International Journal of E-Health and Medical Communications
Volume 12 • Issue 4 • July-August 2021

Vocal Folds Analysis for Detection and Classification of Voice Disorder:
Detection and Classification of Vocal Fold Polyps

Vikas Mittal, National Institute of Technology, Kurukshetra, India
https://orcid.org/0000-0003-3808-5057

R. K. Sharma, National Institute of Technology, Kurukshetra, India

ABSTRACT

The detection and description of pathological voice are the most important applications of voice profiling. Currently, techniques like laryngostroboscopy or surgical microlarynoscopy are popularly used for the diagnosis of voice pathologies but are invasive in nature. Disorders of vocal folds impact the quality of voice, and therefore, the accuracy of voice profiling is reduced. This paper presents a better solution to differentiate normal and pathological voices based on the glottal, physical, and acoustic and equivalent electrical parameters. These parameters have been correlated using mathematical equations and models. Results reveal that the glottal flow is strongly influenced by physical parameters like stiffness and viscosity of vocal folds in case of pathological voice. However, their direct measurement requires complex invasive medical procedures or costly and complex electronic hardware arrangements in case of non-invasive methods. Glottal parameters, on the other hand, facilitate much simpler estimation of vocal folds disorders. In this work, the authors have presented two non-invasive approaches for better accuracy and least complexity for differentiating normal and pathological voices: 1) by using correlation of glottal and physical parameters, 2)by using acoustic and equivalent electrical parameters.

KEYWORDS

Acoustic Circuit, Electrical Circuit, Glottal Parameters, Physical Parameters, Two Mass Model, Vocal Disorders

1. INTRODUCTION

The risk of pathological voice related disorders has increased manifolds. This is due to modern lifestyle, environmental issues, self medications and even a profession. About 25% of the population is engaged in activities that are “vocally demanding” (Amami & Smiti, 2017). The examples include professors, lawyers, auctioneers, aerobics instructors, singers, actors and manufacturing supervisors. For the diagnosis of voice pathologies, invasive endoscopy procedures are the current state of the art. But recently non-invasive digital techniques (like voice profiling and image processing) have evolved and are assisting medical professionals for early detection of voice disorders. In voice based detection, the most common method for extracting voice features is determination of acoustic parameters directly from the voice signal. Since most of the voice disorders are due to vocal fold dynamics, the researchers have started to work with glottal parameters of vocal folds to expedite the detection of related disorders. The detection of voice pathologies needs further improvement so as

DOI: 10.4018/IJEHMC.20210701.oa6

to increase the accuracy of voice detection as well as their classification. This work aims improved detection and classification of voice disorders using vocal folds glottal, physical, acoustic, and equivalent electrical parameters.

Vocal acoustic evaluation is popularly used for the assessment and diagnosis of voice disorders (Teixeira et al., 2020). Xiao Yao et al. claimed that when the speaker is under stress, certain vocal organs are affected (Yao et al., 2015).Xiao Yao et al. further discussed the physical parameter, glottal flow and stress output relationships(Yao et al., 2018). Although the voice parameters like the vibration of the vocal folds, shape of the glottis and the glottal airflow have been extensively researched in literature, yet their individual impacts on the voice quality cannot be accurately computed (Ramsay, 2019). It is a fact that thorough study and evaluation of vocal folds behavior essentially require characterization of vocal folds and their relationship with vocal tract. This paper focuses on the diagnosis of pathological voice using physical, glottal parameters as well as acoustic and its equivalent electrical parameters.

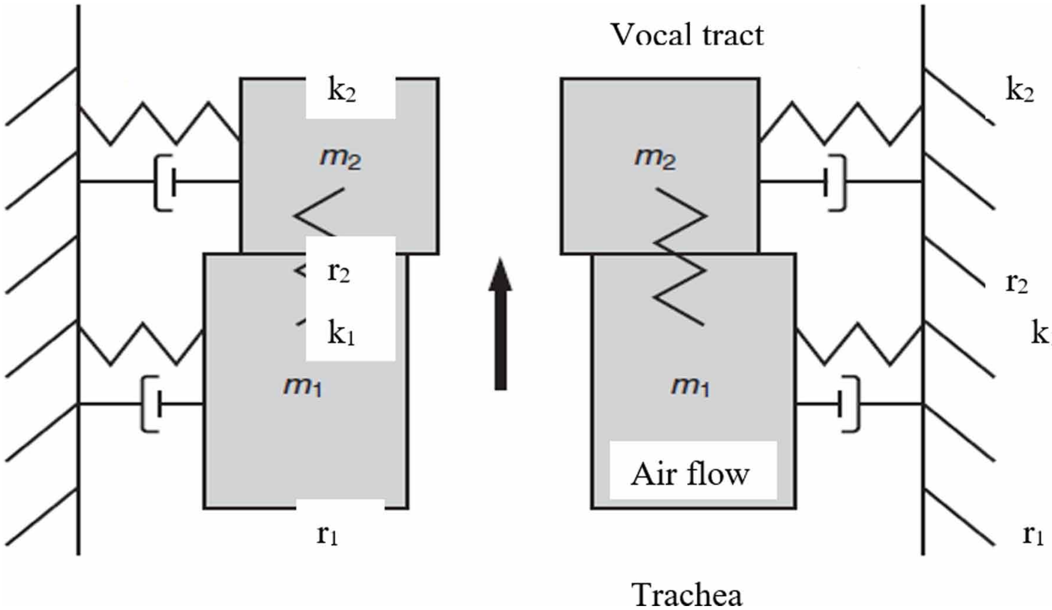
The major contributions of this paper are summarized as under:

- It is a fact that speech disorders in the voice is caused fundamentally by the physiological changes of vocal folds leading to deviation in their natural vibrations and are reflected by glottal flow. The authors have proposed a novel method to find physical parameters of vocal folds. On the other hand IAIF method is used to extract glottal flow parameters from the given voice samples. The physical parameters are then correlated with glottal flow parameters. Any change in glottal parameters reflects change in physical parameters which are then utilized to classify pathological voices. Hence the contents of the paper presents a method for detection and classification of voice disorders based on physical speech production model and characteristics of glottal flow.
- Furthermore, authors have also developed relation between vocal folds length and parameters of the acoustic model of the vocal folds. Change in vocal folds length, due to voice disorders reflects change in acoustic model parameters and the same has been used to classify pathological voices. The results of acoustic parameters variations due to voice disorders are shown in Table 10 at page-15 of the manuscript.
- It has also been experimented that current variation in equivalent electrical model is a function of change in vocal folds length. This feature has been used to classify pathological voices as shown in Figure 16.

2. PHYSICAL MODEL OF VOCAL FOLDS

The vocal folds are situated in an anterior-posterior orientation in the middle of the glottis. There is a “V” shaped form in the right and left folds. The gap in the shape of “V “forms the entrance to the trachea. Each vocal fold is attached with muscles on both sides of the larynx. The contraction and relaxation of the muscles lead to glottal flow and is termed as the glottal source. According to body-cover theory stiffness properties of vocal folds depends on thyroarytenoid (TA) and cricothyroid (CT) muscles. As a result, the behavior of vocal folds depends on the combined effect of body oscillation and layers of cover. This paper considers only the symmetrical two-mass vocal folds model, shown in Figure 1. Each vocal fold is composed of a lower mass (m_1) and an upper mass (m_2), stiffnesses (k_1, k_2), and viscous resistances (r_1, r_2). P_s is Subglottal pressure which defines the pressure of the airflow in the trachea below the glottis. The vocal fold vibration is the phonation source, which further determines the nature of the glottal flow. Since voice disorders result in variations in vocal fold stiffness (k_1), viscous resistance (r_1) and varying airflow patterns in the glottis may affect glottal flow production (Hirano, 1974).

Figure 1. Symmetrical vocal folds two mass model (Riede, 2011)



3. METHODOLOGY

For determining the relation between glottal flow and physical parameters, the procedure shown in Figure 2 has been used.

Two mass model is used to compute physical parameters (Ishizaka & Flanagan, 1972) meanwhile the glottal parameters are estimated from voice samples of normal and pathological voices(Mokhkari et al., 2018), using Aalto Aparat tool. The proposed work makes use of the German database; ‘Saarbrucken Voice Database (SVD)’. This database is freely available online. The authors have also recorded real voice samples from MMIMSR, Mullana hospital. There are total 16 samples that include eight (8) healthy and eight (8) pathological subjects, suffering with vocal folds polyps. Each category includes four (4) male and four (4) female samples all above 18 years. We have considered recordings of **vowels /a/** produced at normal pitch. The durations of samples varies between 1 to 3 seconds. All recordings are sampled at 50 kHz with 16 bit resolution.

