A New Diagnostic Approach: Voice Analysis of Thyroid Patients after Thyroid Surgery without Laryngeal Nerve Injury

Prof Gururaj B Gour¹, Dr V Udayashankara²
¹ BLDEAs CET/ECE, Bijapur, India
² SJCE/IT, Mysore, India

Email: guru_g123@rediffmail.com
Email: v_udayashankara@yahoo.co.in

Abstract—Voice dysfunction is an important complication due to an injury to the recurrent and superior laryngeal nerves following thyroid surgery. Voices of six female patients were studied who underwent surgery with preservation of the recurrent and superior laryngeal nerves during the three months of observation period after surgery. Acoustic voice analyses of thyroid patients who underwent surgery was performed preoperatively and postoperatively one week, one month, 2 months and 3 months after surgery along with formant frequency analysis. The preoperative voice is slightly abnormal and fundamental frequency (F0) and standard deviation of pitch (STD), voice turbulence index (VTI), degree of sub harmonics were decreased significantly after one month of surgery (P<0.05). Acoustic indicators returned to normal after 3 months of surgery.

Index Terms—Voice acoustic analysis, Thyroid lobectomy, Thyroidectomy, Thyroid—surgery, Postoperative dysphonia, Vocal tract steadiness.

I. INTRODUCTION

The thyroid is one of the largest endocrine glands in the body, is found in the neck and sits in front of the trachea (also known as the windpipe). When it doesn't function properly, leads to Thyroid disorder. And 15% to 30% of urban people have prone to thyroid disorder due to existing life style and women are four times more prone to thyroid disorders than men. Parathyroid disease (hyperparathyroidism) causes symptoms in most people, but it will decrease the life expectancy in all patients by about 5-6 years, and the fact that people with hyperparathyroidism have nearly twice the chance of developing breast and prostate cancer, heart disease, and stroke. Vocal dysfunction is one of the most important complications of thyroid patients [5,6,7]. Vocal complications following thyroid surgery has long been received attention. It was thought that injury to recurrent laryngeal nerve (RLN) and injury of the external branch of the superior laryngeal nerve (EBSLN) were the main causes of voice change in patients who underwent thyroid surgery [4, 24]. The results of injury to this nerve can be dramatic and even life threatening. Therefore it became routine step in thyroid procedure to correctly locate the RLN or EBSLN to prevent injury to these nerves. Preservation of both RLN and EBSLN is important in sustaining voice function. But still voice change occurs even without laryngeal nerve injury. And these changes were mainly presented as difficulty in producing a high pitch, low voice, and mild hoarseness, which may include asymmetrical vocal cords, difficulty in swallowing. The complications can be
psychogenic dysphonia as a result of diagnosis or surgery, strap muscles or nerves that have been stretched by a goiter or by surgical retraction, laryngeal bruising caused by intubation, maladaptive voice behavior due to a mass in the neck, or endocrine changes at the time of surgery. Dysphonia following surgery may have significant occupational, psychosocial, and economic consequences for the patient. Furthermore, failure to address these issues (either during patient consent or during the operative procedure) may have serious consequences [22, 23, 25].

Acoustic voice analysis can give important objective data on voice disturbances, especially those with organic and functional origins. Therefore we will find the characteristics of voice changes in Thyroid patients and their possible impacting factors [1, 4]. The objective here is to investigate the features of voice changes in Thyroid patients. Six acoustic parameters that are sensitive to signal properties are used to analyze voices between two groups. The Control group is comprised of normal voices, while the other group comprised of thyroid patients who underwent surgery. The success of applying measures in acoustic voice analysis of the vibrating vocal folds, may be to correlate results with the perceived roughness [2].

II. THYROID GLAND AND VOICE DYSFUNCTION

A. Normal thyroid homeostasis

Number of citations Thyroid gland is controlled by the Thyroid Stimulating Hormone (TSH) and is in turn controlled by TSH releasine hormone (TRH) as shown in Fig.1. Clinically the TSH is important due to close proximity of hypothalamus and pituitary and a disease affecting one involves the other. Pituitary will senses the less thyroid hormone production by thyroid and TSH rises to stimulate the thyroid gland. Here, the primary problem is with thyroid gland and is called as primary hypothyroidism. In secondary hypothyroidism, pituitary fails to produce normal amount of TSH even though thyroid gland is normal but cannot produce thyroid hormone in absence of proper stimuli. In a similar way when thyroid gland produces more than normal thyroid hormone is called primary hyperthyroidism and in case of pituitary tumor, excess TSH stimulates thyroid gland leads to thyroid hormone in circulating blood, called secondary hyperthyroidism.

