Perceptual Evaluation of Voice in Patients with Thyroid Disease

Lisa Allison Melnick

Indiana University of Pennsylvania

Follow this and additional works at: http://knowledge.library.iup.edu/etd

Recommended Citation
http://knowledge.library.iup.edu/etd/443

This Thesis is brought to you for free and open access by Knowledge Repository @ IUP. It has been accepted for inclusion in Theses and Dissertations (All) by an authorized administrator of Knowledge Repository @ IUP. For more information, please contact cclouser@iup.edu, sara.parme@iup.edu.
Indiana University of Pennsylvania
The School of Graduate Studies and Research
Department of Education and Educational Technology

We hereby approve the thesis of

Lisa Allison Melnick

Candidate for the degree of Master of Science

____________________
Lori Lombard, Ph.D., CCC-SLP
Professor,
Chair, Thesis Committee

____________________
Jill Brady, Ph.D., CCC-SLP
Assistant Professor

____________________
David Stein, Ph.D., CCC-SLP
Professor

ACCEPTED

____________________
Timothy Mack, Ph.D.
Dean
School of Graduate Studies and Research
The thyroid gland is positioned directly below the larynx in the anterior portion of the neck (Kumrow & Dahlen, 2002). The nerves and arteries of the thyroid and larynx are intertwined thus damage to one structure could easily affect the other structure. Thyroid dysfunction may also affect the histology of the larynx thus changing vocal quality (Stemple, 2010).

This study measured vocal quality of patients with thyroid disease using the Consensus Auditory Perceptual Evaluation of Voice (CAPE-V). Voice samples from 16 subjects with thyroid disease were compared to an age-matched group of 18 control subjects with no reported history of thyroid dysfunction. An experienced voice clinician rated each sample on the six parameters of the CAPE-V: overall severity, roughness, breathiness, strain, pitch and loudness.

Statistical analysis of the results revealed a significant difference between the thyroid disease group and the control group for overall severity of vocal quality and vocal roughness. These results further strengthen the connection between thyroid disease and vocal dysfunction. Further research is warranted to explore the specific thyroid diagnoses that relate to vocal dysfunction, as well as to strengthen the findings on a larger population of participants.
ACKNOWLEDGEMENTS

Thank you to the following people who so willingly gave their time and talent throughout the progression of this thesis.

Dane Alabran
Jill Brady
Katelyn Fleishman
Hanna Gratzmiller
Seana Hollingsworth
Kayla Kirkpatrick
Lori Lombard
Anne Rhoades
Jill Schager
David Stein
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. REVIEW OF LITERATURE</td>
<td>1</td>
</tr>
<tr>
<td>Thyroid Disease</td>
<td>1</td>
</tr>
<tr>
<td>Incidence/prevalence</td>
<td>1</td>
</tr>
<tr>
<td>Mortality</td>
<td>2</td>
</tr>
<tr>
<td>Common treatment techniques</td>
<td>2</td>
</tr>
<tr>
<td>Symptoms</td>
<td>4</td>
</tr>
<tr>
<td>Related thyroid disease</td>
<td>4</td>
</tr>
<tr>
<td>Anatomy and physiology</td>
<td>5</td>
</tr>
<tr>
<td>Thyroid gland</td>
<td>5</td>
</tr>
<tr>
<td>Larynx</td>
<td>6</td>
</tr>
<tr>
<td>Impact of aging on the larynx</td>
<td>8</td>
</tr>
<tr>
<td>Impact of thyroid disease on perceptual voice quality</td>
<td>9</td>
</tr>
<tr>
<td>Measure of Perceptual Voice Quality (CAPE-V)</td>
<td>9</td>
</tr>
<tr>
<td>Study of Voice and Thyroid disease</td>
<td>11</td>
</tr>
<tr>
<td>Studies of Voice Changes following Thyroidectomy</td>
<td>13</td>
</tr>
<tr>
<td>Interpretive value of studies of voice and thyroid disease</td>
<td>16</td>
</tr>
</tbody>
</table>
2. DEFINITION OF PROBLEM....................................................................................18

Study purpose .......................................................................................................19

Research question................................................................................................19

3. METHOD ..............................................................................................................20

Design ...................................................................................................................20

Participants ...........................................................................................................20

Subject selection ..................................................................................................20

Inclusion and exclusion criteria ............................................................................21

Final sample ..........................................................................................................22

Instrumentation ....................................................................................................22

Procedures ............................................................................................................23

Ethical Use of Data ...............................................................................................24

Statistical Analysis ...............................................................................................24

4. RESULTS .............................................................................................................25

5. DISCUSSION .......................................................................................................27

Conclusion ............................................................................................................28

References ............................................................................................................29
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Comparison of Variables between Experimental and Control Groups</td>
<td>26</td>
</tr>
<tr>
<td>2 Intraclass Correlation Coefficients (ICC) analysis</td>
<td>26</td>
</tr>
</tbody>
</table>
CHAPTER 1
REVIEW OF LITERATURE

Thyroid Disease

Incidence/prevalence

Thyroid nodules are becoming an increasingly prevalent diagnosis. In fact, palpable nodules occur in 3-5% of healthy adults (Grebe, 2009). The true occurrence of thyroid nodules is much more common when considering non-palpable nodules. Current ultrasound studies found a prevalence of 19 to 46% in normal functioning thyroid glands (Kang et al., 2004). Furthermore, autopsy studies have found nodules in greater than 50% of subjects with clinically healthy thyroid glands (Grebe, 2009; Ross, 2002). Other studies estimate that by 80 years of age, 80% of people will have at least 1 nodule (Galofré, Lomvardias, & Davies, 2008). Although nodules are common to a very large population, they are most common among women and the elderly population (Galofré, et al., 2008).

