ACOUSTIC VOICE QUALITY INDEX (AVQI)

IN TAMIL LANGUAGE

Ms. Vishali.P

Register No: 17SLP040

A Dissertation submitted in Part Fulfilment of

Degree of Master of Science (Speech-Language Pathology)

University of Mysore, Mysuru

All India Institute of Speech and Hearing,

Naimisham Campus, Manasagangothri,

Mysuru - 570 006

MAY, 2019

CERTIFICATE

This is to certify that this dissertation entitled "**Acoustic Voice Quality Index (AVQI) in Tamil Language**" is a bonafide work submitted in part fulfilment for degree of Master of Science (Speech-Language Pathology) of the student Registration Number: 17SLP040. This has been carried out under the guidance of a faculty of this institute and has not been submitted earlier to any other University for the award of any Diploma or Degree.

Mysuru
May, 2019

Prof. M. Pushpavathi
Director
All India Institute of Speech and Hearing,
Manasagangothri, Mysuru-570006.

CERTIFICATE

This is to certify that this dissertation entitled "**Acoustic Voice Quality Index (AVQI) in Tamil Language**" has been prepared under my supervision and guidance. It is also certified that this dissertation has not been submitted earlier to any other University for the award of any Diploma or Degree.

Mysuru
May, 2019

Dr. R. Rajasudhakar
Guide
Reader in Speech Sciences
Department of Speech-Language Sciences
All India Institute of Speech and Hearing,
Manasagangothri, Mysuru-570006.

DECLARATION

This is to certify that this dissertation entitled "**Acoustic Voice Quality Index (AVQI) in Tamil Language**" is the result of my own study under the guidance of Dr. R. Rajasudhakar, Reader in Speech Sciences, Department of Speech-Language Sciences, All India Institute of Speech and Hearing, Mysuru, and has not been submitted earlier to any other University for the award of any Diploma or Degree.

Mysuru **Registration No.:** 17SLP040

May, 2019

Dedicated to

My

Appa (Mr. M. Ponnuchamy),

Amma (Mrs. V. Chandra),

and

Almighty

ACKNOWLEDGMENTS

"Feelings of heat and cold, pleasure and pain, are caused by the contact of the senses with their objects. They come and they go, never lasting long. You must accept them."

― Anonymous, Bhagavad Gita

I sincerely thank **Lord Krishna** & **Lord Ganesha** for giving me strength and courage, to overcome all my difficulties and showering immense blessing on me.

Without you this work would not have been possible.

"No one who does good work will ever come to a bad end, either here or in the world to come"

― Anonymous, The Bhagavad Gita

Yes **Lord Shani Dev**- thank you so much for teaching the art of Truthfulness, Facing Reality And Simplicity. I Miss You Ms. Kanna (The Crow)

Special thanks to my Lord of Thursday (Yellow) - **Sri Dakshinamurthy**, My guru who enlightened me with his *knowledge* and wisdom.

I would express my sincere gratitude to Prof. M. Pushpavathi (Director, AIISH) for permitting me to conduct my dissertation.

I would like to thank my guide **Dr. R. Rajasudhakar** (Reader in Speech Sciences) for his guidance and patience during the course of my dissertation. Sir, it was a wonderful experience working under your guidance. Thank you for being calm even in the nth moment and also for correcting all my documents from A-Z, 1-10, fullstop-comma.

I am very thankful to Mr. Santhosha, Statistician for helping me in finding out the outcome of the study.

‘A gift is pure when it is given from the heart to the right person at the right time and at the right place, and when we expect nothing in return’

— Anonymous, The Bhagavad Gita

I’m dedicating this quote to you akka – Niharika akka

Dr. Niharika, Thank you so much akka for dedicating your precious personal time for me and without your motivation and support, this work would not have been possible. Thank you so much akka, for your constant reminding and also for making me even more punctual in my work. I will never meet anyone with this personality akka. I thank lord Krishna and I believe he sent you to me as a physical being in all my bad times.

“It is better to live your own destiny imperfectly than to live an imitation of somebody else's life with perfection.”

— Anonymous, The Bhagavad Gita

This is for you thangachi

Dr. Sasikaran (thangachi), thank you so much for accepting me for who I’m and listening to all my good/bad time stories and remembering them.

