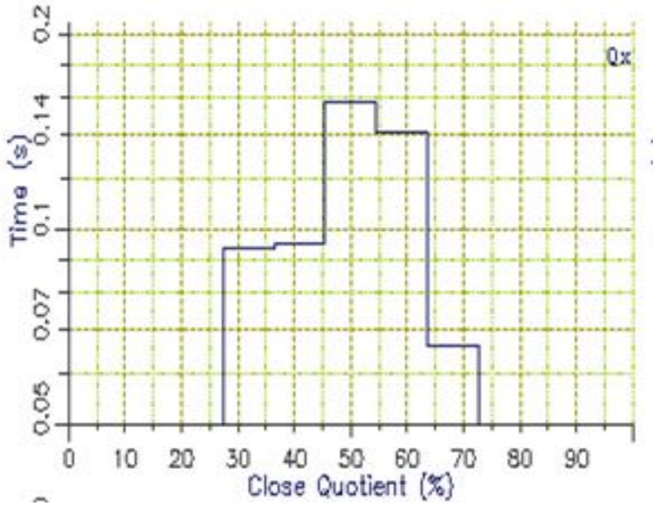


**Figure 6. % Close Quotient wrt Time Kiran_1a**

Figure 7 shows percentile $f_0$ track variations. Percentile $f_o$ track indicates gradient index varying in a step pattern as follows. 0 to 5 percentile range – 1.62 –High, 5 to 70 percentile range-0.05 Low,70 to 75 percentile range-0.72 –High, 75 to 80 percentile range -0.05-Low ,80 to 85 percentile range-0.91 – High, 85 to 95 percentile range -0.05-Low 95 to 100 percentile range -1.2-High

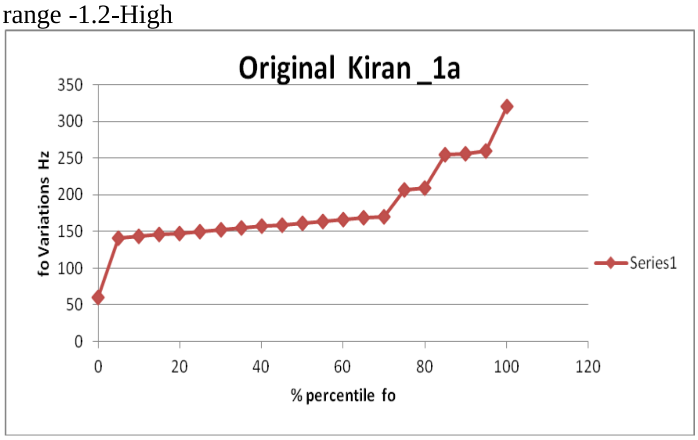


**Figure 7. Percentile $f_o$ track variations Kiran_1b**

Figures 8 to 10 show the graphs for diagnostic marker for disordered speaker 3 named Bipin Patil. Figure 8 shows the %Tx graph for disordered speaker 3. The regular Tx pulses are produced in the 158 Hz to 200 Hz pitch spectrum exactly matching with All Tx waveform but it is not at all produced in other frequency range. The time duration of regular Tx pulses is almost 80% of All Tx graph when they coexist.

**Diagnostic Markers- Bipin Patil 1**

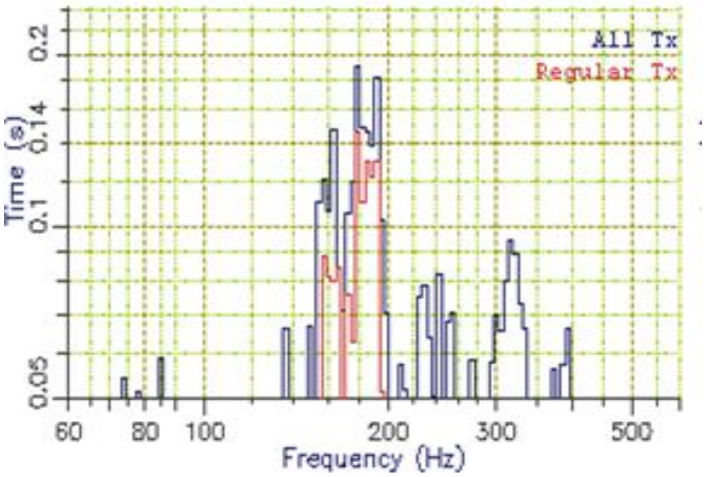


**Figure 8. % Tx Glottal Pulses-Speaker3-Bipin Patil 1**

Figure 9 shows the % Qx graph for the same speaker. % Qx graph indicates 25 % to 62 % wide Qx range with mean value %CQ is 49.7.