B. Laryngeal nerves and Voice dysfunction

The thyroid gland controls body metabolism and has a profound effect on all bodily functions. The thyroid gland is located in the lower anterior neck in close relationship to the larynx, the trachea, the esophagus, the carotid sheath structures, the sympathetic chain, the recurrent laryngeal nerve, and the mediastinal structures. The vagus nerve exits the base of the skull through the jugular foramen and descends in the neck within the carotid sheath. The first branch is the superior laryngeal nerve, which travels in close association with the superior thyroid artery. This nerve innervates the cricothyroid muscle through its external branch and supplies sensation to the supraglottic larynx through its internal branch. Damage to this branch may increase the risk of aspiration and can affect the singing voice but has little effect on the speaking voice. The recurrent laryngeal nerve arises from the vagus nerve and has a different course in the right and left neck as far as the tracheoesophageal groove is concerned. The recurrent laryngeal nerve innervates the intrinsic muscles of the larynx (except for the cricothyroid). Damage to this nerve results in unilateral vocal cord paralysis, which leads to varying degrees of breathy voice, ineffective cough, and risk of aspiration. This nerve is most at risk during thyroid surgery. The main trunk of the vagus may incur injury during neck dissection if the contents of the carotid sheath are not clearly identified or during ligation of the internal jugular vein at the skull base [3]. Therefore the problems related to thyroid disease create a major challenge to the diagnostic capabilities of the clinician after thyroid surgery. In the injury to superior laryngeal nerve leads to dysfunction of cricothyroid muscle, this plays important role in phonation; which in turn results in vocal cord paralysis [3]. Almost 90% of paralyzes are caused by lesions that compress the nerve along peripherally located segments, which reflects the long and redundant routes of the recurrent laryngeal nerves. Specific causes include neoplasm (36%); postoperative complications, such as from parathyroid and thyroid surgery (25%). Direct infiltration of the recurrent or superior laryngeal nerves by local tumor is rare in case of vocal cord paralyzed due to laryngeal cancers [3]. Identifying the course of the RLN is the best way to avoid injury. When RLN arborizes prior to its entrance into the larynx at the cricothyroid joint, injury to a small branch may contribute to the
changes in voice without significant changes in vocal fold motion [22]. But, now it became modern surgery practice to preserve voice without laryngeal nerve injury. Apart from RLN or EBSLN injury, voice changes occur. Removal of the thyroid gland modifies the vascular supply and venous drainage of the larynx. Together with the orotracheal intubation, this could cause alterations of the mucosa as a slight congestion leads to voice changes. Even there may be a fixation of cricothyroid and strap muscles which play important role in proper phonation, would result in voice dysfunction [3].

Figure 1. Thyroid Homeostasis

III. PATIENTS AND METHODS

Six female thyroid patients (between ages 23 to 62 years, mean age = 41± 14 years) who underwent thyroid surgery between November, 2011 and March, 2012 in the department of ENT, JSS hospital, Mysore. The patients had histories of the disease ranging from 1 week to 5 years. Thyroid function (T3, T4, TSH) was in the normal range for patients except one. All patients had no history of throat and vocal fold disease, head and neck surgery, or head and neck trauma. In the control group, 10 subjects (5 males, 5 females, between the ages of 24 and 60 years, mean age 28 ± 05 years) with normal voice function were included. No intraoperative or postoperative injury of the RLN or EBSLN occurred.

A. Voice Acoustics Analysis and Parameters

In the medical field the voice acoustic analysis is used in diagnostic and treatment of voice disorder. The patients are normally reluctant to video laryngoscope (VLS). Moreover VLS has limitations like, irritation to patient, examination also requires topical anesthesia in many patients. The patient's mouth should be open with anterior traction on the tongue, although this position prevents examination of the larynx during speech or swallowing [3]. On the other hand, the stroboscope which provides a flashing light to simulate a slowed vocal-cord vibration, and has a microphone to sense the fundamental frequency of the vibrations and to coordinate the flashing light to the same frequency. If the light flashes at the fundamental frequency, the stroboscope will freeze one image in the vibratory cycle. Thus it will not document the status of vocal function with cycle to cycle information. Moreover, the stroboscope has limitation in patients in whom an unstable voice will not allow for a steady frequency to slow motion. This occurs in patients with vocal cord paralysis or vocal cord scarring and in some patients with spasmodic dysphonia. And stroboscopic parameters include closure pattern, amplitude of vibration, mucosal wave, symmetry, and regularity, and these objective data are visual perceptual ratings and subject to reliability and validity errors.