Nodules are classified by appearance and cytology. When pertaining to appearance, nodules may be solid, cystic, single, or multiple (Galofré, et al., 2008). As for cytology, classification is not solely benign or malignant. Nodules can also be classified as follicular neoplasm, inconclusive, or suspicious for carcinoma (Galofré, et al., 2008). Although most nodules are benign, between 7% and 29% will become malignant (Grebe, 2009; Kang, et al., 2004; Ross, 2002).
Mortality

The incidence of thyroid malignancy has increased from 7,800 cases in 1974 to 33,550 cases in 2007 (Grebe, 2009). Women have a much higher incidence of thyroid carcinoma at 25,480 cases when compared to an incidence of 8,070 cases for men in 2007. The prevalence for the disease was approximately 388,386 cases in 2006, making thyroid disease the most common endocrine malignancy (Grebe, 2009). The disease is estimated to be the 13th most common cause of new malignancies in the United States (Grebe, 2009). However, the mortality rate for the disease remains quite low. Fewer than 4% of all new patients are estimated to die within 10 years of the diagnosis (Grebe, 2009). Correspondingly, the percentage of deaths per year due to thyroid cancer is 4.56% (Grebe, 2009). When comparing to deaths from other types of cancer, thyroid carcinoma only accounts for approximately 0.23% of all cancer related deaths (Ross, 2002). The disease is considered as more of a chronic condition rather than an immediate cause of death (Grebe, 2009).

Common treatment techniques

Management and treatment for thyroid nodules depends on the individual characteristics of the nodule and their effects on the body. Thyroid function tests are often the first step in the management of thyroid disease. Thyroid function tests are conducted to measure the amount of Thyroid Stimulating Hormone or TSH (Galofré, et al., 2008). TSH levels indicate whether the thyroid is functioning properly and can guide physicians to appropriate imaging techniques.
Ultrasound is the most common visualization method but other possibilities include CT scanning, MRI scanning, scintigraphy, and positron emission topography (Galofré, et al., 2008). In the United States, current standards for a full evaluation require a nodule to exceed 1 cm in size. Suspicious nodules are biopsied and then classified as either benign, malignant, follicular neoplasm, inconclusive, or suspicious for carcinoma (Galofré, et al., 2008).

Treatment options for thyroid disease vary based upon the severity of the disease. Benign asymptomatic nodules may only require follow-up monitoring and TSH testing, whereas suspected malignancies require surgery (Galofré, et al., 2008). Aggressive malignancies may require chemotherapy and radiotherapy following excision (Ramirez et al., 2007). Surgery is also indicated for toxic nodules, recurrent cysts, and large goiters. The most common treatments for thyroid disease are radioiodine ablation and hormone treatments with thyroxine. Ethanol injections and laser therapy may also be utilized (Galofré, et al., 2008).

Surgery is the most effective treatment for thyroid nodules. A complete thyroidectomy is often utilized due to the risk of recurrence that is associated with an incomplete thyroidectomy. Although patients may experience relief from both symptoms of and concerns about malignancies, complications can occur (Galofré, et al., 2008). The most common complications are damage to the recurrent laryngeal nerve and permanent hypoparathyroidism (Galofré, et al., 2008). Damage to both the recurrent and superior laryngeal nerves can occur due to the close proximity and interconnectedness of the larynx and thyroid gland, resulting in temporary vocal fold dysfunction or vocal fold paralysis (Sulica
Following thyroidectomy, between 1.0 and 5.1% of patients report temporary vocal fold dysfunction (Sulica & Blitzer, 2006). The incidence of permanent vocal fold paralysis ranges from zero to 5.8% (Sulica & Blitzer, 2006).

**Symptoms**

Thyroid nodules are often asymptomatic. Some people may complain of a growing lump in their neck. Others may experience difficulty or discomfort when swallowing (Galofré, et al., 2008). Symptoms more representative of thyroid function such as hyperthyroidism or hypothyroidism may occur.

**Related thyroid disease**

Related thyroid diseases such as hypothyroidism and hyperthyroidism are much more common than nodules and cancer. Again, both conditions are more common in women (McGrogan, Seaman, Wright, & De Vries, 2008). The etiologies of these diseases may be related to environmental factors, pregnancy, viruses, radiotherapy, inadequate iodine intake, or autoimmune responses (McGrogan, et al., 2008). Often, thyroid nodules can be related to both conditions (Galofré, et al., 2008).

A high TSH level indicates hypothyroidism (Galofré, et al., 2008). Hypothyroidism or underactive thyroid hormone production reduces tissue fluids throughout the body (Stemple, 2010). This loss of fluid may affect the health and physiology of hair, skin, and organs (Stemple, 2010). With respect to vocal symptoms, the larynx often loses fluid in the lamina propria, as the deeper layers experience edema. Due to these physical changes, patients may experience voice changes such as lower pitch, roughness, decreased range, and vocal
fatigue (Stemple, 2010). Another related symptom that patients experience may be a persistent dry cough.