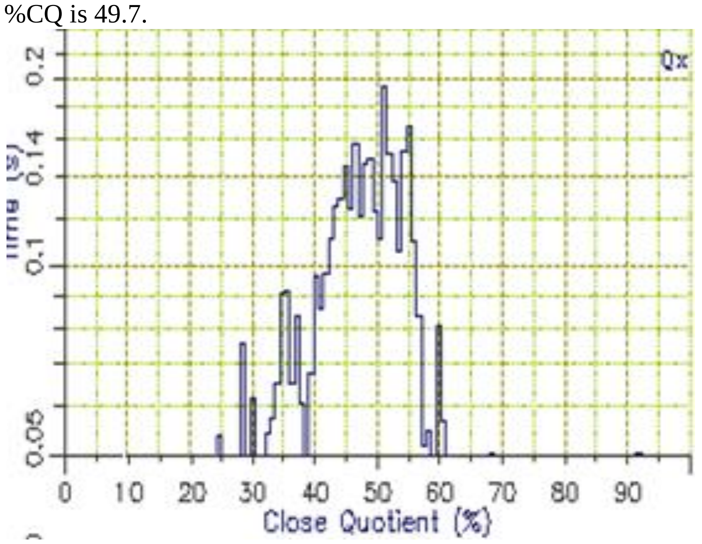


**Figure 9. % Close Quotient wrt Time Bipin Patil1**

Figure 10 shows percentile $f_0$ track variations. Percentile $f_o$ track indicates gradient index varying in a step pattern as follows. 0 to 5 percentile range – 1.75 –High, 5 to 55 percentile range-0.29 Low, 60 to 90 percentile range-0.46 –High, 95 to 100 percentile range -4.07-High

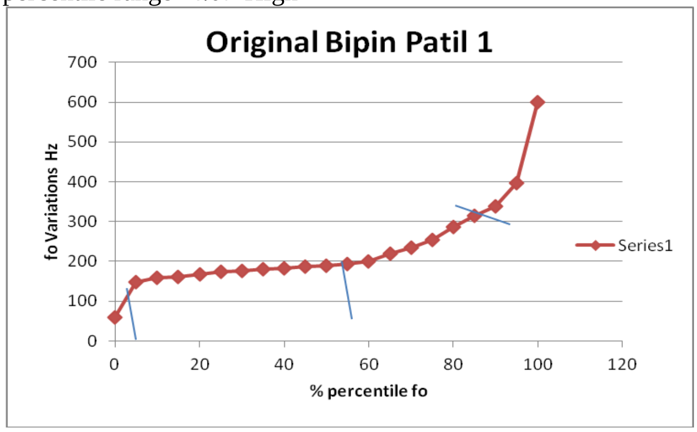


**Figure 10. Percentile fo track variations**

Figures 11 to 13 show the graphs for diagnostic marker for disordered speaker 4 named Ashish. Regular Tx pulses (regular pitch periods) are produced between 190 Hz to 240 Hz as an impulse at time instant 0.05 sec.

**Diagnostic Markers- Aashish Salvekar 14**

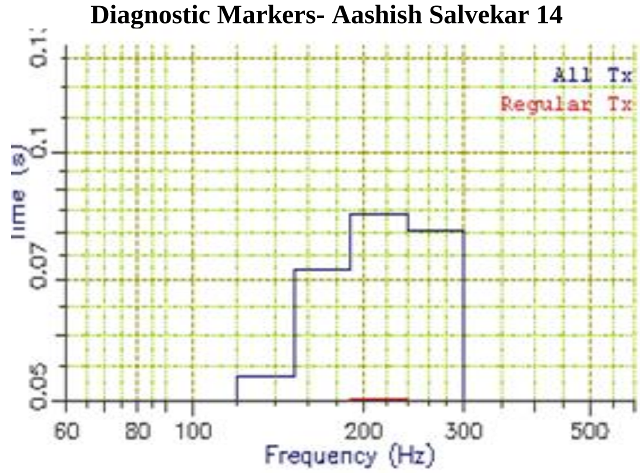


**Figure 11. % Regular Glottal Pulses - Aashish Salvekar 14**

Figure 12 shows % CQ variation. % Qx graph indicates 20 % to 60 % wide Qx range with mean value %CQ is 49.7.

