Parkinson Disease and Voice

J.Sujatha[#1], Dr. S.P.Rajagopalan[*2]
[#1] Research Scholar, Department of Computer Science, Vels University, Chennai, India
[*2] Professor, Department of Computer Science, GKM College of Eng. & Technology, Chennai, India

Abstract *— Parkinson's disease (PD) is a chronic and progressive neurological disorder that affects the motor system. The muscles to speak and sing from the diaphragm to the larynx are affected. This disease is more common in the elderly people though 5-10% cases can be found with an early onset of 20 years of age. It affects around 6.8 million people globally. Parkinson disease destroys the brain cells which affects the brain related activity. The changes in the brain related activity helps to diagnose Parkinson disease at a stage where vital brain cells are significantly damaged. Voice undergoes a change at an earlier stage before the time the brain cells undergoes changes. Diagnosis at an early stage helps in better and effective treatment and prevents damage of up to 60% neurons that control the movement in the brain. This paper discusses about vocal system and the attributes used in detecting Parkinson's disease.*

Keywords *— Dopaminergic, substantia nigra, dysarthria, resonation, articulation*

I. INTRODUCTION

Parkinson disease is a chronic (long term), progressive (progresses slowly) and degenerative (cannot be brought to normalcy) neurological disorder that affects the basal ganglia which is involved in selection, sequencing and implementation of movement. Parkinson Disease can have a profound effect on speech and voice. 89% of people with PD experience some type of speech problems [1]. The voice changes occur before the parts of the brain get affected. The affected person voice changes i.e., they either speak softly or have difficulty forming words. The muscles in the face, mouth, throat and neck are affected giving a husky voice. They find it hard to communicate emotions. The speech disorder is characterized by decreased loudness, reduced pitch, soft voice, less precise articulation, dysfluency, stuttering of words, vocal tremors, monotone in conversation, hoarse voice quality, rigidity and stiffness in laryngeal and ribcage muscle and rapid burst of speech. Dysfluency is the unexpected pause, hesitation, false start, repetition of sound and syllables.

II. PARKINSON'S DISEASE

Parkinson's disease is neurological disorder of central nervous system. It is the second largest neurological disorder that targets neurons that produce dopamine chemicals. The dopamine generating cells called dopaminergic neurons in substantia nigra part of the brain have died which is responsible for coordination and movement. The substantia nigra is a nucleus in the mid brain which is a part of basal ganglia. This causes motor deficits. Dysarthria is neuromuscular speech disorder that involves the speech subsystem. This speech disturbance is experienced by a person with Parkinson disease. Voice disturbance is the first and early indicator of Parkinson disease [2] which occurs before the brain cells are affected. Voice change in Parkinson has 2 components, the larynx becomes weak and secondly the voice becomes low. Max Little, head of Parkinson Voice Initiative performed his research proving that voice can detect Parkinson and it had 99% success rate [3]. The three main symptoms to be detected in the voice are vocal fold tremors, breathiness and weakness and fluctuation of jaw, tongue and lip.

The voice onset time (VOT) differs for PD and control. The opening and closing of mouths during pronunciation of word is smaller in PD patients due to the inefficiency of sequencing and implantation of voice gestures.

III. VOICE

Speech is the sound in pronouncing words or sentences to communicate a message verbally. It involves sound waves through the air to spontaneously communicate ideas and express thoughts. Speech is produced by coordinated actions of head, neck, chest and abdomen. For speech to occur, air must be forced up out of the lungs to the trachea, and into the vocal tract and involves tongue, lip, soft palate and mandible. Voice is the sound produced by humans and other vertebrates using the lungs and the vocal folds in the larynx or voice box. Voice is generated by airflow from the lungs as the vocal folds are brought close together. When air is pushed against the vocal folds with sufficient pressure, the vocal folds vibrate producing speech, which otherwise would have resulted in whisper. Voice is as unique as our fingerprint and defines one's personality, mood, and health. It is generated as a result of combined action of expiratory breath and muscular properties of vocal cord and modified by resonating cavities above vocal cord.

Though there are 5 vowels in English alphabet, 12 different vowel sounds are produced Ex: i in miss and mice produce different sound. These vowels make 38% of speech. The remaining 21 alphabets make up the consonants and forms 62% of speech.

IV. BASIC STRUCTURE OF VOCAL SYSTSEM

Speaking process involves 4 stages

1. Respiratory stage
2. Phonation stage
3. Resonation stage
4. Articulation stage

A. *Respiratory stage:*

This involves breathing and consists of 2 phases' inhalation and exhalation. Lungs are reservoir of air. During inhalation, air is taken into the lungs which compress the diaphragm. Diaphragm is muscle separating chest from abdomen. It is the floor of chest and roof of abdomen that gives pressure to the breath stream.