4. RELATIONSHIP AMONG PHYSICAL & GLOTAL PARAMETERS

To achieve the objective of this work, the following procedure is used to demonstrate a correlation between physical and glottal parameters. Furthermore, the relationships among these parameters have been plotted.

4.1. Glottal Parameters

Aalto Aparat tool is used to extract glottal parameters & these parameters are briefly explained as under:

4.2. Normalized Amplitude Quotient (NAQ)

The parameter NAQ specifies the closing phase of vocal folds, and given as:

Figure 2. Blocks of methodology used

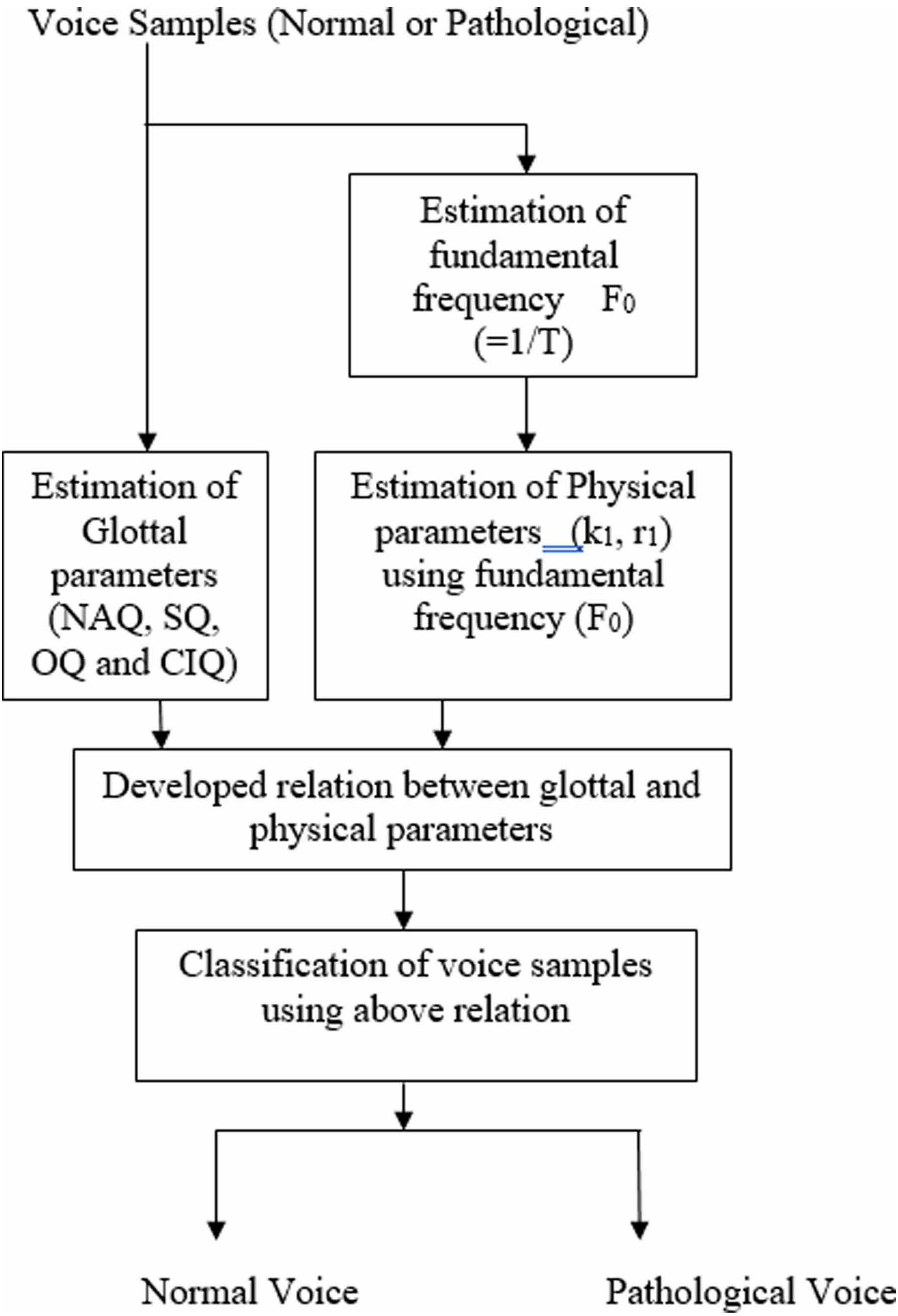
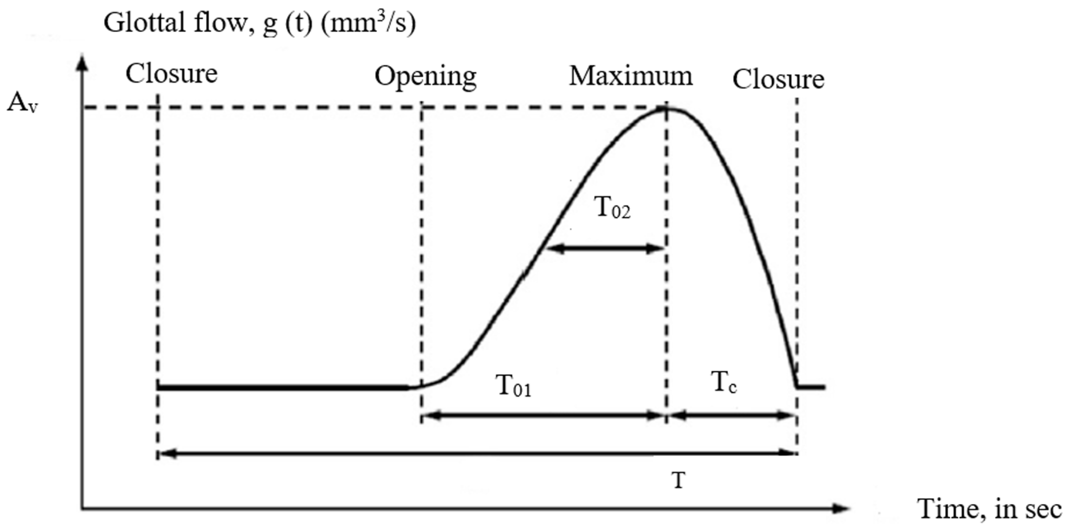


Figure 3. Representation of one glottal pulse (Forero et al., 2014)



$$\text{NAQ} = \frac{AQ}{T} \tag{1}$$

where T is the glottal period, AQ (Amplitude Quotient) is defined by the maximum amplitude of the glottal flow (Mittal & Sharma, 2020).

4.3. Speed Quotient (SQ)

It defines the ratio of opening and closing intervals of vocal folds and expressed as (Mittal & Sharma, 2020):

$$\text{SQ} = \frac{\text{T}_{\text{o1}}}{T_c} \tag{2}$$

Some authors consider two parameters for the speed quotient: SQ1 and SQ2, as:

$$\text{SQ1} = \frac{\text{T} - \text{T}_c}{\text{T}_c} \tag{3}$$

$$\text{SQ2} = \frac{\text{T}_{\text{o1}}}{\text{T}_c} \tag{4}$$

T_{o1} is a time interval between the beginning instant of opening and the instant when the opening is maximum & time interval between max. opening to the complete closure is T_c.

4.4. Opening Quotient (OQ):

It is defined as:

$$\mathrm{OQ} = \frac{\mathrm{T}_{\mathrm{o1}} + \mathrm{T}_{c}}{\mathrm{T}} \tag{5}$$

Some authors consider two parameters for the OQ: OQ1 and OQ2. These are defined as(Mittal & Sharma, 2020):

$$\mathrm{OQ1} = \frac{\mathrm{T}_{\mathrm{o1}} + \mathrm{T}_{c}}{\mathrm{T}} \tag{6}$$

and

$$\mathrm{OQ2} = \frac{\mathrm{T}_{\mathrm{o2}} + \mathrm{T}_{c}}{\mathrm{T}} \tag{7}$$

Where T_{o2} as observed in Figure 3

4.5. Closing Quotient (CIQ)

It is defined as:

$$\mathrm{CIQ} = \frac{\mathrm{T}_{\mathrm{c}}}{\mathrm{T}} \tag{8}$$

4.6. Physical Parameters:

Physical parameters stiffness and viscosity that characterize the vocal folds are discussed as under.