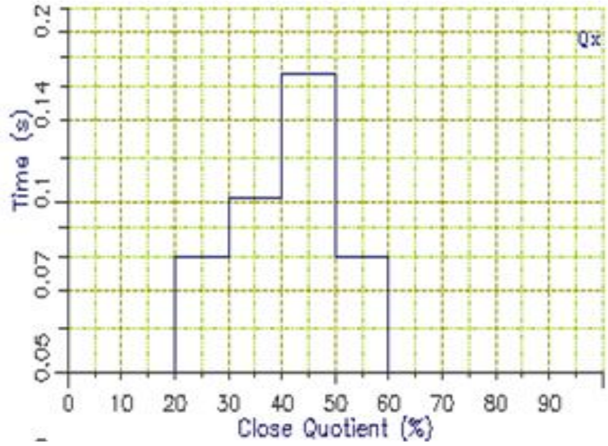


**Figure 12. % Close Quotient wrt Time- Aashish Salvekar 14**

Figure 13 shows percentile $f_0$ variations. Percentile fo track indicates gradient index varying in a step pattern as follows.

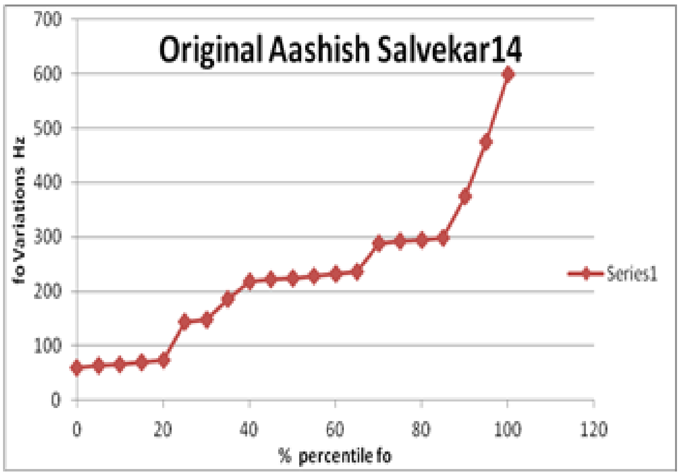


**Figure 13. Percentile fo track variations**

0 to 20 percentile range – 0.07 –Low, 20 to 25 percentile range-1.4 High, 25 to 30 percentile range-0.09 –Low, 30 to 40 percentile range -0.7-High , 40 to 65 percentile range-0.07 –Low, 65 to 70 percentile range -1.04-High , 70 to 85 percentile range -0.07-Low, 70 to 85 percentile range -2.0-High

Figures 14 to 16 show the graphs for diagnostic marker for normal speaker 5. Figure 14 shows the %Tx graph for disordered speaker 5. Regular Tx pulses (regular pitch periods) are produced between 80 Hz to 180 Hz almost similar to All Tx graph.

**Diagnostic Markers- First (Normal speech)**

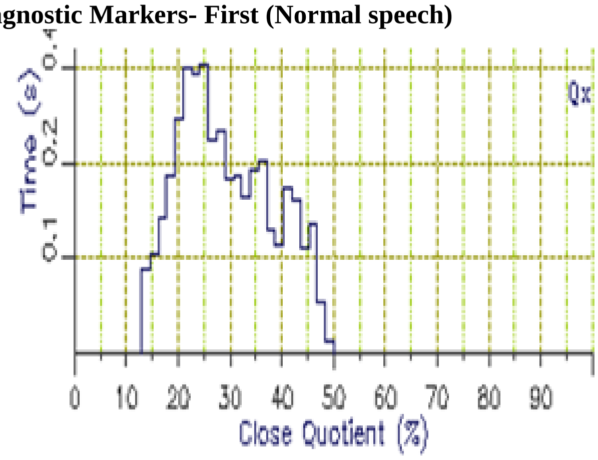


**Figure 14. % Regular Glottal Pulses –First Normal**

Figure 15 shows the %Qx graph for the normal speaker. % Qx graph indicates 13% to 50% wide Qx range.