B. *Phonation stage:*

Voice is produced in speaking as the expiratory air stream from lungs goes up through the trachea or wind pipe to the larynx (voice box). Trachea is passage way where air going up from lungs. Larynx (Adams apple) is principal organ of phonation found at top of trachea. Vocal fold or Vocal cord is pair of bundle of muscles and cartilages that open and close at various degrees. Glottis is the name given to space between vocal fold.

There are two types of phonation:

(i) Voiceless phonation

The glottis and the vocal fold are wide open allowing the air to escape and there is no vibration. This is called voiceless phonation.

(ii) Voiced phonation

Glottis and vocal fold are closed. Air from the lungs creates pressure behind glottis and the vocal fold separate and vibrate. This is called voiced phonation.

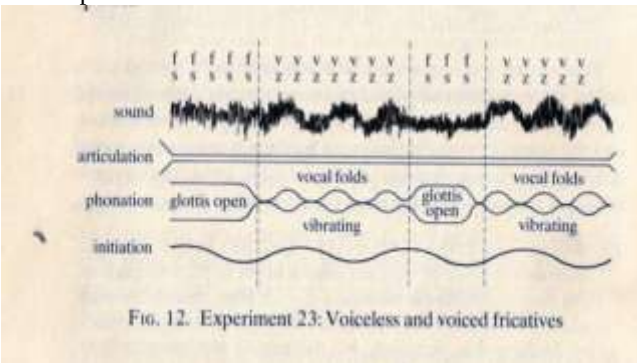

FIG. 12. Experiment 23: Voiceless and voiced fricatives

Fig. 1. Phonation – Basic Component

C. *Resonation stage:*

Voice produced in phonation is weak. So it is amplified by resonators. Resonation is voice amplification and modification. It involves the following organs.

- Pharynx which is the passageway for air and food located behind nose and mouth and includes cavity at back of tongue.
- Nose which consists of nostril (opening of external nose), nasal cavity (internal nose through which air passes to pharynx) and septum (One which divides external and internal nose into two passageways).
- Mouth which consists of vestibule (between teeth and lips) and the Oral cavity.

D. *Articulation stage:*

Tone produced in larynx is changed to specific sound by the articulators. Articulators help in production of different sound. It involves the following organs: lips, teeth, hard palate (dome), uvula, velum (soft palate), tongue and glottis.

1. Lips (upper and lower) to produce different sound
2. Teeth for creating sound
3. Tongue (tip, front, middle and back) to assist in formation of sounds of speech
4. Hard ridge behind the upper front teeth
5. Hard palate, a thin horizontal bony palate of the skull located in the roof of mouth
6. Soft palate which separates oral cavity from noise allowing air to escape through nose during speech and speech is perceived
7. Uvula to create guttural sounds with air coming from lungs and prevents air escape through nose when making sound
8. Glottis, a combination of vocal folds and space in between folds producing buzz to the speech.

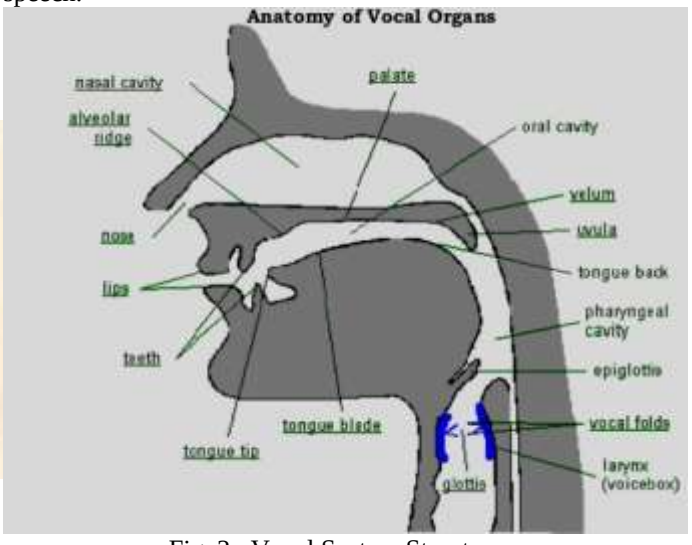

Fig. 2. Vocal System Structure

V. VOICE PRODUCTION

Voice is produced because of the above elements. The different stages of voice production are

1. Compression
2. Vibration
3. Amplification

4. Modification

A. ***Compression***

We inhale the air which fills the lungs and compresses the diaphragm. This inhaled air is compressed for exhalation.