VLS may not be cost-effective in patients who have subjective voice disturbances. However, it should be used in cases with abnormalities of objective voice measures, or damage seen on indirect laryngoscopy, or both [23]. In regard to above discussion the voice analysis will help the clinicians to document the status of vocal function with cycle to cycle information that is not available with stroboscope [3].

The voice acoustic analysis began with the thyroid patients assuming a natural and comfortable position with the mouth positioned 1 cm from the microphone. The vowel /a/ was produced sustainedly at a comfortable pitch and intensity for 3 to 5 seconds. The signal was digitized, analyzed, and stored using PRAAT software. The analysis indicators included the average fundamental frequency (F0), the F0 standard deviation (STD), the fundamental frequency perturbation (Jitter), the amplitude perturbation (Shimmer), the harmonic/noise ratio (HNR), the voice turbulence index (VTI), and the degree of sub harmonics (DSH), as follows.

Fundamental frequency (F0): Pitch or fundamental frequency is calculated by using autocorrelation method, searching for a peak in normal pitch range in speech of 50Hz to 500 Hz. This estimation is improved by the algorithm to find a period time, interpolated between the sampling instants [10, 11, 12].
Jitter and Shimmer: Jitter and Shimmer are used as indicators of harshness of voice and relative jitter and shimmer are given as a percentage that represents the maximum deviation from a nominal frequency or amplitude [13, 14, 15]. If \( N \) is the total number of Pitch periods, \( T \) be the fundamental frequency and \( A \) the amplitude of particular period then relative jitter and shimer are calculated using “(1)” and “(2)”.

**Harmonic to Noise Ratio (HNR):** HNR is used to assess vocal fold behavior and is closely related to whisperiness, harshness of voice quality [17, 18]. The HNR (dB) is estimated by frequency domain approach. As first cepstral peak correlates with overall voice quality and breathiness, and is directly proportional to average harmonic to noise ratio [16, 19].

**Voice Turbulence Index (VTI):** Voice Turbulence Index is the average ratio of spectral noise energy in high frequency range to the spectral harmonic energy [20]. VTI correlates with incomplete or loose adduction of vocal folds.

**Degree of Sub Harmonics (DSH):** DSH(%) is computed based on the algorithm used to determine pitch by sub harmonic to harmonic ratio (SHR). SHR can also be used to describe degree of irregularity or roughness that reflects the voice quality [21]. According to Titze [5], sub harmonic generation can occur when there is left-right asymmetry in the mechanical properties of the vocal folds.

**Formant Dispersion (Df):** Df(\( \text{Hz} \)) an estimator of vocal tract length, was calculated by finding the average distance between successive formants [26].

**Formant Frequencies:** Formant frequencies are resonant frequencies of air in vocal tract. \( F1, F2 \) and \( F3(\text{Hz}) \) were calculated by using LPC method [27, 28].

The STD, Jitter, Shimmer, HNR, VTI, and DSH were mainly related to the stability of the vocal fold vibration. The value of these indicators increased when the vocal function was abnormal [7]. Voice acoustic analysis was performed preoperatively, postoperatively one week after surgery, one month, 2 months, 3 months subsequently following thyroid surgery. The voice acoustic analysis was performed only once in case of control group.

**B. Statistical Method**

A \( t \) test was performed to analyze and compare data obtained from experimental group before and after surgery and the data from control group, with \( P<0.05 \) [8, 9]. And \( t \) test applied to groups of unequal lengths.