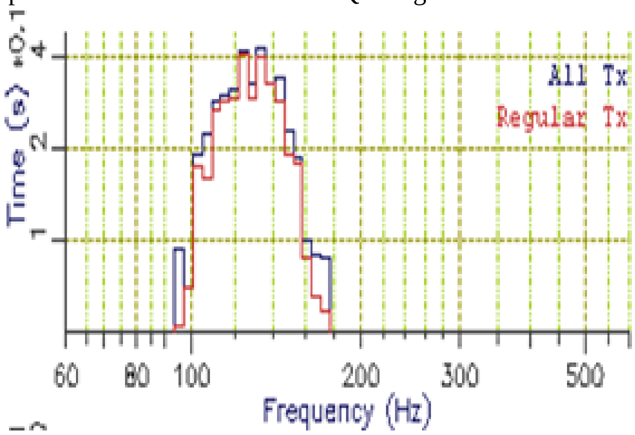


**Figure 15. % Close Quotient wrt Time-First Normal**

Figure 16 shows percentile $f_0$ variations for normal speaker. Percentile fo track indicates linear very low gradient index as follows. 5 to 95 percentile range – 0.07 –Very Low

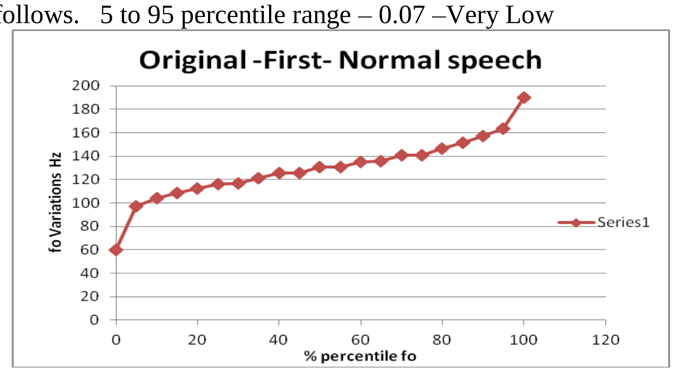


**Figure 16. Percentile fo track variations**

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

1. It is observed that Laryngograph comprising of regular Tx pulses contains less than 50 % of the frequency spectrum as compared to all Tx.
2. Time periods Vs closed quotient graph indicates closed quotient range more than 50 % wide with mean %CQ within 42% to 48%.
3. Percentile $f_o$ track indicates gradient index varying in a Staircase pattern with 3 to 6 steps as follows. Every individual has shown different step pattern varying in following range.

0 to 20 percentile range – 0.05 to 0.09 –Low
20 to 40 percentile range-0.8 to 1.4 High
45 to 65 percentile range-0.05 to 0.09 –Low
70 to 80 percentile range -0.7 to 1.8-High
85 to 90 percentile range-0.05 to 0.09 –Low
95 to 100 percentile range -1.0 to 2-High

However as the severity of the disorder increases the $f_o$ track exhibits very high gradient and nonlinear nature throughout the range

**System Developed for Correction**

We have used MATLAB based developmental tool E-System compatible with SFS developmental software for trying methods for correcting Dysarthric 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 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
- 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. The specifications for the systems are given below.

Band Pass Filter- The lower cut off frequency is in the range 10 Hz to 100 Hz and 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 Dysarthria of Speech Disorder The performance of Band Pass Filter was found to be better in comparison with Resonator Filter and Vocal Tract Filter during the Training mode. Hence the Band Pass Filter is applied for correction of Dysarthric speech. The Band Pass Filter is designed with the lower cut off frequency is in the range 10 Hz to 100 Hz and 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 the categories adult male, female, children or elderly people based on their respective pitch frequency range.

After applying the Band Pass Filter to Dysarthric Speech the following observations are made.

4. It is observed that Laryngograph comprising of regular Tx pulses graph shows small improvement with respect to time duration and frequency spectrum of pitch.

5. Time periods Vs closed quotient graph indicates closed quotient range 10 % more wide with BPF processed with respect to original with mean %CQ within 41% to 45% with 5% improvement.