Hyperthyroidism, or overactive thyroid hormone production, is indicated by low TSH levels (Galofré, et al., 2008). This causes physical symptoms such as weight loss, tremor, tachycardia, and goiter (Stemple, 2010). An overactive thyroid may cause a patient to experience vocal symptoms such as a shaky voice, breathy vocal quality, and reduced intensity. This condition is also normally treated with pharmacological intervention (Stemple, 2010). Some patients initially present with hyperthyroidism, and then later develop hypothyroidism. This condition is known as Hashimoto’s thyroiditis (Stemple, 2010).

**Anatomy and physiology**

**Thyroid gland**

The thyroid gland is positioned directly below the larynx in the anterior portion of the neck (Kumrow & Dahlen, 2002). This butterfly-shaped organ is fixed to both the larynx and the trachea via fibrous tissue (Kumrow & Dahlen, 2002; Shagam, 2001). The gland consists of two lobes connected by a small band of tissue known as the isthmus (Kumrow & Dahlen, 2002; Shagam, 2001). The lobes of the thyroid protrude anterior and laterally on either side of the trachea (Kumrow & Dahlen, 2002). The posterior wall of the thyroid lobes houses four parathyroid glands (Shagam, 2001). The main function of the thyroid
is to secrete thyroxine, triiodothyronine, and calcitonin, which control the metabolic rates of almost every cell in the body (Kumrow & Dahlen, 2002).

The thyroid gland is closely related to the larynx via both innervation and blood supply. Blood is supplied through four and in some cases five arteries. The gland is innervated by the superior laryngeal nerves and the recurrent laryngeal nerves (Kumrow & Dahlen, 2002). The recurrent laryngeal nerve courses very closely with the inferior thyroid artery (Sulica & Blitzer, 2006). The exact relationship between the two structures is debatable as some report the recurrent laryngeal nerve to course superficial to the inferior thyroid artery; while others report the recurrent laryngeal nerve to course either deep to or between the branches of the artery (Sulica & Blitzer, 2006). Due to these contrasting reports, the precise location of the nerves and arteries are thought to vary from person to person (Sulica & Blitzer, 2006). However, the nerves and arteries of the thyroid and larynx are intertwined, thus damage to one structure could easily affect the other structure.

**Larynx**

The larynx is located inferior to the vocal tract and superior to the trachea (Stemple, 2010). It is composed of muscles, cartilages, connective tissue and mucosa which are all suspended from the hyoid bone (Stemple, 2010). The larynx contains three large cartilages (i.e. epiglottis, thyroid, cricoid) and three smaller paired cartilages (i.e. arytenoids, corniculates, cuneiforms). The musculature of the larynx includes several extrinsic muscles located both above and below the hyoid. The five intrinsic muscles of the larynx are the cricothyroid,
thyroarytenoid, lateral cricoarytenoid, interarytenoid, and posterior cricoarytenoid (Stemple, 2010). Blood is supplied from the superior thyroid artery, superior laryngeal artery, and the inferior laryngeal artery (Stemple, 2010).

The histological structure of the larynx consists of five layers including the epithelium, superficial lamina propria, intermediate lamina propria, deep lamina propria, and vocalis muscle (Stemple, 2010). The epithelium requires mucus lubrication to oscillate properly. Hydration of the remaining layers of the larynx is also important for optimum vocal quality. Thyroid dysfunction may decrease the fluid in the lamina propria, thus changing vocal fold vibration (Stemple, 2010).

The larynx shares much of its innervation with the thyroid gland. Both are innervated by branches of cranial nerve (X), the vagus. The two vagus branches involved are the superior laryngeal nerve and the recurrent laryngeal nerve. The internal branch of the superior laryngeal nerve is responsible for sensation in the larynx above the vocal folds (Kumrow & Dahlen, 2002). The external branch of the superior laryngeal nerve controls the motor function of the cricothyroid muscle in the larynx (Stemple, 2010). Correspondingly, the recurrent laryngeal nerve provides sensation to the larynx below the vocal folds and provides motor innervation to the remaining intrinsic laryngeal muscles (i.e. thyroarytenoid, posterior cricoarytenoid, lateral cricoarytenoid, and interarytenoids) (Stemple, 2010).

The larynx serves three main functions. First, the structure provides airway preservation via opening and closing during respiration (Stemple, 2010).
Second, it prevents particles from entering the airway by closing during swallows. Third, it is the source of phonation via vocal fold vibration (Stemple, 2010).

Phonation is controlled mainly through the muscles of the larynx. The main function of the extrinsic laryngeal muscles is to adjust the position of the larynx in the neck (Stemple, 2010). The intrinsic muscles of the larynx are responsible for altering the tension, length, and mass of the vocal folds through a series of muscular movements (Stemple, 2010). These movements consist of adduction, abduction, and tension (Stemple, 2010). Thus, the coordinated movements of the extrinsic and intrinsic muscles alter the pitch, quality, and intensity of phonation (Stemple, 2010).