4.6.1. Stiffness

Stiffness is related to muscle tension and fundamental frequency (Cataldo et al., 2006). It, therefore, affects the closing and opening of vocal folds.

$$m_1 \frac{d^2 x_1}{dt^2} + r_1 \frac{dx_1}{dt} + s_1\left(x_1\right) + k_c\left(x_1 - x_2\right) = \mathrm{F}_1 \tag{9}$$

$$m_2 \frac{d^2 x_2}{dt^2} + r_2 \frac{dx_2}{dt} + s_2\left(x_2\right) + k_c\left(x_1 - x_2\right) = \mathrm{F}_2 \tag{10}$$

Where m_i, r_i, s_i and F_i are viscous resistance, elasticity, and airflow respectively. x_i is the horizontal displacement from the balance of the two masses. k_c denotes to the stiffness of the coupling between the two masses. The elasticity can be computed as:

$$s_i\left(x_i\right) = k_i\left(x_i + \eta x_i^{3}\right) \mathrm{i} = 1, 2 \tag{11}$$

where k_i is the stiffness parameter.

In the following discussion the relations of the fundamental frequency, physical parameters, and glottal parameters have been derived.

The fundamental frequency F_0 is expressed as under,

$$F_0 = \frac{1}{T} \tag{12}$$

T is given as:

$$T = \frac{AQ}{NAQ} \tag{13}$$

As per in Ishizaka and Flanagan model 1972, standard value of m_1 can be considered equal to 0.125g. Further the researchers (Ishizaka & Flanagan, 1972) relate m1& m_2 and k_1 & k_2 as expressed below:

$$m_2 = \frac{m_1}{5} \text{ and } k_2 = \frac{k_1}{10} \tag{14}$$

where k_1 is lower spring stiffness and m_1 is lower mass. Similarly, k_2 is upper spring stiffness and m_2 is upper mass.

F_0, as a function of **k** and **m,** can be defined as:

$$F_0 = \frac{1}{2\pi}\sqrt{\frac{k}{m}} \tag{15}$$

where $k = k_1 + k_2$ and $m = m_1 + m_2$.Consequently, F_0 reduces to:

$$F_0 = \frac{1}{2\pi}\sqrt{\frac{1.1k_1}{1.2m_1}} \tag{16}$$

From Equation (16), k_1 can be calculated as:

$$k_1 = \frac{\left(F0 * 2\pi\right)^2 * \left(1.2m_1\right)}{1.1} \tag{17}$$

In above developed formula k_1 is proportional to F_0 so it is inversely proportional to T. NAQ, OQ and CIQ glottal parameters are inversely proportional to T and thus stiffness parameter k_1

is proportional to NAQ. SQ parameter is proportional to glottal time period T, so it is inversely proportional to k_1.

4.6.2. Viscosity

Viscosity of vocal fold tissues plays an important role in vocal fold oscillation and it varies during phonation. Kaneko,et al. (Kaneko et al., 1972) and Isshiki (Isshiki, 1977) have measured the viscous resistance. The vocal folds viscous resistance reflects the stickiness of surfaces during vocal fold contraction, which can be calculated as:

$$r_1 = 2\zeta_1\sqrt{m_1 \mathrm{k1}} \qquad r_2 = 2\zeta_2\sqrt{m_2 k_2} \tag{18}$$

where ζ_1 , ζ_2 refer to damping ratios for the viscous resistances r_1 and r_2 .

r_1 is computed assuming damping ratio ζ_1 =0.1 as in Ishizaka and Flanagan model 1972(Ishizaka & Flanagan, 1972) . Viscous resistance r_1 is proportional to k_1. So, r_1 is also proportional to NAQ and inversely proportional to SQ glottal parameter.

5. RESULTS AND DISCUSSION

The mathematical relations develop above are used to process voice samples of healthy and pathological subjects. In this work, Saarbrucken Voice Database (SVD) (Barry & Putzer, 2015) samples are used to test relationships for the speakers in the database. Using the same set of speech samples, relevant physical parameters, and glottal parameters are derived and plotted.

The computed values of stiffness (k_1) and NAQ of Normal and Pathological Voices are shown in Table 1 and their graphical representations are shown in Figure 4. Stiffness (k_1) represents the tension in the cricothyroid (CT) muscle. The high value of stiffness (k_1) is responsible for the contraction of the cricothyroid (CT) muscle which causes slow vocal folds closure during vibration. An increase in k_1 raises NAQ. Stiffness (k_1) for normal and pathological voices is obtained using Equation17 whereas NAQ for both types of voices is obtained using inverse filtering.

Table 1. Computed values of k_1 and NAQ

Sample No.	Normal Voice		Pathological Voice	
	Stiffness k_1(kdyn/cm)	NAQ	Stiffness k_1(kdyn/cm)	NAQ
1	44.5	0.035	44.5	0.021
2	55.9	0.032	52.7	0.027
3	76.1	0.027	77.4	0.042
4	102.4	0.039	149.9	0.039
5	174.2	0.036	194.1	0.105
6	176.1	0.075	226	0.073
7	237.1	0.041	228.2	0.048
8	253.2	0.060	262.6	0.049

The computed values of k_1 and SQ of Normal and Pathological Voices are shown in Table 2.

Figure 4. Relationship between NAQ and k_1

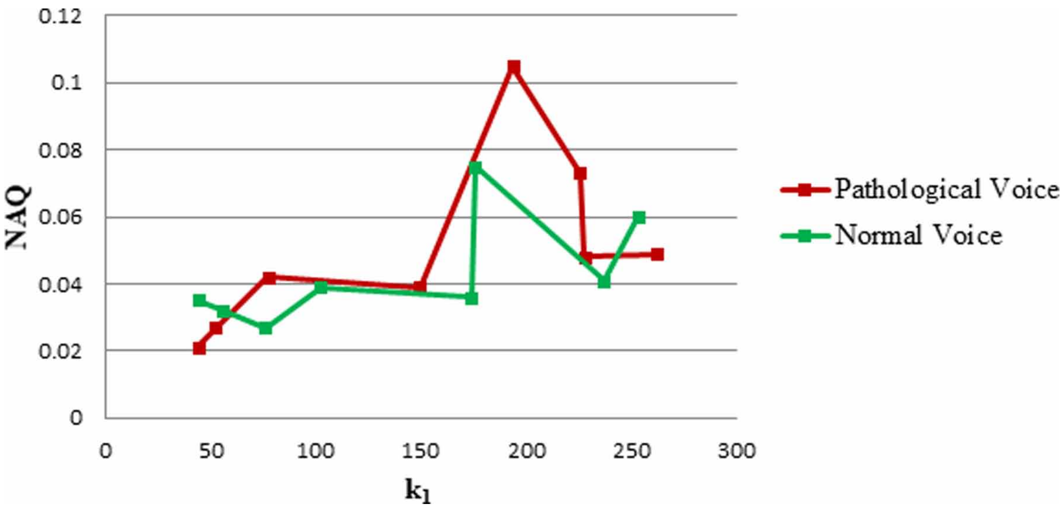


Figure 5 shows graphical representation of the relation of stiffness (k_1) with speed Quotient (SQ). The value SQ parameter is higher in pathological voice. It is clear from Figure 5 that shrinking in the structure of vocal folds due to the presence of vocal diseases.

The values of k_1 and OQ of Normal and Pathological Voices are shown in Table 3.

Figure 6 shows the relation of stiffness k_1 with the Primary Opening Quotient (OQ). It is obvious from Figure 6 that value of the OQ parameter is higher in pathological voice. The increased value of OQ is due to the partial closing of vocal folds.

The computed values of k_1 and CIQ are shown in Table 4 and Figure 7 shows the relation of stiffness (k_1) with Closing Quotient (CIQ). The value CIQ parameter is higher in pathological voice.

The computed values of r_1 and NAQ are shown in Table 5.

Figure 8 demonstrates the relationship between the glottal parameter NAQ and the physical parameter viscous resistance r_1. The value of NAQ is high in the case of pathological voices. An increase in viscous resistance r_1 will lead to an increase in NAQ. Due to that, the vocal folds activity of adduction and abduction is delayed.

The computed values of r_1 and SQ are shown in Table 6.