6. Percentile f0 track has a low gradient (slope)- In 0 to 35 percentile range – 0.17 to 0.35 and a high gradient (slope) in 40 to 95 percentile range –0.45 to 1.4 with approximately linear pattern instead of step pattern .

Improvement in the diagnostic markers due to application of Band Pass Filter is indicated below with the help of Regular Tx graph, % Close Quotient (CQ) graph and percentile $f_o$ track graphs through Figure 17 to Figure 22.

**Diagnostic Markers -Aashish Salvekar14-processed using BPF-Dysarthria**

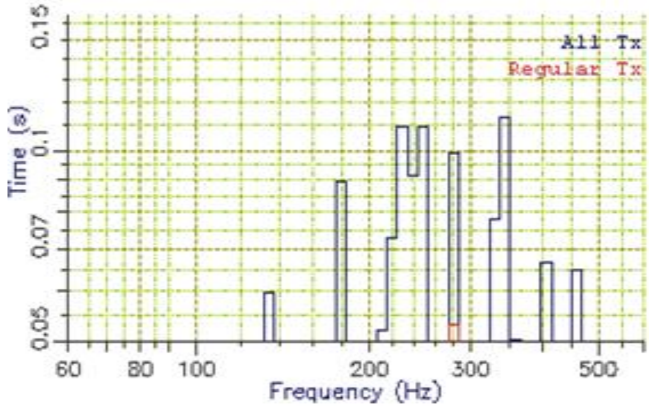


**Figure 17. % Regular Glottal Pulses processed by BPF**

Note

Regular Tx pulses (regular pitch periods) are produced between 275 Hz to 290 Hz during small time duration 0.05 sec to 0.055 sec

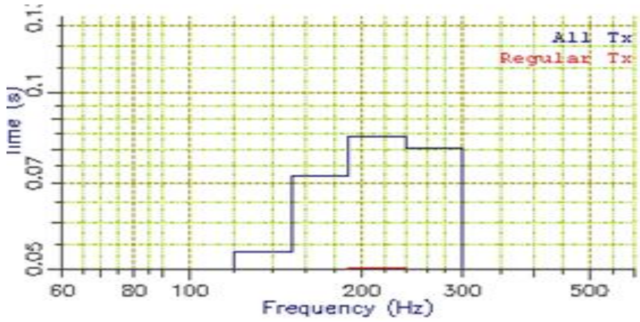


**Figure 18. % Regular Glottal Pulses Original**

Regular Tx pulses (regular pitch periods) are produced between 190 Hz to 240 Hz as an impulse at time instant 0.05 sec. It is observed that the % Regular Glottal Pulses Tx graph processed by BPF is in 275 Hz to 290 Hz during small time duration 0.05 sec to 0.055 sec as compared with Regular Tx pulses (regular pitch periods) produced between 190 Hz to 240 Hz as an impulse at time instant 0.05 sec in original.

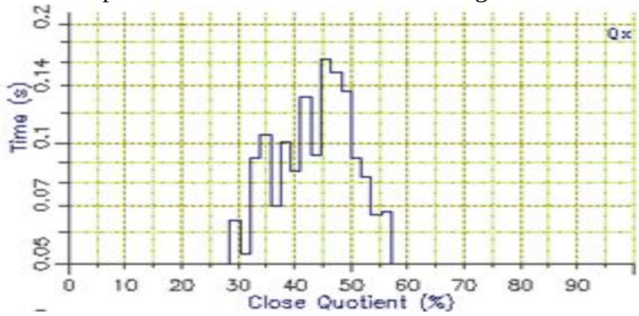


**Figure 19. % Close Quotient Vs Time Original**

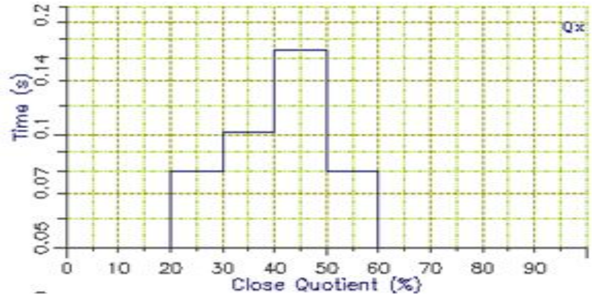


**Figure 20. % Close Quotient Vs Time periods Processed by BPF**

Note that the % Close Quotient (CQ) graph shows 28 % to 58% wide pulse in original and 20% to 60 % wide smooth pulse when processed through BPF.