B. ***Vibration***

Air from the lungs is exhaled out through the larynx or voice box called as vibrator. Vocal folds are present inside the larynx which vibrates when the air passes through. The vocal folds are called as vibrators proper. Vibration results in the production of initial sound of voice.

C. ***Amplification***

Initial sound is weak and amplified by air chambers called resonators. Vestibule, pharynx, nasal cavity and mouth act as resonators.

D. ***Modification***

Sound is intercepted as intelligible sounds by the modifiers or articulators such as lips, teeth, tongue, jaw, hard palate and soft palate.

VI. VOICE CHARACTERISTICS

A. Pitch is the relative tone of person voice i.e. high or low.
B. Intensity is the force or volume of tone i.e., loud or soft.
C. Quality is the character or distinctive attribute of sound resonance.
D. Vocal range is the span from highest to lowest pitch a voice can produce.
E. Speaking rate is the rate at which a person speaks and it can be fast or slow. Understandable speed is 140 to 160 words per minute. Women speak faster than men.
F. Prosody is the rhythm i.e. variation of length and accentuation of series of sound.
G. Intonation is the pitch variation

VII. ACOUSTICS MEASURES

Acoustic measure is a non-invasive measure of voice output, pressure signal and provides indirect quantifiable information about physical property of sound.

A) Period is the time taken by one vibratory cycle of the vocal folds.

B) Fundamental frequency is the rate of vibration of the vocal cord.

C) Amplitude is the height of sound wave and measure of energy in the wave.

D) Jitter is the amount of perturbation (instability) in the frequency of voice. It is the deviation or displacement of pulse in high frequency digital signal in terms of amplitude, phase timing or pulse width.

E) Shimmer is the amount of perturbation (instability) in the amplitude of voice.

F) Harmonicity (Harmonics to noise ratio) is the measure of hoarseness of voice.

According to Multi-Dimensional Voice Program (MDVP) for a normal individual, Jitter <=1.04%, shimmer <= 3.81%, HNR < 20, pitch is 128 Hz (Between 85-196 Hz) for male, and 225 Hz (Between 155-334Hz) for female. This is called thresholds of Pathology [4]. Any score above these thresholds is a cause of concern.

VIII. ATTRIBUTES IN VOICE MEASURE

A) Local Jitter (jitt) is the average absolute difference between two consecutive periods divided by average period. MDVP local jitter threshold value is 1.040% for pathology.

B) Absolute Jitter (jitta) is the average absolute difference between two consecutive periods. Threshold value is 83.2 µs for pathology.

C) Relative Average Perturbation(RAP) Jitter is the average absolute difference of one period and the average of the period with its two neighbours divided by average period. Threshold value is 0.68 % for pathology. It is absolute average perturbation divided by mean period.

D) PPQ5 Jitter is the five-point period perturbation quotient which is the average absolute difference between a period and the average of it along with its four closest neighbours divided by average period. Threshold value is 0.840 µs for pathology. It is absolute 5 point average perturbation divided by mean period and represents the ratio of disturbance within five periods.

E) DDP Jitter is the average absolute difference between consecutive differences between consecutive periods divided by average period. The value is three times RAP. It is one third of mean absolute difference of consecutive intervals divided by mean period

F) Local Shimmer is the average absolute difference between amplitudes of two consecutive periods divided by average amplitude. MDVP local shimmer threshold value is 3.81% for pathology.

G) DB Shimmer (ShDB) is the average absolute difference of the base 10 logarithm of the difference between amplitudes of two consecutive periods multiplied by 20. Threshold value is 0.350 dB for pathology.

H) APQ3 Shimmer is three-point amplitude perturbation quotient which is average absolute difference between amplitude of period and the mean amplitude of its two neighbors divided by average amplitude and represents the quotient of amplitude disturbance within three periods.

I) APQ5 Shimmer is five-point amplitude perturbation quotient which is average absolute difference between amplitude of period and the mean amplitude of it and its four closest neighbours divided by average amplitude.

J) APQ11 Shimmer is eleven-point amplitude perturbation quotient which is average absolute difference between amplitude of period and the mean amplitude of it and its ten closest neighbours divided by average amplitude. Threshold value is 3.810% for pathology.

K) DDP Shimmer is the average absolute difference between consecutive differences between amplitudes of consecutive periods. The value is three times APQ3.