Table 2. Computed values of k_1 and SQ

Sample No.	Normal Voice		Pathological Voice	
	Stiffness k_1 (kdyn/cm)	SQ	Stiffness k_1 (kdyn/cm)	SQ
1	44.5	3.7	44.5	6.9
2	55.9	3.1	52.7	3.4
3	76.1	2.3	77.4	2.8
4	102.4	2.2	149.9	3.2
5	174.2	2.1	194.1	2.9
6	176.1	0.4	226	2.2
7	237.1	1.1	228.2	2
8	253.2	1.7	262.6	2.2

Figure 5. Relationship between SQ and k_1

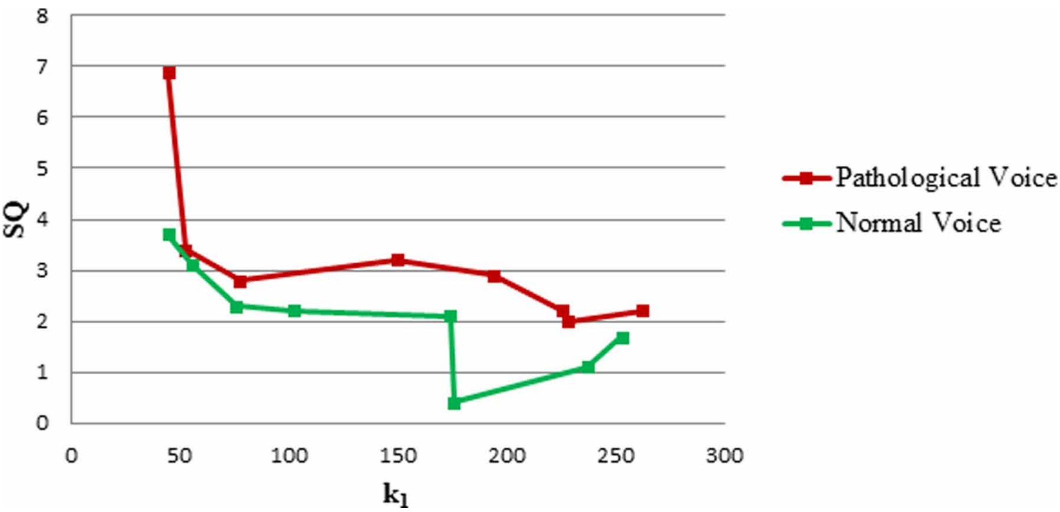


Figure 9 demonstrates the relationship between SQ and the viscous resistance r_1. An increase in viscous resistance r_1 will lead to a decrease in SQ. The value of SQ is high in the case of pathological voices. It implies that a more asymmetric glottal flow is developed.

The computed values of {r_1 and OQ} and {r_1 and CIQ} are shown in Table 7 and Table 8.

Figure 10 and Figure 11 demonstrates the relationship between the OQ, CIQ and the viscous resistance r_1. The value of OQ and CIQ is high in case of pathological voices. An increase in viscous resistance r_1 will lead to an increase in OQ and CIQ that mean slowing down the vibration speed of the vocal folds.

Authors have correlated physical parameters (stiffness, viscous resistance) and glottal flow parameters (NAQ, SQ, OQ1 and CIQ) for normal and pathological voice.

Table 9 shows the average values of computed parameters for normal and pathological voice data. The increase in k_1 for pathological voices is due to the contraction of CT which in turn happens due to distorted muscle under tension. Increased k1 leads to deceleration of vocal folds and thus asymmetrical glottal flow. This behavior of glottal flow is also reflected by larger NAQ, SQ, OQ1,

Table 3. Computed values of k_1 and OQ

Sample No.	Normal Voice		Pathological Voice	
	Stiffness k_1 (kdyn/cm)	OQ	Stiffness k_1 (kdyn/cm)	OQ
1	44.5	0.305	44.5	0.243
2	55.9	0.201	52.7	0.224
3	76.1	0.230	77.4	0.233
4	102.4	0.268	149.9	0.330
5	174.2	0.124	194.1	0.521
6	176.1	0.247	226	0.446
7	237.1	0.304	228.2	0.299
8	253.2	0.349	262.6	0.338

Figure 6. Relationship between OQ and k_1

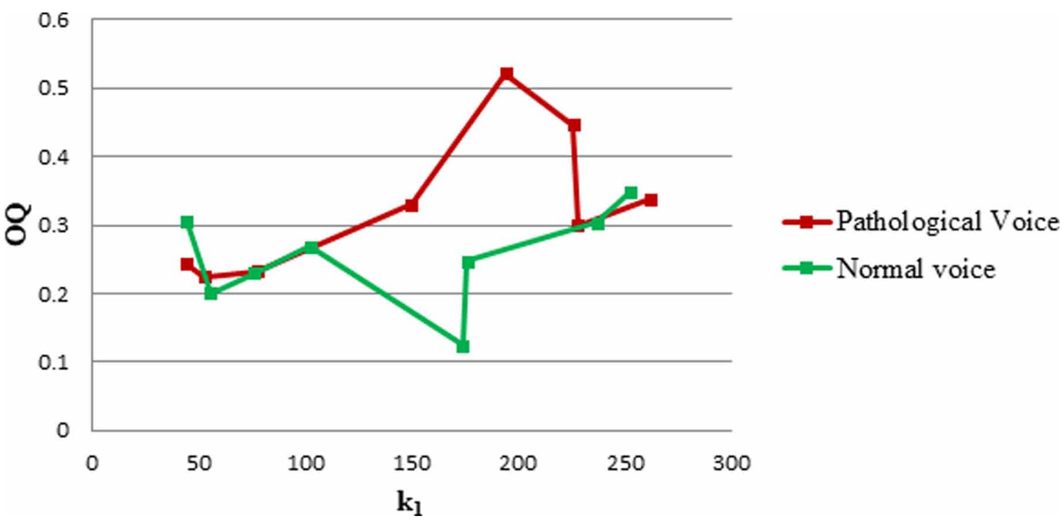


Table 4. Computed values of k_1 and CIQ

Sample No.	Normal Voice		Pathological Voice	
	Stiffness k_1 (kdyn/cm)	CIQ	Stiffness k_1 (kdyn/cm)	CIQ
1	44.5	0.037	44.5	0.167
2	55.9	0.051	52.7	0.073
3	76.1	0.057	77.4	0.069
4	102.4	0.076	149.9	0.065
5	174.2	0.17	194.1	0.016
6	176.1	0.244	226	0.35
7	237.1	0.10	228.2	0.118
8	253.2	0.106	262.6	0.110

Figure 7. Relationship between CIQ and k_1

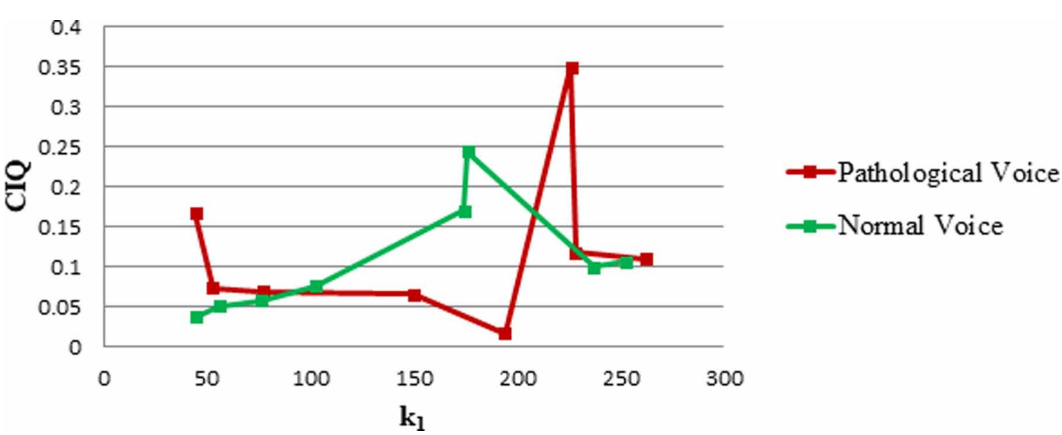


Table 5. Values of r_1 and NAQ

Sample No.	Normal Voice		Pathological Voice	
	Viscous resistance (r_1)	NAQ	Viscous resistance (r_1)	NAQ
1	14.9	0.035	14.9	0.021
2	16.7	0.032	16.2	0.027
3	19.5	0.027	19.6	0.042
4	22.6	0.039	27.3	0.039
5	29.5	0.036	31.1	0.105
6	29.6	0.075	33.6	0.073
7	34.5	0.041	33.7	0.048
8	35.5	0.060	36.2	0.049