I have no words to thank you **thangachi** right from your motivation to determination; you were there by my side. Thank you so much for all the reinforcements; it boosted me to finish all my work on time.

I convey my special thanks to Ms. Priyadarshini. V for being my Pillayar Suzhi. Akka, I researched this topic with your timely support. I sincerely thank Mr. Jesnu (Sir) and Ms. Seshasri (akka) for teaching the concepts on AVQI.

I’m highly indebted to all the participants for their time and co-operation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Voice is a key aspect of communication between people. It serves as a means of expressing the speech. Voice is produced by lung air that vibrates the larynx's vocal folds. Voice disorders are the most common speech disorders affecting about 6% of children under the age of 14 and 3-9% of the adult population in the UK population. The prevalence of hoarseness was 41% in the age group of 50-79 years whereas 10-29 years age group, it showed fewer incidences (24%) (Rameshkumar & Rosmi, 2017). "Voice disorders occur when the voices of other people differ in quality, pitch, loudness or flexibility from similar age, sex, and cultural group voices" (Casper & Leonard, 2006). In the narrow sense, voice refers to sounds or voiced sounds produced by vocal fold vibration. Vocal fold vibration modulates airflow through the glottis for voice production and produces sound that propagates through vocal tract and at different frequencies is selectively amplified or attenuated. This voice basis spectrum modification produces clear contrasts that are used to express different linguistic sounds and meanings. While this modification is an essential component of voice production, it focuses on the source of voice and its control in the larynx.

The voice contain important information such as pitch, loudness, prosody, and voice quality (Kreiman & Sidtis, 2011), biological information (size), and paralinguistic information such as social status of the speaker, individual characteristics, and emotional state (Sundberg, 1987; Kreiman & Sidtis, 2011). People tend to use diverse voices for different speakers on different occasions, and from the tone of their voice it is often probable to tell if somebody is happy or sad. One of the key objectives of voice research is to know how the vocal system produces the voice of different sources and how people associate perceptions with these characteristics (Kreiman, Gerratt, Garellek, Samlan, & Zhang, 2014). Dejonckere et al. (2001) proposed for all "common" dysphonias a multidimensional set of necessary measurements. It includes five different approaches: (i) perception (ii) video-stroboscopy (iii) acoustics (iv) aerodynamics, and (v) patient subjective rating.

Based on a thorough review of the literature, Dejonckere et al. (2001) establish the practice within the' European Laryngological Society.' The primary aim of a Speech-Language Pathologist (SLP) is to carry out a comprehensive assessment for the person with dysphonia and to make an appropriate diagnosis to aid in successful management. The quality of a voice can be measured qualitatively by listening and quantitatively by using instruments (Hakkesteegt, Brocaar, Wieringa, & Feenstra, 2008). However, in a clinical setup, it is difficult for the clinician to assess all the variety of assessment tools mentioned above. Thus, the evaluation of voice disorders involves a frequent combination of perceptual and acoustic measurements. The advantage of perceptual analysis is that it is easily available for SLP and laryngologists for daily use in their clinical setup. Most widely used perceptual measures include Darley Rating System (Darley, Aroson, & Brown, 1969), GRBAS scale (Hirano, 1981), Buffalo Voice Screening Profile (Wilson, 1987), and Consensus Auditory Perceptual Voice Evaluation (CAPE-V) scale (Kempster, Gerratt, Verdolini, Barkmeier, Kraemer, & Hillman, 2009). While perceptual voice assessment is a gold standard when evaluating a person with dysphonia, it is vulnerable to many variations due to multiple listeners, subjects or task factors.

Acoustic measurements of voice, on the other hand, have been used in the assessment due to their advantage such as non-invasive and affluence of use. Acoustic measurements quantify the degree of severity of dysphonia and to monitor improvements in voice quality with medical or therapeutic management. The acoustic measure has been regarded as the most reliable objective measure of voice quality (Carding, Wilson, Mackenzie, & Deary, 2009). Objective evaluation of acoustic measures includes;

- Frequency-related measures include fundamental frequency, habitual frequency, and frequency range.
- Amplitude-related measures include habitual intensity, extent, and fluctuation of intensity.
- Perturbation related measures include jitter and shimmer. Harmonic related measures include harmonic to noise ratio widely used by various researchers (Hirano, Hibi, Yoshida, Hirade, Kasuya, & Kikuchi, 1988; Rabinov, Kreiman, Gerratt, & Bielamowicz, 1995; Wolfe, Cornell, & Fitch, 1995; Dejonckere & Lebacq, 1996).