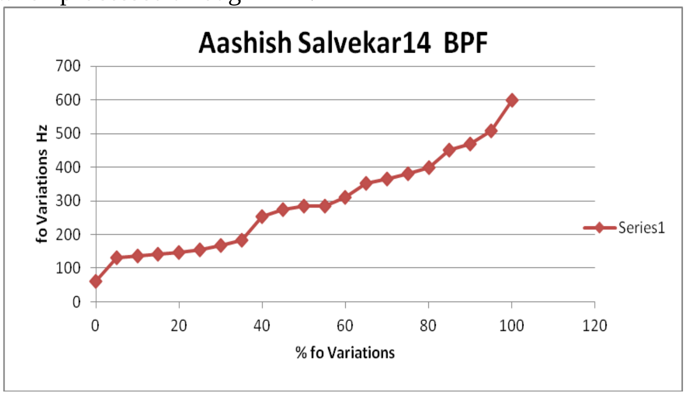


**Figure 21. Percentile $f_o$ track variations –BPF Processed**

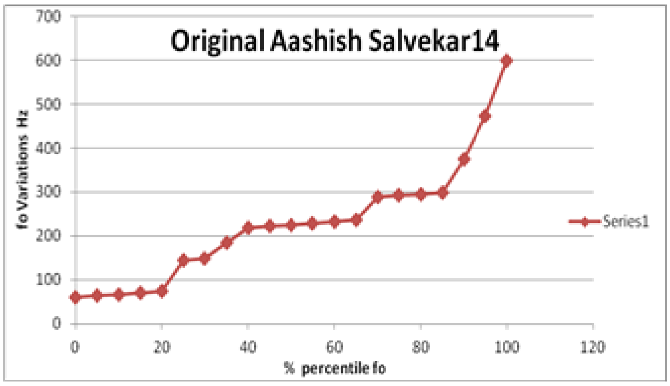


**Figure 22. Percentile $f_o$ track variations –original**

Note that percentile glottal frequency fo graph indicates gradient index improved by BPF as indicated by reduced gradient levels and extended percentile levels as shown below. 0 to 20 $f_o$ percentile range with low gradient 0.68 is extended to 0 to 35 $f_o$ percentile range with improved low gradient 0.35. The step pattern is improved and it is approximately linear in 40 to 95 $f_o$ percentile range with low gradient 0.47.

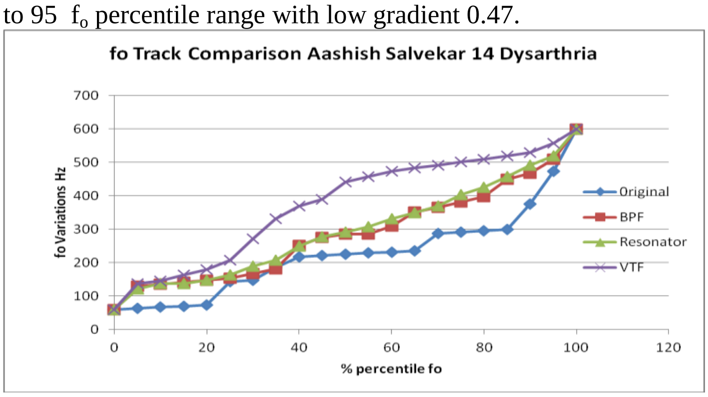


**Figure 23. Comparative Response of the Correction System for Dysarthric Speech**

**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 |

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

**Table 2. range of diagnostic markers for 50 % of Dysarthric speech samples and 50 % of normal speech**

| Diagnostic marker | Range of values for Dysarthric speakers | Range of values for normal speech |
|---|---|---|
| Time Vs % CQ | 15 % to 75 % | 10 % to 60 % |
| Time Vs frequency Tx graph | Very poor Regular Tx graph for less than 30 % frequency range and for very small time periods as compared to all Tx graph | Regular Tx matches with all Tx for more than 90 %. |
| Percentile fo graph variation | Percentile fo graph indicates Staircase pattern with low-high-low-high … gradients with 3 to 6 steps | For 5 to 95 percentile range – Linear with gradient in 0.05 to 0.3 range |

Results for testing of remaining 50 % samples are shown in Table 3 as % samples confirmed for Dysarthric or normal speakers.