Figure 8. Relationship between NAQ and r_1

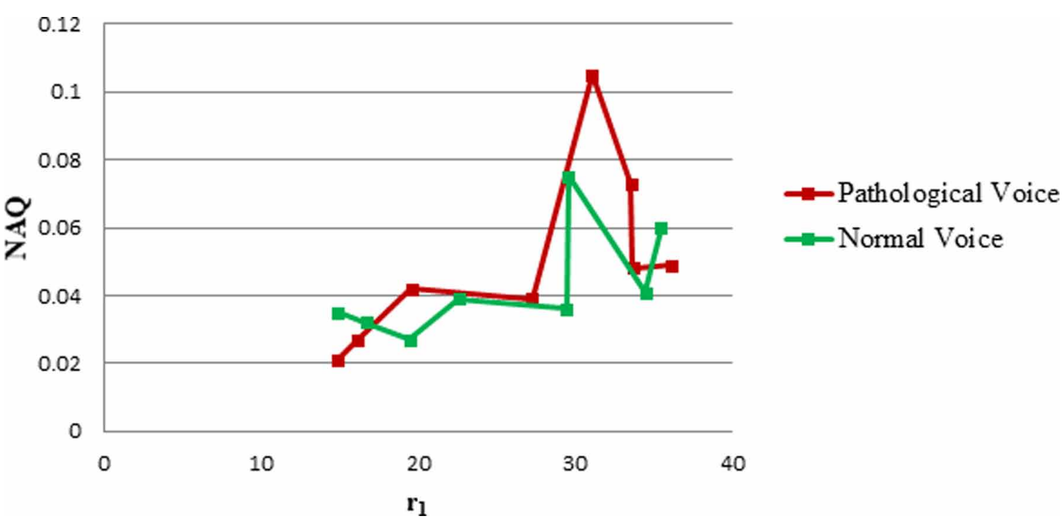


Table 6. Computed values of r_1 and SQ

Sample No.	Normal Voice		Pathological Voice	
	Viscous resistance (r_1)	SQ	Viscous resistance (r_1)	SQ
1	14.9	3.7	14.9	2.8
2	16.7	3.1	16.2	3.1
3	19.5	2.3	19.6	3.8
4	22.6	2.2	27.3	3.2
5	29.5	2.1	31.1	2.9
6	29.6	0.4	33.6	2.2
7	34.5	1.1	33.7	2
8	35.5	1.7	36.2	6.9

Figure 9. Relationship between SQ and r_1

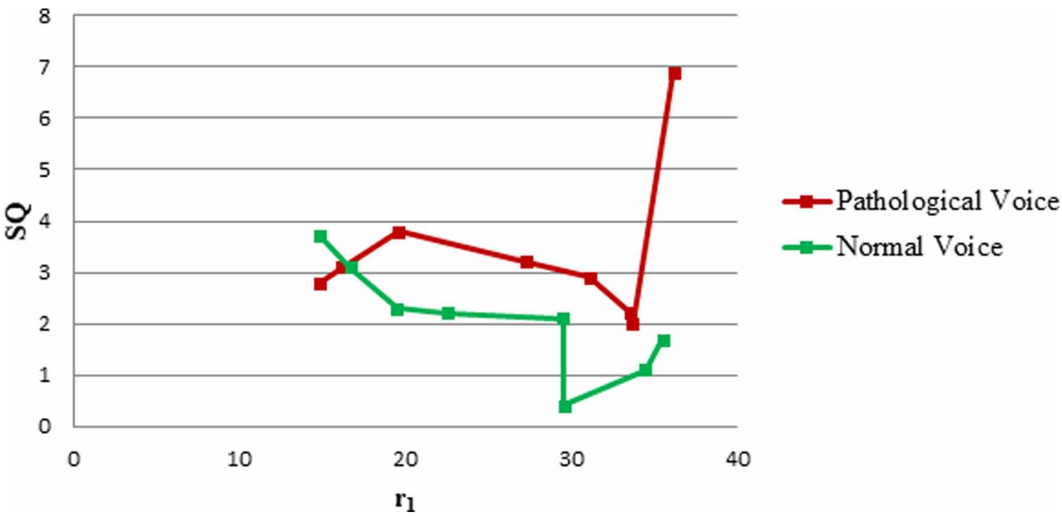


Table 7. Computed values of r_1 and OQ

Sample No.	Normal Voice		Pathological Voice	
	Viscous resistance (r_1)	OQ	Viscous resistance (r_1)	OQ
1	14.9	0.305	14.9	0.243
2	16.7	0.201	16.2	0.224
3	19.5	0.230	19.6	0.233
4	22.6	0.268	27.3	0.330
5	29.5	0.124	31.1	0.521
6	29.6	0.247	33.6	0.446
7	34.5	0.304	33.7	0.299
8	35.5	0.349	36.2	0.338

and CIQ. The Viscous resistance r_1 for pathological voice is substantially greater than normal voice. This leads to a stickier surface of vocal folds.

The physical parameters reflect the physical characteristics in the physiological system but their computation is complex. In the detection of voice pathology, the glottal parameters often perform well and the estimation method is simple. In this work, the glottal parameters are estimated for normal and pathological voice samples. Further, these parameters have been related to physical parameters. As seen from results, these parameters are effective in detecting vocal folds disorder as presented above in Table 9.

In subsequent sections, a method based on acoustic and electrical circuits is proposed to differentiate between normal and pathological voices.

Flanagan and Landgraf in 1968(Flanagan & Landgraf, 1968) were the first researchers who modeled the acoustic behavior of vocal folds using one mass model and assumed lungs as a constant-pressure source should it be denoted as P_s . Further Van den Berg (Van den Berg, 1958), experimented on one mass model and studied the impact of constant and variable pressure on air volume velocity

Figure 10. Relationship between OQ and r_1

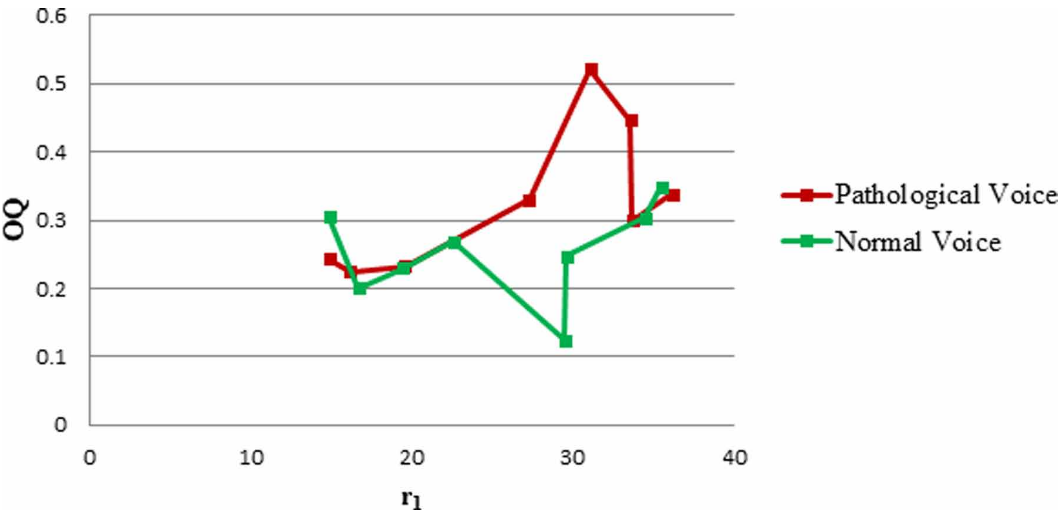


Table 8. Computed values of r_1 and CIQ

Sample No.	Normal Voice		Pathological Voice	
	Viscous resistance (r_1)	CIQ	Viscous resistance (r_1)	CIQ
1	14.9	0.037	14.9	0.167
2	16.7	0.051	16.2	0.073
3	19.5	0.057	19.6	0.069
4	22.6	0.076	27.3	0.065
5	29.5	0.170	31.1	0.016
6	29.6	0.244	33.6	0.35
7	34.5	0.100	33.7	0.118
8	35.5	0.106	36.2	0.110

(U_g) and express the time-varying glottal impedance as function of viscous non-flow dependent resistance (R_v), a kinetic flow dependent resistance (R_k) and inertance (L_g).