Although different parameters have been used to objectively document voice quality, the Dysphonia severity index (DSI) (Wuyts et al., 2000) is a multiparametric measurement reported to be a robust measure in different studies (Timmermans, De Bodt, Wuyts, & Heyning, 2004). DSI considers Maximum Phonation Time (MPT), highest frequency, lowest intensity and jitter to reach a numerical value that reflects the voice quality of a given individual. In a study by Hakkesteegt et al. (2008), it was revealed that the mean DSI scores could be highly stable across subject groups were well correlated with the available perceptual measures. In the Indian context, Jayakumar and Savithri

(2012) developed the normative for DSI and compared it with the European population in terms of Highest F0, MPT as well as the DSI values. The function of increasing the ecological validity of the analysis can serve the acoustic qualities of sustained phonation, as well as the connected speech varies, including a connected speech sample. The acoustic voice quality index (AVQI) (Maryn, Corthals, Van Cauwenberge, Roy, & De Bodt, 2010) is one recently introduced technique for measuring the severity of overall dysphonia involving sustained phonation and connected speech.

Maryn et al. (2010) investigated the feasibility and diagnostic accuracy of combining continuous phonation and speech in an overall voice quality assessment consisting of perceptual and acoustic methods. Sustained phonation and continuous speech samples were collected in this study from 251 subjects including 229 patients with voice disorder and 22 normal vocal individuals. For at least 5 seconds, these participants were asked to sustain the vowel / a/ and read a Dutch text. The samples were trimmed to comprise only the middle 3 seconds and the first two sentences. These samples were rated perceptually by five SLPs who had five years of experience using the GRBAS scale in clinically assessing voice quality and overall severity of dysphonia. Reliability for perceptual assessment was found to be moderate to high intra-rater and fair to moderate reliability between raters. It was found that there was a positive relationship between AVQI and G of GRBAS with a correlation of about 0.78 and higher than the AVQI score, the more disturbed was the overall voice quality and vice versa. Using the receiver operating characteristics (ROC) curve, cut-off point was achieved with the best balance between sensitivity and specificity and provided optimal discrimination between normal and pathological groups. In this study, AVQI cut off score was 2.95 and found sensitivity of 74 percent was reported and specificity of 96 percent. Almost all normal voice quality has been properly classified, but only 74 % of dysphonic patients have been properly classified as such. Likelihood analysis resulted in Likelihood Ratio+ = 19.98 and Likelihood Ratio- = 0.27 for this AVQI cut off score. Authors concluded that continuous speech assessment should also be considered in order to improve ecological validity. The acoustic algorithm must be further refined. The future direction of this study suggests that the validity of acoustic and perceptual analysis results would be increased by longer samples of continuous speech.

Maryn et al. (2014) examined the impact of language on the AVQI, originally built on native Dutch speakers. 12 normo-phonics and 38 dysphonics were requested to read six text samples aloud in four different languages, i.e. English, Dutch, German and French, using a standardized reading passage and to phonate a sustained vowel /a/. Three experienced Dutch voice clinicians rated these samples using GRBAS rating scale. The results revealed that, although it was found to be superior in English and German and less in French, the AVQI measures did not differ significantly across languages. The authors also confirmed the good diagnostic accuracy and cross-linguistic validity of AVQI.

Benoy (2017) developed AVQI reference data and also validated in the Indian context with perceptual measurements between normal voice quality and dysphonia. A sum of 120 people were taken part in the study, 100 of whom were individuals with normal voice quality (50 Malayalam & 50 Kannada speakers) and 20 were individuals with dysphonia. In the age range of 20-35 years as Group I and 36-50 years as Group II, the normo-phonic individuals were divided into two groups. Reading a text and phonating the vowel /a/ were asked. A sentence was edited with 13 syllables and a portion of the sustained vowel phonation in mid-vowel 3-seconds. Both speech and phonation tasks were concatenated and assessed for the severity of dysphonia by five voice clinicians. The grade (G) was rated from the GRBAS protocol. Authors reported that the G scores and AVQI were highly correlated. Gender, language, and age did not affect AVQI. The mean AVQI score was 3.02(±0.33) for the age range of 20-35 years and 3.05(±0.31) for 36-50 years. The mean AVQI score was 4.43 (±0.78) for the age range of 20 –35 years and 5.03 (±1.03 for the age group of 36 –50 years. Author concluded that dysphonic individuals have higher AVQI values than normo-phonic individuals. However, this study was limited to a mild and moderate degree of dysphonics; in order to obtain voice sample for acoustic measurements, participants were also required to be literate. This study paves the way for exploring AVQI in other Dravidian languages and professional voice users.