**Table 3. % confirmation for 50 % of Dysarthric and normal speech samples**

| Parameter used | % samples confirmed for Dysarthric speech | % samples confirmed for normal speech |
|---|---|---|
| Time Vs frequency TX graph | 95% | 100 % |
| Percentile fo track variation | 100 % | 100 % |
| All 3 parameters % CQ, Tx, $f_o$ | 96 % | 100% |

**Table 4. Results of correction for Dysarthric speech after applying correction using BPF**

| Diagnostic marker | Range of values for Dysarthric speakers | Range of values for Dysarthric after applying correction using BPF |
|---|---|---|
| Time Vs % CQ | 15 % to 75 % | 15 % to 75 % graph is more smooth |
| Time Vs Frequency TX graph | Very poor Regular Tx graph for less than 30 % frequency range and for very small time periods as compared to all Tx graph | Regular Tx graph matches with All Tx graph for more than 60 %.<br>No intermittent pulses, graph is smooth and present for all time |
| Percentile Glottal Frequency fo variation | Percentile fo graph indicates Staircase pattern with low-high-low-high … gradients with 3 to 6 steps<br>Values vary for every subject | Initial Low Gradient range is improved -i.e. extended and Gradient values are improved-i.e. reduced<br>Overall smooth Characteristics |

The Figure 23 indicates that the BPF system with lower cut off frequency 10 Hz and upper cut off frequency 2500 Hz has improved the original step pattern and high gradient Dysarthric speech to extend initial low gradient 0.35 graph in 0 - 35 percentile $f_0$ range and the step pattern is improved and it is approximately linear in 40 to 95 $f_o$ percentile range with low gradient 0.47. Similar experimentation on other speech samples has proved that BPF is a better solution or correction system for Dysarthric speech.

**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 and12 Dysarthric disabled 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 $f_0$ track variation observations the system is designed for confirmation of Dysarthria. Results for training system are shown in Table 2.

**Observations**

The existing speech enhancement algorithms like spectral subtraction do not help in enhancement of pathological speech [17]. The pathological speech due to Dysarthria 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.
- The minimum intensity level does not drop much as there is no silence region due to the presence of breathing voice segments.
- The speakers have to put more efforts for the motor movements of articulators .Hence the utterances of different words is not appropriate.
- 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 Dysarthria disordered speech database is not available.

1. We got the database labeled by the doctors.
2. We have evaluated and analyzed the speech of the Dysarthria affected people with the help of few segmental an supra segmental acoustic indices like Percentile $f_o$ track Characteristics, Laryngeal Quality represented by % Close

3. Quotient characteristics, All 'Tx'-Regular 'Tx' time-frequency Histogram and % voice Regularity.
4. Evaluation and confirmation of Dysarthria 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 Dysarthria of Speech disorder can be identified by evaluation of speech of 25 normal and 12 pathological Dysarthric subjects containing 100 speech samples. We have confirmed the Dysarthria speech disorder with the help of the threshold values of the acoustic indices as described below.

1. Percentile $f_o$ characteristics are piecewise linear and indicate step pattern. Gradient index varies as follows.-

0 to 20 percentile range – 0.05 to 0.07 –Low,
20 to 25 percentile range-0.5 to 1.4 High,
25 to 30 percentile range-0.05 to 0.09 –Low,
30 to 40 percentile range -0.5 to 1.2-High,
40 to 65 percentile range-0.05 to 0.09 –Low,
65 to 70 percentile range -0.7 to 1.2-High,
70 to 85 percentile range -0.05 to 0.09-Low,
85 to 100 percentile range -2.0 and more -High.

Every Dysarthric speaker exhibits different step pattern in $f_o$ characteristics.

2. % Close Quotient (CQ) mean for Dysarthric speech lies in the range 40 % to 48% with range of variation 15% to 70% and the graph becomes more wide and smooth after being processed by BPF.
3. The indication of Glottis performance judged with the help of % Regular Tx and All Tx pulse waveforms show that Regular Tx pulses are produced for a very short frequency range and for a small time for Dysarthric speech.