**IV. RESULTS**

The voice acoustic analyses of thyroid patients before and after surgery and of patients in the control group are shown in Table I. In comparison to the control group, the preoperative F0 levels in the thyroid group decreased significantly \( (P<0.05) \), whereas the other indexes were significantly higher than those of the control group \( (P<0.05) \). One week after surgery F0 and STD increased in comparison to preoperative levels \( (P<0.05) \). One month after surgery, the F0 value returned to the level observed preoperatively, but it was still lower than the control group level. The postoperative relative Jitter (%) has shown slight increase and relative Shimmer (%) decrease without much deviation from control group. VTI and DSH levels are significantly decreasing from one month after surgery, and are higher than those of control group \( (P<0.05) \). HNR is slightly increased from preoperative levels and higher than control group \( (P<0.05) \). Three months after surgery, all the acoustic indicators returned to normal.

With formant frequency analysis, \( F1 & F2 \) after one month after of surgery have shown significant difference \( (P<0.05) \) in comparison with control group as shown in Table I.

**IV. DISCUSSION**

Abnormal objective voice acoustic analyses are more common in women [3,4,7]. As it is observed from Table I, that F0 will significantly decreased but returned to normal after 3 months. The cricothyroid (CT) muscle, when activated, elongates the vocal fold thus tensing the cover and elevating the frequency of vibration. Fine control of the degree of tension is accomplished by balancing these CT contraction forces.
against thyroarytenoid, lateral cricoarytenoid, and interarytenoid muscle forces to maintain the vocal folds in an appropriate position for phonation. Unopposed CT muscle contraction will lead to increases in the glottal width, which will negatively affect the vibratory cycle [3]. The author believes that the decrease in F0 levels was related to decreases in the vocal folds contributing to a lower pitch; and (2) the human voice and acoustic detection can be affected by emotion and behavioral conditions. Lower pitch may be caused by patients being unwilling to exert their voices for fear of impacting the wound. (3) And having secure and safe airway is important in patients undergoing surgery. Orotracheal intubation is one of the most widely used techniques and devices used in clinical practice to secure an airway. This method maintains airway patency by way of manipulation of the structures of the upper airway. Success in maintaining a patent airway using this technique requires a detailed understanding of their interaction with the structures of the upper airway. The failure to achieve this goal inevitably results in hypoxic brain injury or death. Thus, together with the orotracheal intubation, the removal of the thyroid gland modifies the vascular supply and venous drainage of the larynx; this could cause alterations of the mucosa as a slight congestion that may help to explain the lower pitch [3]. (4) The strap muscles extending below the larynx function as accessory muscles of

<table>
<thead>
<tr>
<th>Analysis Indicators</th>
<th>Before Surgery</th>
<th>One Week after Surgery</th>
<th>One Month after Surgery</th>
<th>Two Months after</th>
<th>Three Months after</th>
<th>Control Group (n=10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>f0 (Hz)</td>
<td>425.76±68.14</td>
<td>444.20±38.93</td>
<td>425.76±8.87</td>
<td>413.46±24.40</td>
<td>402.79±65.27</td>
<td>447.48±52.23</td>
</tr>
<tr>
<td>STD f0 (Hz)</td>
<td>141.93±24.45</td>
<td>146.41±2.07</td>
<td>165.50±33.58</td>
<td>159.38±18.20</td>
<td>157.94±34.43</td>
<td>146.06±29.68</td>
</tr>
<tr>
<td>Relative Jitter (%)</td>
<td>0.37±0.12</td>
<td>0.38±0.12</td>
<td>0.44±0.11</td>
<td>0.50±0.13</td>
<td>0.45±0.14</td>
<td>0.36±0.18</td>
</tr>
<tr>
<td>Relative Shimmer (%)</td>
<td>0.07±0.02</td>
<td>0.07±0.01</td>
<td>0.04±0.01</td>
<td>0.04±0.01</td>
<td>0.05±0.02</td>
<td>0.06±0.02</td>
</tr>
<tr>
<td>VTI</td>
<td>0.03±0.05</td>
<td>0.01±0.01</td>
<td>0.52±1.27</td>
<td>0.00±0.00</td>
<td>0.00±0.00</td>
<td>0.00±0.01</td>
</tr>
<tr>
<td>HNR (dB)</td>
<td>0.10±0.02</td>
<td>0.11±0.01</td>
<td>0.12±0.02</td>
<td>0.11±0.01</td>
<td>0.11±0.02</td>
<td>0.09±0.02</td>
</tr>
<tr>
<td>DSH (%)</td>
<td>0.21±0.07</td>
<td>0.21±0.10</td>
<td>0.20±0.12</td>
<td>0.19±0.06</td>
<td>0.19±0.09</td>
<td>0.15±0.10</td>
</tr>
<tr>
<td>Df (Hz)</td>
<td>816.10±92.63</td>
<td>531.66±471.85</td>
<td>441.04±835.39</td>
<td>316.01±751.43</td>
<td>638.10±638.35</td>
<td>592.83±427.32</td>
</tr>
<tr>
<td>F1 (Hz)</td>
<td>712.02±274.78</td>
<td>1353.78±1256.42</td>
<td>2568.04±1822.75</td>
<td>2079.44±2032.13</td>
<td>1565.79±1727.83</td>
<td>926.99±721.18</td>
</tr>
<tr>
<td>F2 (Hz)</td>
<td>1606.58±241.35</td>
<td>1784.58±559.20</td>
<td>3330.66±2033.50</td>
<td>2452.67±1003.30</td>
<td>2348.77±1518.63</td>
<td>1553.60±902.32</td>
</tr>
<tr>
<td>F3 (Hz)</td>
<td>2344.22±320.30</td>
<td>2417.09±375.478</td>
<td>3450.13±2072.48</td>
<td>2711.46±1064.52</td>
<td>2841.99±1471.40</td>
<td>2112.67±1221.46</td>
</tr>
</tbody>
</table>