1.1. Need for the study

Maryn et al. (2010) developed AVQI for European normo-phonics and dysphonic subjects for the first time. In recent years, Acoustic voice analysis has gained attention of researchers due to its objectivity in measuring the voice quality. For some Western European languages (i.e., Dutch, German, English, French, and Finnish) (Maryn, De Bodt, Barsties, & Roy, 2014), one Altaic language group (i.e., Korean) (Maryn, Kim, & Kim, 2016; Kim et al., 2018), and one Indo-European language (i.e., Lithuanian) (Uloza et al., 2016), AVQI's validity across different phonetic structures has been studied. Tamil is the oldest, longest-surviving classical language in the world of the four South Indian languages in the Dravidian family, with examples from the early Common Era. More than 66 million people spoke Tamil in the early 21st century, mostly residing in India (Krishnamurti, 2018).

Similar to other Dravidian languages, Tamil is characterized by a series of retroflex consonants (/ḍ/, /ṇ/, /ṭ/) produced by curling the tip of the tongue backwards and having contact with the mouth. Tamil has no aspirated phonemes. There is no distinction between the Tamil script for voiced and unvoiced stop, though both are present as allophones in the spoken language. There is a complementary distribution of the voiced and unvoiced stops, and the places they may occur do not intersect. For example, at the beginning of the words, the voiceless stop [p] occurs and the voiced stop [b] cannot. There are 12 vowels, 18 consonants in the Tamil language.

However, there is a dearth of studies in the Tamil context using AVQI with normo-phonic as well as dysphonic subjects. Further, physiological variations continuously make an influence in adults in the aging processes, it is essential to establish age and gender-specific normative values for this parameter. Therefore, there is a need to evaluate AVQI on normo-phonic as well as dysphonic Tamil speaking population and to develop reference data in them. Hence, the present study attempted to provide an understanding of the measures of AVQI on normo-phonic subjects in the Tamil speaking population.

1.1. Aim of the study

The objective of this study is to determine the Acoustic Voice Quality Index (AVQI) in native Tamil-speaking normo-phonics adults between 20 to 50 years of age.

1.3. Objectives of the study

1. To establish standard reference data for the Acoustic Voice Quality Index (AVQI) in native Tamil speakers between 20 to 50 years of age.
2. To investigate the effect of gender on the Acoustic Voice Quality Index (AVQI).
3. To investigate the effect of age on the AVQI Index.
4. To measure the validity (Field test) of the AVQI in a few dysphonic patients whose native language is Tamil.

CHAPTER II

REVIEW OF LITERATURE

"Voice" is the sound perceived when the respiratory air stream drives the vocal folds into vibration.

2.1. Assessment of Voice

Speech-Language Pathologists (SLPs) play a crucial part in the assessment, diagnosis, as well as management of voice disorders. SLPs' professional roles and activities include clinical services, advocacy and prevention, education, administration, and research. SLPs conduct initial voice disorder screening depending on individuals, parents, teachers, or health care providers' concerns. When deviations from clinically normal voice during screening are detected, further evaluation is required. The screening procedure includes evaluation of vocal characteristics related to the respiratory system, phonatory system, and resonatory system, as well as vocal range and flexibility (pitch, loudness, pitch range, and endurance). SLPs may use a formal screening tool (Lee, Stemple, Glaze, & Kelchner, 2004) or use informal tasks to obtain data. For more thorough screening, standardized self-report questionnaires may be included (Jacobson, Johnson, Grywalski, Silbergleit, Jacobson, Benninger, & Newman, 1997; Hogikyan & Sethuraman, 1999; Deary, Wilson, Carding, & Mackenzie, 2003).