The acoustic circuit of vocal folds based on two- mass model (Ishizaka & Flanagan, 1972) is shown in Figure 12. The elements of the acoustic circuit are defined as,

$$R_{v1} = 12\frac{\mu L^2 d_1}{A_{g1}^3} \tag{19}$$

$$R_{v2} = 12\frac{\mu L^2 d_2}{A_{g2}^3} \tag{20}$$

Figure 11. Relationship between CIQ and r_1

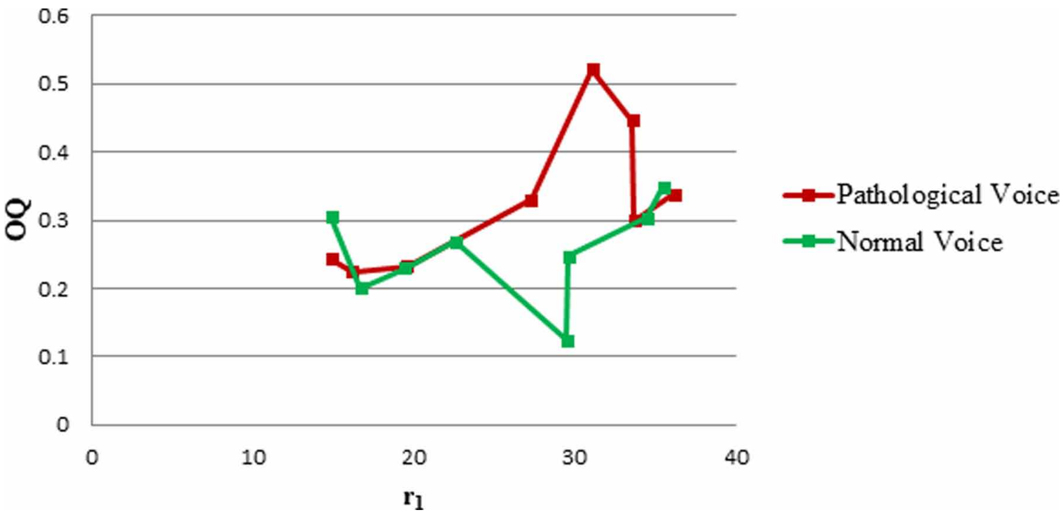


Table 9. Average value of physical and glottal parameters

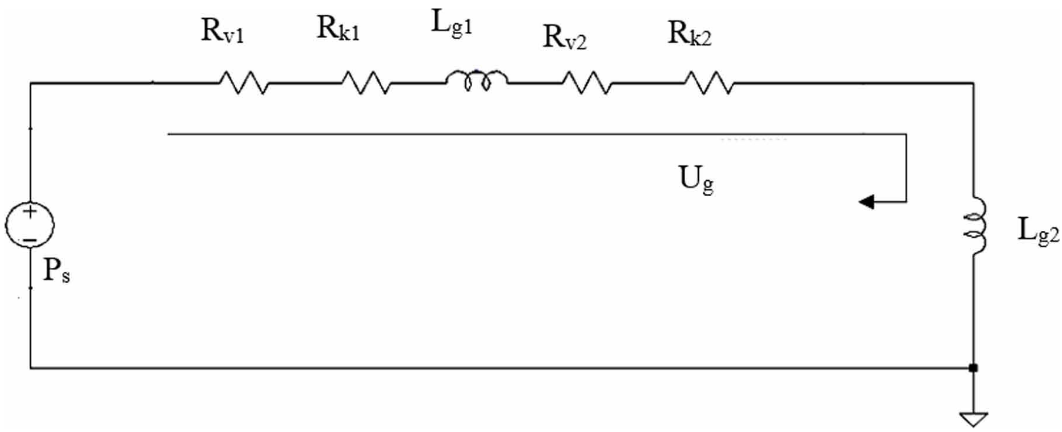
Physical and glottal Parameters	Normal Voice	Pathological Voice	(%) increase in pathological parameters
NAQ	0.03	0.05	66
SQ	2.07	3.36	62
OQ1	0.25	0.32	28
CIQ	0.08	0.10	25
k_1(kdyn/cm)	139.9	154.4	10
r_1	25.3	26.5	04

$$L_{g1} = \frac{\rho d_1}{A_{g1}} \tag{21}$$

$$L_{g2} = \frac{\rho d_2}{A_{g2}} \tag{22}$$

$$R_{k1} = \frac{0.19\rho}{A_{g1}^2} \tag{23}$$

Figure 12. The acoustic circuit representation for the vocal folds (Ishizaka & Flanagan, 1972)



$$R_{k2}= \frac{\rho\left[0.5-\left(\frac{A_{g2}}{A1}\right)\left(1-\left(\frac{A_{g2}}{A1}\right)\right)\right]}{A_{g2}^{2}} \tag{24}$$

Where,

d_1 and d_2: thickness of mass m_1& m_2,respectively,
$A_{g1}$and A_{g2}: cross-sectional areas at the two masses (m_1& m_2),
L_g: effective length of the vocal folds,
$R_{v1}$and R_{v2}: viscous non-flow dependent resistances,
L_{g1} and L_{g2}: inertances due to the two masses (m_1& m_2),
R_{k1} and R_{k2}: kinetic flow dependent resistances
Viscosity of air (μ) = $1.86*10^{-4}$ dyne-sec/cm^2, at 30^0C
Air density (ρ) = $1.14*10^{-3}$ g/cm^3, 30^0C

In the acoustic circuit given in Figure 12, if there is any change in subglottal pressure or in the values of components, the volume velocity (U_g) changes. Change in U_g reflects the change in the fundamental frequency of normal or pathological voices. From the above circuit, it can be shown that:

$$\left(R_{k1}+R_{k2}\right)\left|U_{g}\right|U_{g}+\left(R_{v1}+R_{v2}\right)U_{g}+\left(L_{g1}+L_{g2}\right)\frac{dU_{g}}{dt}=\text{Ps} \tag{25}$$

Assuming constant lung pressure P_s of the acoustic circuit, if the length of the vocal folds (L) changes, the values of components of acoustic circuit also changes {eqns. 19 to 24}. The changed values of the components imply a change in volume velocity U_g and consequently, change in the fundamental frequency. Table 10 shows values of R_{v1}, R_{v2}, R_{k1}, R_{k2}, L_{g1}, and L_{g2} w.r.t different vocal folds length (L) and a corresponding change in fundamental frequencies for normal and pathological voices.

Table 10. Computed values of R_{v1}, R_{v2}, R_{k1}, R_{k2}, L_{g1} and L_{g2} for varying vocal lengths (L)

Normal Voices							
L (cm)	R_{v1}	R_{v2}	R_{k1}	R_{k2}	L_{g1}	L_{g2}	F_0
0.94	0.0039	0.118	0.00095	0.063	0.60	0.60	180
0.8	0.0055	0.13	0.0013	0.088	0.70	0.70	210
1.6	0.0027	0.069	0.00035	0.021	0.35	0.35	102
1.23	0.0036	0.09	0.00056	0.037	0.46	0.46	138
0.78	0.00057	0.14	0.0014	0.09	0.73	0.73	217
1.7	0.0025	0.064	0.00033	0.017	0.32	0.32	98
1.42	0.0032	0.077	0.00042	0.027	0.40	0.40	119
0.93	0.0039	0.11	0.00095	0.063	0.60	0.60	181
Pathological Voices							
L (cm)	R_{v1}	R_{v2}	R_{k1}	R_{k2}	L_{g1}	L_{g2}	F_0
0.82	0.0055	0.13	0.0013	0.088	0.71	0.71	205
0.76	0.0057	0.14	0.0014	0.09	0.73	0.73	221
1.4	0.0032	0.077	0.00042	0.027	0.40	0.40	120
1.0	0.0044	0.108	0.00084	0.055	0.57	0.57	167
0.89	0.0057	0.14	0.0013	0.088	0.71	0.71	190
1.7	0.0027	0.067	0.00037	0.023	0.38	0.38	99
1.26	0.0040	0.108	0.00084	0.055	0.57	0.57	134
0.82	0.0057	0.14	0.0014	0.09	0.73	0.73	206

Equation 26 relates fundamental frequency (F_0) and vocal folds length (L). Decrease in L results into increase in F_0.

$$F_0 = \left(\frac{8.67}{L}\right)\left(1 + 5.69\frac{A^2}{L^2}\right)e^{4.61\left(\frac{L}{L0}\right)} \tag{26}$$

Where,

L_0 is abducted reference length,
A is vibration amplitude

For the human vocal folds, amplitude to length (A/L) ratio is of the order of 0.1 and 0.5<L/L0<1.0(Titze, 1989a).