**Impact of aging on the larynx**

As the body ages, the larynx begins to deteriorate and atrophy. The vocal folds become looser and thinner, which decreases tissue bulk (Stemple, 2010). These changes to the folds result in decreased pitch and loudness range (Stemple, 2010). Women experience a 10-15 Hz drop in fundamental frequency following menopause, which is perceptually experienced as lower pitch of voice (Linville, 2004, October 19). Correspondingly, men also experience a 10 Hz drop in fundamental frequency between young adulthood and middle age (Linville, 2004, October 19). However, the fundamental frequency of male voices increases approximately 35 Hz with a perceptually higher pitch as they reach old age (Linville, 2004, October 19). Furthermore, poor perceptual vocal quality and vocal endurance have also been observed with the aging process (Stemple, 2010). Indications of poor vocal quality due to aging include objective measures
of increased jitter, shimmer, breathiness, hoarseness, and tremor (Linville, 2004, October 19). Loss of endurance is related to decreased airflow caused by the weakening of the respiratory muscles and rigidity of the thorax (Linville, 2004, October 19). These physical changes are associated with changes in perceived vocal quality.

**Impact of thyroid disease on perceptual voice quality**

Thyroid disease may or may not have a noticeable impact on laryngeal functioning. If a patient develops a massive tumor, compression of the vocal folds results in dysphonia, stridor and dyspnea (Ramirez, et al., 2007). When cervical nodes metastasize, vocal cord paralysis can occur (Ramirez, et al., 2007). However, thyroid nodules are often asymptomatic and do not produce any noticeable changes to the larynx (Galofré, et al., 2008). Related diseases that often co-occur with nodules such as hypothyroidism and hyperthyroidism may impact the laryngeal mechanism. The loss of fluid to the larynx which occurs during hypothyroidism results in the perception of lower pitch, roughness, decreased vocal range, and patient perceptions of vocal fatigue (Stemple, 2010). Hyperthyroidism changes perception of voice stability, breathiness, and loudness. (Stemple, 2010).

**Measure of Perceptual Voice Quality (CAPE-V)**

“The Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) is a clinical and research tool developed to promote a standardized approach to evaluating and documenting auditory-perceptual judgments of voice quality”. (Kempster, Gerratt, Verdolini Abbott, Barkmeier-Kraemer, & Hillman,
The CAPE-V measures three vocal tasks: sustaining the vowels /a/ and /i/ three times each, reading six different sentences containing differing phonemic contexts, and responding to a standard question in a natural conversational style (Kempster, et al., 2009). These tasks are measured by six different voice quality features: overall severity, roughness, breathiness, strain, pitch and loudness. The six quality measures are documented by the listener (i.e. voice expert) via a 100 mm visual analogue scale. The furthest point on the left of the scale (0/100 mm) represents normal voice whereas the right end of the scale (100/100 mm) represents the most deviant voice pattern (Kempster, et al., 2009).

The CAPE-V has good empirical validity when compared to the GRBAS (i.e. Grade, Roughness, Breathiness, Asthenia, Strain) protocol, a well-established method for measuring perceptual features of voice (Zraick et al., 2011). As established by Zraick et al., the interrater reliability of the CAPE-V is slightly higher than that of the GRBAS (2011). However, both measures’ reliability coefficients fluctuate across parameters and raters (Zraick, et al., 2011). The CAPE-V has also been established as being more sensitive to small differences among patients when compared to the GRBAS (Karnell et al., 2007). In a study measuring postthyroidectomy parameters of voice, professionals with experience in voice disorders were found to use the CAPE-V both reliably and similarly (Helou et al., 2010). Thus, the CAPE-V is currently considered a reliable and valid measure of the perceptual qualities of voice.
Auditory-perceptual measures of voice are common and widely accepted among voice experts for use in clinical practice and research. Recent findings have justified the use of the CAPE-V for evaluating and documenting the auditory-perceptual qualities of voice (Zraick, et al., 2011). Establishing the qualities of voice measured by the CAPE-V is necessary to document deviations of voice quality as experienced by the listener.

**Study of Voice and Thyroid disease**

Current research has sought to measure changes in patients with thyroid disease. Voice related outcomes were reported post-operatively in patients who received thyroid surgery (McIvor, Flint, Gillibrand, & Morton, 2000). Fifty consecutive patients who received treatment for thyroid surgery in Green Lake Hospital were initially enrolled in the study. Forty-four patients completed the study, for which no exclusion criteria were mentioned. Patients were tested prior to surgery and postoperatively with the Visipitch oscilloscope, an objective tool used for voice analysis, which measures pitch and amplitude over time. Patients were also given a questionnaire following surgery. As stated by McIvor et al., patients were considered to be “worse” if there was a decrease in maximum phonation time more than 2 seconds, an increase in perturbation greater than .05%, or a decrease in percentage voiced by 2%. “Improvement” was judged by using the same guidelines in the opposite direction (McIvor, et al., 2000). Results concluded that 26 of 44 patients had no subjective postoperative voice change (McIvor, et al., 2000). Objective assessment found that 17 were the same, 13 were better, and six were worse. Preoperatively, 15 patients were
reported to have voice abnormalities. Six of the 15 patients improved post-operatively. According to McIvor et al. (2000), “Patients may have voicing abnormalities before thyroid surgery is performed. Surgery may improve or worsen the voice irrespective of the pre-operative voice status” (p. 179).