L) Detrended Fluctuation Analysis(DFA): Time series are highly non stationary. A simplified definition characterizes time series as stationary if mean, standard deviation and correlation function are invariant under time translation. Signals that don't obey this are non-stationary. DFA is a method to determine self-affinity of a signal. It analyses the time signal.

M) Harmonic to Noise ratio: The proportion of periodic and non-periodic waves in vocal note is called HNR. The vocal note produced by vibration of vocal fold is complex and periodic (regular and repetitive) produced by vibration of vocal cords and non-periodic (irregular and non-repetitive) sound waves from the glottal noise. For periodic signal, signal to noise ratio equals harmonics to noise ratio. If noise is more, hoarse is more and HNR is low. Threshold value is less than 7 dB.

N) Recurrence Period Density Entropy (RPDE): It determines periodicity and repetitiveness of signal. It characterizes the extent a time series repeat the same sequence.

O) Pitch period Entropy(PPE): Healthy voice exhibit natural pitch variation which is measured by regular jitter measure. To identify PD pitches, we calculate the entropy of probability distribution. Increase in entropy measures identifies PD.

P) Noise to Harmonics Ratio: It is a measure of hoarseness.

Q) Average fundamental frequency F0: It is the number of times a sound waves produced by vocal cord repeats during a given time period and is the average value of all extracted period to period fundamental frequency values and is used to convey prosody.

R) Lowest Fundamental Frequency (Flo): It is the lowest of all extracted period to period fundamental frequency values.

S) Highest Fundamental Frequency (Fhi): It is the greatest of all extracted period to period fundamental frequency values.

T) Correlation Dimension D2: It is a measure of complexity which gives the number of independent variables to describe the system behavior.

U) Spread1: Spread1 is the log of the variance of whitened pitch periods.

V) Spread2: Spread2 is the entropy of whitened pitch periods.

IX. PARKINSON'S DATASET ATTRIBUTES

These are the vocal attributes that are used to identify Parkinson's disease. [5]

Name	ASCII subject name and recording number
MDVP: Fo(Hz)	Fundamental frequency of standard vocal
MDVP: Fhi(Hz)	Fundamental frequency of greatest vocal
MDVP: Flo(Hz)	Fundamental frequency of lowest vocal
MDVP: Shimmer (dB)	Amplitude of peak-to-peak in terms of decibels
Shimmer:APQ3	Quotient of amplitude perturbation in 3-point.
Shimmer:APQ5	Quotient of amplitude perturbation in 5-point.
MDVP:Jitter (Abs)	Fundamental frequency variation in cycle-to cycle
Jitter (relative)	Difference between the average period average and consecutive periods.
MDVP:RAP	Comparative Perturbation
MDVP:PPQ	Perturbation Quotient of Period in 5-point.
NHR	Ratio of Noise-to-Harmonic
HNR	Ratio of Harmonic-to- Noise
MDVP: Jitter(%)	Inconsistency of period-to-period.
MDVP: Fo(Hz)	Standard Fundamental Frequency
DFA	Detrended fluctuation analysis-based on casual walk
Spread1,Spread2,PPE	Quantify the fundamental frequency in variation.
Status	Health status of the subject 1- Parkinson's 0- healthy

X. CONCLUSION AND FUTURE ENHANCEMENTS

The vocal attributes to identify Parkinson's disease were discussed. Superstar singer Linda Ronstadt could not sing and gave up singing when she was diagnosed with Parkinson's at a later stage [6]. These vocal attributes help to identify the disease at an earlier stage and prevents the brain cells from getting degenerated. Thus vocal analysis can be used as a reliable and effective tool for identification of Parkinson disease.

As a future enhancement, machine learning algorithms are to be used with these vocal attributes for the identification of Parkinson disease. This would serve the physicians as an aid to identify persons with neurological problems.

REFERENCES

[1] Rajesh Pahwa, Kelly E. Lyons, Handbook of Parkinson's Disease Fourth Edition

[2] Kris Tjaden, Speech and Swallowing in Parkinson's Disease, Top Geriatr Rehabil 2008, 24(2) Pages 115-126.

[3] http://www.parkinsonsvoice.org/media.php

[4] Jaoao Paulo Teixeir, Vocal Acoustic Analysis – Jitter, Shimmer and HNR Parameters, Procedia Technology,Volume 9, 2013, Pages 1112-1122

[5] https://archive.ics.uci.edu/ml/datasets/Parkinsons.

[6] Ferrar, Tim. "Linda Ronstadt Has Parkinson's, Can No Longer Sing." (2013)