Inspiration in varying degrees. Cervical strap muscle activity affects glottic function, and downward traction on the trachea with inspiration results in vocal fold abduction, and play an important regulatory role in phonation. The removal of the thyroid gland results in fixation of the strap muscles and larynx into a form of complex with reduced laryngeal elevation[3,24,25].

The F0 levels of thyroid patients were significantly lower than the control group levels, and the indexes for the stability of vocal fold vibrations (Jitter, Shimmer, VTI, DSH) were significantly higher than the normal control group. This demonstrates the presence of an abnormal voice in the experimental group. But, as VTI
and DSH levels are significantly decreasing from one month after surgery, will indicate disappearance of asymmetrical vocal fold structure and normal returning of vocal folds after 3 months of surgery [5]. As there is only a significant increase of STD after one month followed by slight increase of Jitter, HNR will indicate that vocal function was not much abnormal. And is correlating with present data that only 33% patients were complaining in 3 months follow up with slight postoperative voice symptoms. One month after surgery, the F0 level in female patients increased slightly, although it was still lower than the preoperative and control group levels; however, other postoperative acoustic indexes decreased from the preoperative value(VTI and DSH), indicating an improvement in the postoperative voice of the women. This may be related to ossification of thyroid cartilage in females. The larynx of females is soft because of incomplete ossification; it is thus easily subjected to mechanical effects before and after surgery. The large preoperative thyroid nodules imposed a significant compression on the thyroid cartilage and compression is progressive. After surgery, the compression was removed, and the larynx and trachea displayed improved vibratory activity. Thus voice quality also improved.

The formant analysis has been added for the estimation of vocal tract length. The formant dispersion is decreased after surgery indicating longer vocal tract length. This is correlated with decreased formant frequencies as shown in Table I.

V. CONCLUSION

In short, the preoperative voice acoustic measures were abnormal in female thyroid patients. Performing routine voice acoustic analyses has proved to be especially important in female patients, and it can be helpful in diagnosing preoperative voice disorders without obvious subjective symptoms and and recognizing potential postoperative voice changes. Although the formant frequencies showed no changes in this study, formant frequency analysis would appear to be a theoretically sound method for assessing change in the vertical position of the larynx in the vocal tract. The results indicate that further investigation using a larger sample size is merited.

ACKNOWLEDGEMENT

Authors would like to thank Dr. Basavaraj S, Dr. Prakash B G, Dr. Bharathi M B, Dr. Sandhya of ENT department, JSS Medical College, Mysore, for their contribution to this study.

REFERENCES


Peter J. Murphy and Olatunji O. Akande, “Quantification of glottal and voiced speech harmonic-to-noise ratios using cepstral-based estimation”, Department of Electronic and Computer Engineering University of Limerick, Limerick, Ireland.


Xuejing Sun, “Pitch determination and Voice quality analysis using Sub harmonic to Harmonic ratio”, Northwestern University, USA.