There were several shortcomings in the McIvor et al. (2000) study. The measures used in the study had limitations. The Visipitch cannot obtain information on breath control, scarring, and laryngeal movement (Mclvor, et al., 2000). In addition, the post-operative measure was a non-validated questionnaire (Mclvor, et al., 2000). Further, the investigators failed to utilize any perceptual measures of voice. Perceptual measures are often a key component of classifying vocal quality as deviant. Thus, the measurements lacked in both scope of measure and validity.

The McIvor et al. (2000) study failed to utilize any type of inferential statistics; therefore any changes to the voice samples pre or post operatively could not be categorized as being significantly different. Although the descriptive statistics used included clearly defined parameters (change in maximum phonation time of more than 2 seconds, increase in perturbation of more than 0.5%, or a reduction in percentage voiced by at least 2%), the reasoning behind the chosen parameters was not provided. Furthermore, the results were presented simply as: better, worse, or no change. These data fail to mention which specific criteria (maximum phonation time, perturbation, or percentage voiced) differed from the preoperative data.
Another limitation in the McIvor, et al. (2000) study was the absence of a control group. There is no accurate way to measure whether or not voice quality is abnormal without age and gender matched normative data. Although 15 patients’ Visipitch values measured outside of the normal Visipitch range preoperatively, this may not deviate from the normal population. A sample of age-matched controls could very well contain the same amount of Visipitch deviant voice ranges.

**Studies of Voice Changes following Thyroidectomy**

Several studies have examined voice changes in patients following thyroidectomy. Stojadinovic et al. (2008) evaluated patient-reported and clinician-determined voice assessments for identifying postthyroidectomy dysphonia. Patients were evaluated using patient-reported symptoms (Voice Case History [VCH]), patient-perceived voice handicap (Voice Handicap Index [VHI]), clinician-perceived voice deficits (CAPE-V) and video laryngoscopy (VLS) (Stojadinovic et al., 2008). The purpose of the study was, “To examine the utility of patient-reported and clinician-determined voice assessment in identifying postthyroidectomy voice dysfunction.” (Stojadinovic, et al., 2008, p. 732). Fifty military health beneficiaries scheduled to undergo thyroid resection were included in the study. Persons with preoperative laryngeal dysfunction or prior neck/thyroid surgery were excluded. Participants were evaluated preoperatively and 3 times postoperatively (1-2 weeks, 3 and 6 months). Fisher’s exact test was utilized for categorical variables and the Wilcoxon rank-sum was performed for ordinal data. Spearman’s rho correlation coefficient was calculated to
measure other associations between variables (Stojadinovic, et al., 2008). Results indicated that eight patients had early transient voice dysfunction and one had permanent postoperative voice dysfunction. The VCH had high negative predictive value (NPV) and low positive predictive value (PPV). Both the CAPE-V and the VHI had high NPV and PPV; the VHI being slightly higher. The study concluded that patient self-assessment via VHI identifies voice dysfunction after thyroidectomy. The CAPE-V also identified voice change after thyroidectomy; however, it is more difficult to administer due to the necessity of an experienced voice expert to rate samples.

The study determined that both the CAPE-V and VHI were good measures for identifying dysphonia after thyroid resection. Eight patients had early transient voice dysfunction, and one had permanent voice dysfunction following thyroidectomy (Stojadinovic, et al., 2008).

The Stojadinovic et al. (2008) study had some limitations. The small number of patients with postoperative dysphonia was a limitation to the study. Further, the study did not use a control group to compare patients undergoing other types of surgeries. Moreover, including the patients experiencing dysphonia preoperatively (thereby eliminating exclusion criteria of laryngeal dysfunction) could have added an interesting dimension to this study, and perhaps provided further insight into the role of surgical thyroidectomy on voice function. The need to eliminate those experiencing dysphonia implicates the presence of a cohort containing both dysphonia and thyroid disease. This strengthens the assertion that thyroid dysfunction impacts laryngeal physiology.
DePedro Netto et al. (2006) also studied voice quality and vocal self-assessment after thyroidectomy. One hundred patients who underwent thyroidectomy and a control group of 30 patients who underwent breast surgery were included in the study (de Pedro Netto et al., 2006). Patients were excluded if they had vocal complaints or previous head and neck surgery prior to thyroidectomy. Patients with postoperative laryngeal alterations were also excluded. Both experimental and control groups were evaluated using three perceptual measures and one set of objective measures: (a) videolaryngoscopic examination, (b) Voice Handicap Index (VHI), (c) GRBAS, and (d) acoustic analysis (MDVP). Both groups were measured before and two weeks after surgery. The McNemar test was used to compare variables before and after surgery. The Chi-square test and the Fisher’s exact test were utilized to verify the associations between independent variables. Wilcoxon matched-pairs signed rank test was also performed to compare acoustic measures. Results were that 28% of the thyroidectomy group showed larynx alterations (via videolaryngoscopic examination) post-operatively while the control group showed no significant change. Using the GRBAS, 29.7% of the thyroidectomy group showed subjective voice changes (excluding those with vocal fold immobility) postoperatively, while no statistically significant changes were found for the control group. Acoustic analysis indicated significantly increased values in voice turbulence index parameter in both groups, although higher values were noted in thyroid group. Also, the VHI registered higher voice complaints in the thyroid group. The study concluded that voice differences are common after
thyroidectomy even when excluding those with vocal fold immobility (de Pedro Netto, et al., 2006).