**References**

[1] Prasad D. Polur and Gerald E. Miller "Experiments with Fast Fourier Transform, and Linear Predictive and Cepstral Coefficients in Dysarthric Speech Recognition Algorithms Using Hidden Markov Model" IEEE Transactions On Neural Systems And Rehabilitation Engineering, VOL.13,NO. 4, December 2005

[2] Santiago Omar Caballero Morales and Stephen J. Cox "Modeling Errors in Automatic Speech Recognition for Dysarthric Speakers" Hindawi Publishing Corporation EURASIP Journal on Advances in Signal Processing Volume 2009, Article ID 308340, 14 pages doi:10.1155/2009/308340

[3] Cees H. Taal, Richard C. Hendriks, Richard Heusdens, and Jesper Jensen "An Algorithm for Intelligibility Prediction of Time–Frequency Weighted Noisy Speech" IEEE Transactions on Audio, Speech and Language Processing, VOL. 19, NO. 7, September 2011

[4]Kinfe Tadesse Mengistu and Frank Rudzicz" Adapting Acoustic and Lexical Models to Dysarthric Speech" 978-1-4577-0539-7/11©2011 IEEE 4924 ICASSP 2011

[5]Tiago H. Falk, Richard Hummel, Wai-Yip Chan" Quantifying Perturbations In Temporal Dynamics For Automated Assessment Of Spastic Dysarthric Speech Intelligibility" 978-1-4577-0539-7/11©2011 IEEE 4480 ICASSP 2011

[6]Sridhar Krishna Nemala and Mounya Elhilali "A Joint Acoustic and Phonological Approach To Speech Intelligibility Assessment" 978-1-4244-4296-6/10/$25.00 ©2010 IEEE 4742 ICASSP 2010

[7]Oscar Saz,1 Javier Simon, W.-Ricardo Rodriguez et.al. "Analysis of Acoustic Features in Speakers with Cognitive Disorders and Speech Impairments" Hindawi Publishing Corporation EURASIP Journal on Advances in Signal Processing Volume 2009, Article ID 159234, 11 pages doi:10.1155/2009/159234

[8] Santiago-Omar, Caballero-Morales "Estimation of Phoneme-Specific HMM Topologies for the Automatic Recognition of Dysarthric Speech" Hindawi Publishing Corporation , Computational and Mathematical Methods in Medicine , Volume 2013, Article ID 297860,15 pages http://dx.doi.org/10.1155/2013/297860

[9] Lingyun Gu, John G. Harris,"Disordered Speech Assessment Using Automatic Methods Based on Quantitative Measures" EURASIP Journal on Applied Signal Processing 2005:9, 1400–1409 c_2005 Hindawi Publishing Corporation

[10] M. Hafidz M. J., S.A.R.Al-Haddad, Chee Kyun Ng "Speech Recognition System for Cerebral Palsy" Recent Researches in Computer Science, 2010 ISBN: 978-1-61804-019-0

[11] W.A. Simmt, P.E. Roberts and M.J. Joyce "Signal Processing For Use in the Assessment of Dysarthric Speech" 2005 ,The Nuclear and Biomedical Engineering Research Group, Department of Engineering, University of Lancaster, UK

[12] Bon K, Sy, David M, Horowitz "A Statistical Causal Model for the Assessment of Dysarthric Speech and the Utility of Computer-Based Speech Recognition" 1993

[13] Philip Loizou," COLEA: A MATLAB software tool for speech analysis" Department of Electrical Engineering University of Texas at Dallas Copyright © 1998-99

[14] Pascal van Lieshout "PRAAT Short Tutorial" October 2003 University of Toronto, Graduate Department of Speech-Language Pathology, Faculty of Medicine

[15] SFS Manual Free software developed by University of London

[16] 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

[17] Philipos C. Loizou, Gibak Kim "Reasons Why Current Speech Enhancement Algorithms Do Not Improve Speech Intelligibility and Suggested Solutions" IEEE Transactions on Audio, Speech and Language Processing, Vol 19, No 1, Jan 2011

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

[19] Justice "Communication Sciences and Disorders: An Introduction Motor Speech Disorders: Apraxia and Dysarthria" Pearson Education 2006

**Biography**

[1] Dr.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 18 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.