Another simpler relationship connecting fundamental frequency (F_0) and vocal folds membranous length (L_m) is (Titze, 1989b) shown in Equation 27.

$$F_0 = \frac{1700}{L_m} \tag{27}$$

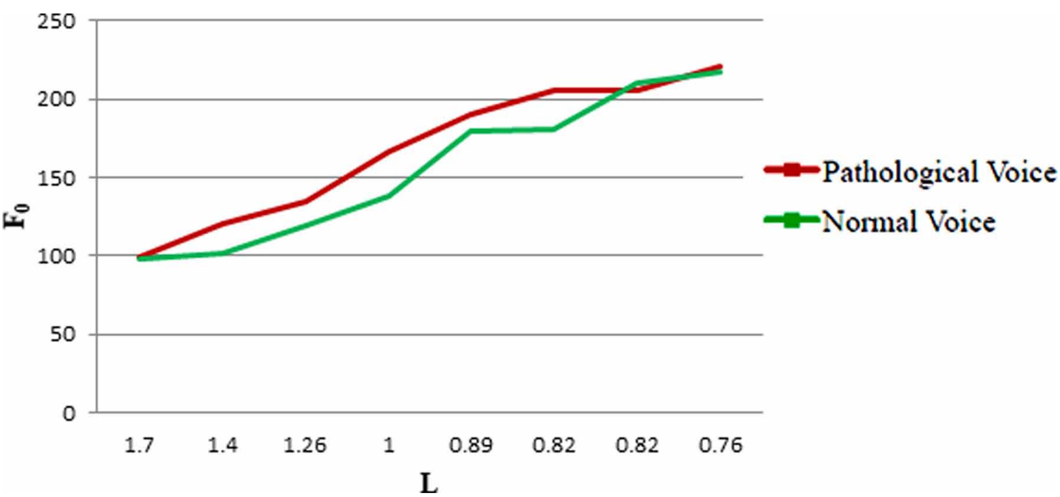
Table 11 and Figure 13 show that if vocal folds length reduces, the fundamental frequency increases.

Table 11. Computed values of fundamental frequency and vocal folds length

Normal Voices		Pathological Voices	
*fundamental frequency(F_0) Hz	** vocal folds length L(cm)	*fundamental frequency(F_0) Hz	** vocal folds length L(cm)
98	1.7	99	1.7
102	1.6	120	1.4
119	1.42	134	1.26
138	1.23	167	1
180	0.94	190	0.89
181	0.93	205	0.82
210	0.8	206	0.82
217	0.78	221	0.76

*Fundamental frequencies are obtained using Aalto Aparat tool.
** Length of vocal folds is obtained using equation 27.

Figure 13. Fundamental frequency vs Vocal folds Length



Now, if the lung pressure P_s in the acoustic circuit changes then volume velocity (U_g) also changes and leading to change in fundamental frequency. These observations are depicted in Table 12 and Figure 14. Here it is assumed that the values of circuit components are constant.

The value of lung pressure P_s is calculated with help of following relationships as given by Hocine Teffahi (Baken & Orlikoff, 2000),

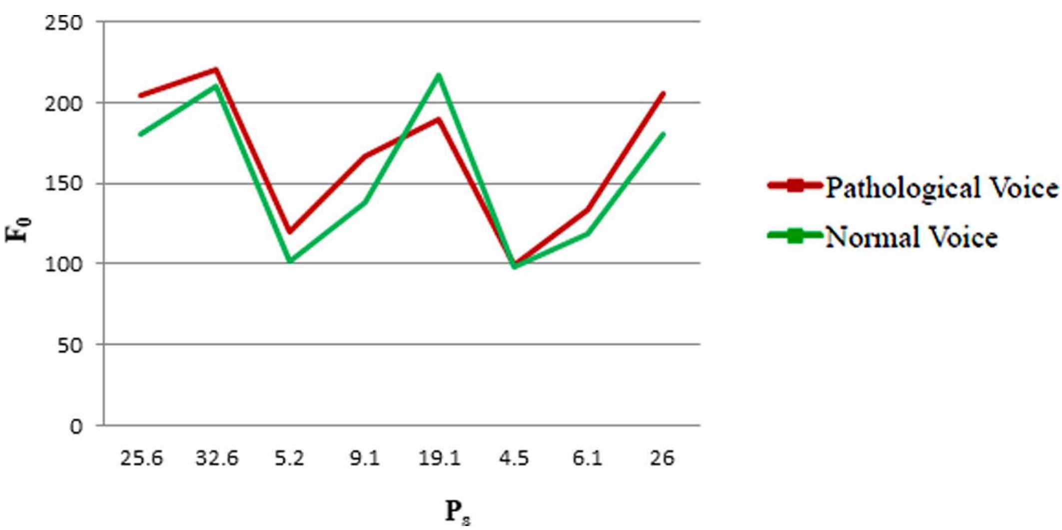
$$F_0\,(\text{Hz}) = 2.3Ps + 48Q + 1.98 \tag{28}$$

Table 12. Computed values of fundamental frequency and subglottal pressure using Equation 28.

Normal Voices		Pathological Voices	
*fundamental frequency(F_0)Hz	**subglottal pressure (P_s) in cmH_2O	*fundamental frequency(F_0) Hz	**subglottal pressure (P_s) in cmH_2O
180	14.7	205	25.6
210	27.8	221	32.6
102	4.9	120	5.2
138	6.2	167	9.1
217	30.8	190	19.1
98	4.5	99	4.5
119	5.2	134	6.1
181	14.8	206	26

*Fundamental frequencies are obtained using Aalto Aparat tool
**subglottal pressure P_s is obtained using Equation 28

Figure 14. Fundamental frequency vs subglottal pressure



Assume, Q (=3) is vocal cord tension (Teffahi, 2009), which is proportional to Fundamental frequency (F_0).

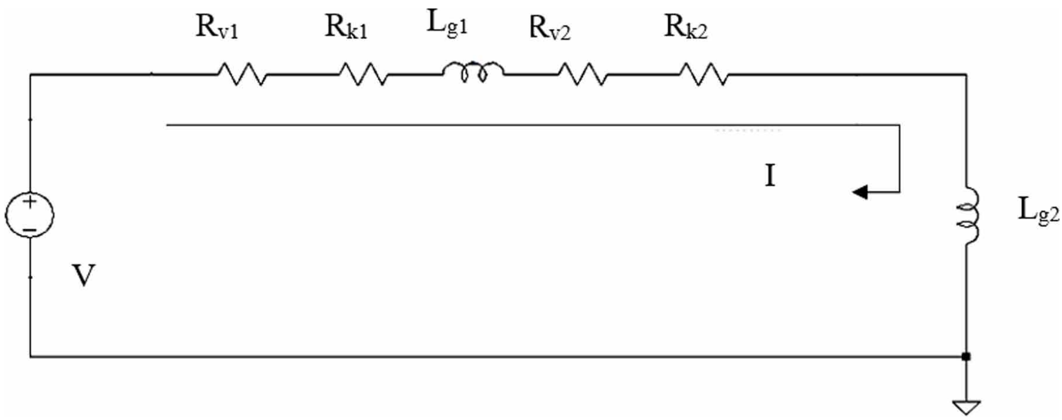
The equivalent electrical circuit derived from the acoustic circuit (Figure 12) is shown in Figure 15. Circuit components values are a function of the length of the vocal fold. The voltage (V) of the circuit which is equivalent to lung pressure (P_s) is calculated by using the following relation (Baken & Orlikoff, 2000),

$$1 cmH_2O \text{ Pressure} = 1.27 * P_{alv} + 5.94 \tag{29}$$

where P_{alv} is the alveolar pressure approximated as dc voltage equivalent.

The mathematical model of the circuit is given by the following relation:

Figure 15. Equivalent electrical circuit for the vocal folds



$$\left(R_{k1} + R_{k2}\right)\left|I\right|I + \left(R_{v1} + R_{v2}\right)I + \left(L_{g1} + L_{g2}\right)\frac{dI}{dt} = V \tag{30}$$

To analyze the above circuit two scenarios are considered.

a) *When voltage is constant and components values are variable*: The current in the circuit is different for normal and pathological voices. Current variations are inversely related to variations in fundamental frequencies of normal or pathological voices.