This study utilized a well-controlled research design with appropriate measures and statistical analyses. In order to isolate the effect of thyroidectomy surgery without vocal fold alterations due to nerve damage, those with altered vocal folds postoperatively were excluded from the study. Subjects with voice disorders prior to surgery were also excluded (de Pedro Netto, et al., 2006). The presence of subjects with voice disorders both prior to surgery and post-surgery supports the assertion that thyroid function may impact voice function. Thus, the incidence of voice disorders associated with thyroid disease was not directly measured in this study, but the connection can be implied.

**Interpretive value of studies of voice and thyroid disease**

Based upon current literature, evidence supports the assertion that thyroidectomy often results in some degree of dysphonia (de Pedro Netto, et al., 2006; McIvor, et al., 2000; Stojadinovic, et al., 2008). This connection implies a strong relationship between thyroid function and laryngeal functioning. The etiology of dysphonia related to thyroid dysfunction remains unclear as studies have controlled for preoperative voice dysfunction, operative nerve damage, and general surgery/intubation. However, the occurrence of postoperative voice dysfunction remains. Studies have found that the CAPE-V, VHI, and GRBAS were good indicators of voice dysfunction following thyroidectomy (de Pedro Netto, et al., 2006; McIvor, et al., 2000; Stojadinovic, et al., 2008). Objective measures did not yield differing results between groups, as de Pedro Netto, et al.
(2006) found increased acoustic measures for both experimental and control
groups.

Most of the studies did not aim to assess vocal symptoms prior to surgery. McIvor, et al. (2000) did assess patients with vocal dysfunction prior to surgery; however the limitations of the study made the results less powerful. The association between thyroid disease and vocal dysfunction prior to surgery remains unclear. Acquisition of documented voice function prior to thyroid surgery is vital for understanding the connection between thyroid disease and dysphonia.
CHAPTER 2
DEFINITION OF PROBLEM

The current literature lacks information on the incidence of voice disorders in those with thyroid disease. Ramirez, et al. (2007), reports the possibility of dysphonia, stridor and dyspnea in those with thyroid tumors. McIvor, et al. (2000) found voice dysfunction in patients with thyroid disease, but the results lacked powerful statistical analysis. Hypothyroidism can cause perceived voice changes such as lower pitch, roughness, decreased range, and vocal fatigue; whereas hyperthyroidism can cause shaky voice, breathy vocal quality, and reduced intensity (Stemple, 2010). The impact of aging on the larynx can also cause perceptual changes in the voice (Linville, 2004, October 19).

The CAPE-V has been shown to be a reliable measure for identifying perceived vocal changes. The ease of use and high predictive value of this measure makes it ideal for measuring voice parameters as experienced by the listener.

If those with thyroid disease often display vocal dysfunction, detection of these symptoms could aid in early detection of thyroid disease. Further, obtaining a baseline for voice dysfunction in those with thyroid disease could facilitate documentation of the recovery process. Moreover, baseline data could help surgeons determine the success of their surgical techniques based on postoperative vocal function.
Study purpose

The purpose of this study is to identify auditory perceptual changes in voice for those with thyroid disease and determine how these changes differ from a control group. Specifically, this study aims to investigate the differences in perceived vocal quality between patients with thyroid disease and the normal population.

Research question

Do patients with thyroid disease differ in perceptual voice quality from an age and gender controlled sample without thyroid disease?
CHAPTER 3

METHOD

Design

This study consisted of a differential group comparative research design. The study compared differences between two preexisting groups of participants, subjects with thyroid disease and subjects without thyroid disease, in a constrained setting. The variables were measured in exactly the same manner for the two age-matched groups. The experimental group’s members had a diagnosis of thyroid disease, whereas the control group did not.

Participants

Subject selection

Potential candidates were identified by co-investigator Dr. Bosco Noronha at the Indiana office of University Ear, Nose, and Throat Specialists (UENTS). After completing a Consent to Participate in a Clinical Research Registry form, participants were given a written description of the study and the investigators’ contact information. Potential subjects indicated their agreement by notifying an investigator and completing the project’s Informed Consent Form. In 2002, the Consent for the Clinical Research Registry was initiated and has been renewed by the University of Pittsburgh Institutional Review Board annually. This registry form is provided to all patients in the UENTS offices. Signing this form provides patients’ consent to release their demographics and medical record information for use in the clinical study database. The subjects’ signatures indicated their
willingness to be contacted for participation in related clinical studies, which facilitated recruitment for this project.  

After contacting an investigator and consenting to the procedures, participants signed Indiana University of Pennsylvania’s Informed Consent Form. A flyer containing basic study and contact information was posted in the waiting room of UENTS office for recruitment of controls. Family members and colleagues who met the same inclusion and exclusion criteria and agreed to the same terms as those recruited through the flyer were also recruited as control participants in order to age-match the two groups.  

**Inclusion and exclusion criteria**

All participants were adults aged between 33 and 79 years. The inclusion criterion for the thyroid disease cohort was a diagnosis of thyroid disease by a physician. No evidence of thyroid disease was required for the inclusion of participants in the normal subject comparison group. The following inclusion criteria applied to participants in both groups: (a) demonstration of functional speech, language, and hearing as judged by the participants’ ability to listen to and follow directions provided for task completion; (b) no history of voice or swallowing disorder; and (c) no history of previous thyroid or laryngeal surgery. Participants who exhibited symptoms of cold or other illness on the day of testing were excluded from the study. Participants considered to be sensitive subjects were also excluded from the study. A participant was considered a sensitive subject if they were a(n): (a) pregnant woman, (b) test subject for new drugs or clinical devises, (c) abortus, (d) individual engaging in illegal behavior, (e)
individual who was incarcerated, (f) individual who was mentally disabled, or (g) an educationally or economically disadvantaged person.

**Final sample**

Thirty-four participants were included in this study. The experimental group consisted of 16 participants with a mean age of 55 years (range = 35-79 years; $SD = 11.0$). The control group consisted of 18 participants with a mean age of 52 years (range = 33-65; $SD = 9.5$). The experimental group’s male to female ratio was 1:3 and the control group’s male to female ratio was 5:4. The gender ratio for the experimental group is to be expected due to the higher incidence of thyroid disease in women.

**Instrumentation**

This study compared the results of the perceptual measure of voice, the Consensus Auditory Perceptual Evaluation of Voice (CAPE-V), in patients with thyroid disease and an age-matched control group with no reported thyroid disease. This measure included completion of the following vocal tasks:

1. Sustained vowels, /a/ and /i/ for 3-5 seconds durations each

2. Sentence production:
   a) The blue spot is on the key again.
   b) How hard did he hit him?
   c) We were away a year ago.
   d) We eat eggs every Easter.
   e) My mama makes lemon muffins.
   f) Peter will keep at the peak.
3. Spontaneous speech in response to “Tell me about your voice problem.” Or “Tell me how your voice is functioning.”

The CAPE-V scoring was completed by a voice expert who is a licensed speech-language pathologist. The clinician was blind to the group status while rating samples. The voice samples were scored on a 100mm visual analogue scale for the following parameters: overall severity, roughness, breathiness, strain, pitch, and loudness. Each parameter was measured and received a numeric score (__/100) (Kempster, et al., 2009).

**Procedures**

Subjects were given written instructions via a handout which contained the three tasks of the CAPE-V. Subjects were given demonstrations for the first task. For the second task, subjects were told to read the sentences as naturally as possible. For the third task, subjects were simply asked to speak about their day for 90 seconds. The standard CAPE-V poses the spontaneous speech probe, “Tell me how your voice is functioning.” For the purpose of this study, the probe was changed to “Tell me about your day yesterday,” so the rater would not be biased by the response.

Data were collected in a private room with the Roland CD2 CF/CD Recorder. The Roland recorder was constructed to record and transfer audio signals to a CD-R/RW drive. An ART20 Audio-technica, Cardioid low impedance microphone was utilized for recording. The microphone’s Cardioid polar pattern reduces pickup of undesired off-axis sounds.
Ethical Use of Data

The investigators did not reveal any identifying information during data collection. No participant’s identifying information will be published or revealed to any agencies. Access to these data was restricted only to the co-investigators responsible for the study. All personal identifying information was stored in a locked cabinet behind two locked doors. The information will be destroyed following completion of a larger series of studies, of which this study is a portion.

Statistical Analysis

The data collected from the CAPE-V were analyzed with SPSS software version 15.0. Data were not normally distributed. Therefore, Mann-Whitney U test was utilized to evaluate the differences between the median scores of the six parameters of the CAPE-V between the experimental and control groups. Significance was determined by a p-value of less than 0.05. Intra-rater reliability was measured with blinded, test-retest conditions for 6 of the subjects.
The six parameters of the CAPE-V were analyzed to determine significance between the two groups (see Table 1). The results revealed a significant statistical difference between the experimental and control groups for the parameters of overall severity ($p = 0.033$) and roughness ($p = 0.031$). The following provides the median scores, first quartiles (1Q), and third quartiles (3Q) for overall severity and roughness: (a) experimental group overall severity ($Mdn = 4$, $1Q = 2.75$, $3Q = 6.26$), control group overall severity ($Mdn = 2$, $1Q = 1$, $3Q = 4.75$); (b) experimental group roughness ($Mdn = 7$, $1Q = 5$, $3Q = 12.3$), control group roughness ($Mdn = 4.5$, $1Q = 2.25$, $3Q = 7.75$).