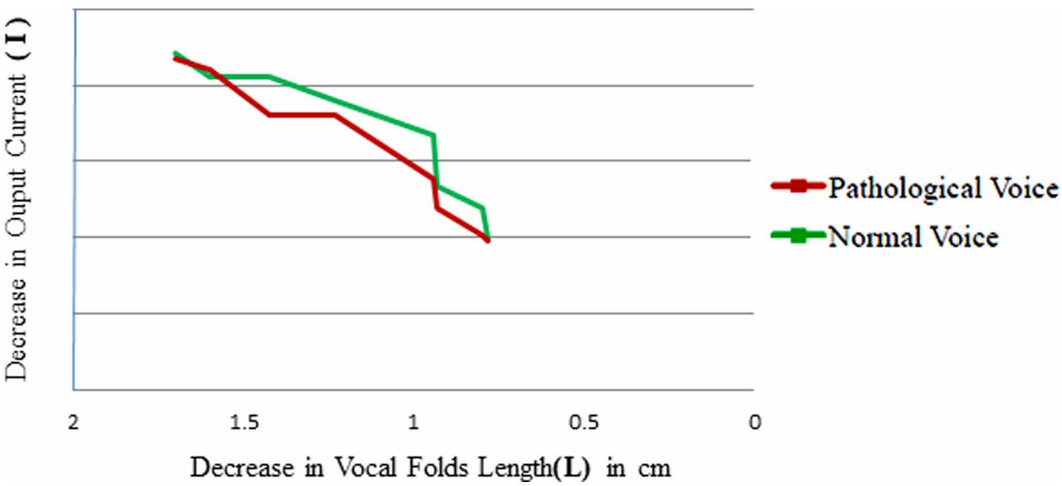
Figure 16 shows the variations of output current (I) of the equivalent electrical circuit as a function of the vocal folds length for normal and pathological voices. It is clear from Figure 16 that as the vocal folds length is decreasing the output current (I) is also decreasing.

b) *When voltage is variable and components values are constant*: The current in the circuit is different for normal and pathological voices. This is inversely related to variations in fundamental frequencies of normal or pathological voices

6. CONCLUSION AND FUTURE SCOPE

Improvement of health and non-invasive diagnosis and treatment of chronic diseases are the major requirements in Biomedical Field. Voice disorders can have a significant negative impact on the social and professional life. Although such disorders are often underestimated, their early detection and accurate diagnosis are necessary to reduce serious consequences. The health of persons may be severely affected by their individual pathological voice conditions. This may financially burden such patients and even the society at large. One of the most frequently utilized tools to diagnose these vocal disorders is a laryngoscope. Laryngoscopy, an invasive and painful technique, is an expensive time-consuming process that requires trained personnel to perform the test. To address these issues, researchers have been experimenting with non-invasive techniques for detecting vocal disorders. This paper has successfully presented an alternate non invasive diagnostic method to accurately and quickly classify voice disorders. The proposed method provides an opportunity to further improve the existing medical techniques that are necessary to diagnose voice disorders. The results obtained in this paper reveal the essence and mechanism for the pathological voice and establish a theoretical

Figure 16. Trend of output current of electrical circuit for normal and pathological voices



and experimental foundation for pathological voice detection. Physical & glottal parameters are defined, related and their impact on normal & pathological voices has been presented in the paper. Stiffness (k_1) and Viscous Resistance (r_1) are higher for pathological voices. The Dependence of glottal parameters on physical parameters is also studied and verified. It is also concluded that if k_1, r_1, or both change then F_0, and NAQ, SQ, OQ and CIQ change. All these variations enable us to distinguish between normal and pathological voices. Compared to normal voices, the pathological voices have shown increased values of both physical & glottal parameters. Further acoustic & electrical models have been synthesized that also detect pathological voices. As future work, authors plan to develop database and try other classification methods to reach a definitive diagnosis for Vocal folds disorders.

ACKNOWLEDGMENT

The authors are thankful to the Special Manpower Development Program, Chip-to-System Design (SMDP-C2SD), initiated by the Ministry of Electronics & Information Technology (MeitY), Govt. of India, for providing lab facilities in the School of VLSI Design and Embedded Systems, NIT, Kurukshetra.

REFERENCES

Amami, R., & Smiti, A. (2017). An incremental method combining density clustering and support vector machines for voice pathology detection. *Computers & Electrical Engineering*, *57*, 257–265. doi:10.1016/j.compeleceng.2016.08.021

Baken & Orlikoff. (2000). Clinical measurement of speech & voice (2nd ed.). Cengage Learning.

Barry, W. J., & Putzer, M. (2015). *Saarbrucken Voice Database.* Institute of Phonetics, University of Saarland. http://www.stimmdatenbank.coli.uni-saarland.de/

Cataldo, Leta, Lucero, & Nicalato. (2006). Synthesis of voiced sounds using low-dimensional models of the vocal cords and tim-varying subglottal pressure. *Mechanical Research Communications, 33*, 250-260. doi: 10.1016/j.mechrescom.2005.05.007

Flanagan, J., & Landgraf, L. (1968). Self Oscillating source for vocal-tract synthesizers. *IEEE Transactions on Audio and Electroacoustics*, *16*(1), 57–64. doi:10.1109/TAU.1968.1161949

Forero, Kohler, Vellasco, & Cataldo. (2014). Classification of Vocal Aging Using parameters extracted from glottal signal. *Journal of Voice, 28*, 532-537. doi: 10.1016/j.jvoice.2014.02.001

Hirano. (1974). Morphological structure of the vocal cord as a vibrator and its variations. *Folia Phoniatr., 26*(2), 89–94. doi: 10.1159/000263771

Ishizaka, K., & Flanagan, J. L. (1972). Synthesis of voiced sounds from a two-mass model of the vocal cords. *The Bell System Technical Journal*, *51*(6), 1233–1268. doi:10.1002/j.1538-7305.1972.tb02651.x

Isshiki, N. (1977). Functional Surgery of the Larynx. Kyoto University.

Kaneko, T., Asano, H., Naito, J., Kobayashi, N., Hayashi, K., & Kitamura, T. (1972). Biomechanics of the vocal cords-on damping ratio. *J. Jpn. Bronchoesophagol. Soc.*, *25*(3), 133–138. doi:10.2468/jbes.25.133

Mittal, V., & Sharma, R. K. (2020). Voice Signal Analysis with the Application in Biomedicine. *Sensor Letters*, *18*(2), 122–127. doi:10.1166/sl.2020.4187

Mokhkari, Story, Alku, & Ando. (2018). Estimation of the glottal flow from speech pressure signals: Evaluation of three variants of iterative adaptive inverse filtering using computational physical modeling of voice production. *Speech Communication, 104*, 24-38. doi: 10.1016/j.specom.2018.09.005

Ramsay. (2019). Mechanical speech synthesis in early talking automata. *Acoustics Today, 15*, 11-19. doi: 10.1121/AT.2019.15.2.11

Riede. (2011). *Subglottal pressure and fundamental frequency control in contacts calls of juvenile alligator mississippiensis*. doi:10.1242/jeb.051110

Teffahi, H. (2009). A Two-Mass model of the vocal cords: determination of control parameters. *Proceeding International Conference on Multimedia Computing and Systems*. doi:10.1109/MMCS.2009.5256726

Teixeira, J. P., Alves, N., & Fernandes, P. O. (2020). Vocal Acoustic Analysis:ANN Versos SVM in Classification of Dysphonic Voices and Vocal Cords Paralysis. *International Journal of E-Health and Medical Communications*, *11*(1), 37–51. doi:10.4018/IJEHMC.2020010103

Titze, I. R. (1989a). On the relation between subglottal pressure and fundamental frequency in phonation. *The Journal of the Acoustical Society of America*, *85*(2), 901–906. doi:10.1121/1.397562 PMID:2926005

Titze, I. R. (1989b). Physiologic and acoustic diffrences between male and female voices. *The Journal of Acoustical Society of America, 85*, 1699-1707. doi: 10.1121/1.397959

Van den Berg, J. (1958). Myoelastic-aerodynamic theory of voice production. *Journal of Speech Hearing Research, 1*, 227-244. doi: 10.1044/jshr.0103.227

Yao, X., Jitsuhiro, T., Miyajima, C., Kitaoka, N., & Takeda, K. (2015). Modeling of Physical Characteristics of Speech under stress. *IEEE Signal Processing Letters*, *22*(10), 1801–1805. doi:10.1109/LSP.2015.2434732

Yao, X., Xu, N., Liu, X., Jaing, A., & Zhang, X. (2018). Research on speech under stress based on glottal source using a physical model. *IEEE Access : Practical Innovations, Open Solutions*, *6*, 44473–44482. doi:10.1109/ACCESS.2018.2860130

Vikas Mittal received his M.Tech in Electronics and Communication Engineering from Kurukshetra University Kurukshetra. Currently, he is pursuing Ph.D. at the National Institute of Technology (NIT) in the School of VLSI Design and Embedded Systems. His research interests include Biomedical Signal Processing and Circuits Design.

R. K. Sharma received his M.Tech in Electronics and Communication Engineering and PhD degree in electronics and communication from Kurukshetra University, Kurukshetra (through National Institute of Technology Kurukshetra), India in 1993 and 2007, respectively. Currently, he is a Professor with the Department of Electronics and Communication Engineering, NIT Kurukshetra, India. His main research interests are in the field of embedded applications, low power, digital design, and disease/ stress detections using voice profiling of human beings.