Intra-rater reliability is reported by the single measures Intraclass Correlation Coefficients (ICC) and single measures p-values (see Table 2). Intra-rater reliability for strain (ICC = .796, $p = .016$), pitch (ICC = .913, $p = .002$) and loudness (ICC = .800, $p = .015$) were high. Intra-rater reliability for overall severity (ICC = .505, $p = .124$) was moderate, and intra-rater reliability for roughness (ICC = .165, $p = .362$) was poor. Intra-rater reliability could not be calculated for breathiness because all values were 0; however, complete agreement between values occurred.
### Table 1

**Comparison of Variables between Experimental and Control Groups**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Experimental Group</th>
<th>Control Group</th>
<th>Mann-Whitney U</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Median</td>
<td>1Q</td>
<td>3Q</td>
</tr>
<tr>
<td>Severity</td>
<td>16</td>
<td>4</td>
<td>2.75</td>
<td>6.25</td>
</tr>
<tr>
<td>Roughness</td>
<td>16</td>
<td>7</td>
<td>5</td>
<td>12.3</td>
</tr>
<tr>
<td>Breathiness</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Strain</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pitch</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Loudness</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

*Groups significantly different at the p<0.05 level*

### Table 2

**Intraclass Correlation Coefficients (ICC) analysis**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Chronbach’s Alpha</th>
<th>Single Measures ICC</th>
<th>Single Measures 95% Confidence Interval</th>
<th>Single Measures p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severity</td>
<td>.671</td>
<td>.505</td>
<td>-.403 -.912</td>
<td>.124</td>
</tr>
<tr>
<td>Roughness</td>
<td>.284</td>
<td>.165</td>
<td>-.673-.818</td>
<td>.362</td>
</tr>
<tr>
<td>Breathiness</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Strain</td>
<td>.886</td>
<td>.796</td>
<td>.104-.969</td>
<td>.016</td>
</tr>
<tr>
<td>Pitch</td>
<td>.954</td>
<td>.913</td>
<td>.508-.987</td>
<td>.002</td>
</tr>
<tr>
<td>Loudness</td>
<td>.889</td>
<td>.800</td>
<td>.115-.969</td>
<td>.015</td>
</tr>
</tbody>
</table>

*Values could not be calculated*
CHAPTER 5
DISCUSSION

This study sought to determine whether patients with thyroid disease differed in perceptual voice quality from an age and gender controlled sample without thyroid disease. The results showed that patients with thyroid disease did differ perceptually in overall severity and roughness when compared to controls. These results suggest that thyroid disease impacts laryngeal functioning. The close proximity of the thyroid to the larynx may be the cause of this dysfunction. Another possibility could be the presence of hypothyroidism in those subjects tested. The loss of fluid to the larynx which occurs during hypothyroidism results in lower pitch, roughness, decreased range, and vocal fatigue (Stemple, 2010). Roughness was a parameter found to be significantly higher in the thyroid group of this study.

The largest limitation of this study was the sample size. A larger number of participants could have yielded different or more powerful results. Due to the small sample size, participants were not grouped into subcategories based on their diagnosis. Subjects were included with various degrees of thyroid dysfunction such as carcinoma, nodules, or associated thyroid diseases. The lack of subcategory distinction could have limited the findings of this study. One or more specific diagnoses may perhaps yield different vocal symptoms. Furthermore, intra-rater reliability was poor for roughness. Higher intra-rater reliability across all variables could also impact the results.
This study is a part of a larger project which aims to measure changes in voice and swallowing for those with thyroid disease. The project is ongoing and will continue to analyze different measures such as acoustic analysis, laryngeal diadochokinesis, the Voice Handicap Index, and the SWAL-QOL. A more robust picture of the impact of thyroid disease on the laryngeal mechanism may be apparent when all portions of the project are combined.

Further research is necessary to attain a complete picture of the relationship between thyroid disease and laryngeal dysfunction. Future endeavors should include studies with larger numbers of participants. Also, the separation of types of thyroid disease is an important feature that should be tested in the future. Furthermore, adding more measures, such as those included in the larger project, should also be applied to future research endeavors.

**Conclusion**

The results of this study concluded that the perceptual measure of voice, the CAPE-V, revealed significant differences between the roughness and overall severity of voice dysfunction between the individuals with thyroid disease and those in an age-matched control group. This study was limited by its sample size, intra-rater reliability, and lack of specific categorization of thyroid disease. Further research is suggested to fully understand the connection between thyroid disease and vocal dysfunction. Understanding this relationship could aid in early detection of thyroid disease, as well as providing a means for documenting recovery from treatment procedures.
References

de Pedro Netto, I., Fae, A., Vartanian, J. G., Barros, A. P. B., Correia, L. M.,
after thyroidectomy. *Head & Neck, 28*(12), 1106-1114.

thyroid nodules: A clinical guide. *Mount Sinai Journal of Medicine: A
Journal of Translational and Personalized Medicine, 75*(3), 299-311.

Grebe, S. (2009). Diagnosis and management of thyroid carcinoma: A focus on
serum thyroglobulin. *Expert Review of Endocrinology and Metabolism,
4*(1), 25-43.

Helou, L. B., Solomon, N. P., Henry, L. R., Coppit, G. L., Howard, R. S., &
Stojadinovic, A. (2010). The role of listener experience on consensus
auditory-perceptual evaluation of voice (CAPE-V) Ratings of
postthyroidectomy voice. *American Journal of Speech-Language
Pathology, 19*(3), 248.

Prevalence, clinical and ultrasonographic characteristics of thyroid
incidentalomas. *Thyroid, 14*(1), 29-33.

Karnell, M. P., Melton, S. D., Childes, J. M., Coleman, T. C., Dailey, S. A., &
and patient-based (V-RQOL and IPVI) documentation of voice disorders.