Following a screening, comprehensive assessment of voice is conducted for individuals suspected of having a voice disorder. Comprehensive assessment of voice uses both standardized and non-standardized measures. Norms are based on age, gender, instrumentation type, cultural background and dialect. These assessments can be subjective or objective; subjective assessments are based on the perceptual competence whereas objective assessments are based on the results of instruments or software used. However, a combination of subjective and objective assessment would fetch a better diagnostic picture of individuals with or without voice disorders.

2.1.1. Subjective assessment

A gold standard clinical voice assessment method, Auditory-perceptual evaluation is commonly used for the documentation of voice related issues. Owing to its subjective nature, perceptual evaluation has been heavily criticized. Consequently, perceptual evaluation is not always accurate and is affected by various listener related factors. The auditory perceptual assessment shows a clear relationship with the underlying physiology of the vocal tract and acoustic parameters due to high intra and inter-rater reliability and agreement. The first tangible outcome of these efforts was the development of the Consensus Auditory Perceptual Evaluation of Voice (CAPE-V; Kempster et al., 2009).

The Consensus Auditory Perceptual Evaluation of Voice (CAPE-V) has been developed as a tool for clinical voice auditory perceptual evaluation. Its primary function is to describe the auditory perceptual attributes and severity of a voice problem so that clinicians can communicate. Its secondary function is to contribute to hypotheses regarding the anatomical and physiological basis of voice problems and evaluate the need for further testing. Over the years, some similar rating scale protocols have been proposed for quality assessment. For example, users of CAPE-V voices on visual analog scales for overall severity, roughness, breathiness, strain, pitch and loudness. The GRBAS protocol evaluates voices on scales for essentially the same qualities as grade, roughness, breathiness, aesthetics and strain, but uses four-point rating scales.

2.1.2. Objective assessment

Objective assessment of voice majorly comprises of aerodynamic, acoustic and laryngeal imaging procedures. Voice production is a result of Aerodynamic activities in respiratory and phonatory systems. The exhaled air acts as a source which activates phonation with a direct effect upon the vocal folds causing it to vibrate. Assessing aerodynamic activities that form the basis for voice production includes measuring volumes of the lung and capacity of the lung. There are four important lung volumes measured: tidal volume, expiratory reserve volume, inspiratory reserve volume, and residual volume. Capacity is two or more volume measurements. The various capacities of the lung are: forced vital capacity (FVC), vital capacity (VC), inspiring capacity (IC), functional residual capacity (FRC) and total lung capacity (TLC). Other parameters measured during aerodynamic measurement of voice (Jiang & Titze, 1994) are the glottal airflow, maximum phonation time, sub-glottal pressure and laryngeal resistance in addition to lung volumes and capacities. The disclosure of the structure and function of vocal folds has become a crucial procedure in the clinical voice assessment protocol. Indirect mirror laryngoscope is the traditional method used to perform laryngeal examinations on dysphonic patients. Vocal folds have been efficiently visualized using laryngeal imaging techniques including endoscopy, videostroboscopy, video kymography and high-speed video endoscopy (Mehta, Deliyski, Quatieri & Hillman, 2011). There are softwares which record voice and effectively measure pitch, loudness and quality related parameters. Voice assessment is one such speech analysis program for speech-language pathologists and speech and voice scientists to do speech/voice assessments. Kay Pentax's MDVP (Multi-Dimensional Voice Profile) software calculates 33 measures and literature measures of major interest. There are a large number of literature studies employed the MDVP software as an acoustic analysis evaluation tool. First, studies that evaluate the program's reliability and its measures can be located. Researchers are widely using the fundamental frequency, jitter, shimmer, and NHR related voice measures. Such objective parameters are important for voice disorder patients (Christmann et al., 2015).

2.1.3. Acoustic voice analysis

The main clinical application of acoustic analysis can be classified into three broad categories: (1) screening, (2) diagnosis, (3) assessment of the effectiveness of different management approaches, and (4) assessment of progress throughout management (Laver, Hiller & Beck, 1992). Screening procedures are often used to detect voice pathology early. Numerous techniques have been described that are able to discriminating normal with rationally good precision against abnormal voices (Hadjitodorov & Mitev, 2002), although some investigators have questioned the usefulness of automating this type of decision-making (Hirano et al., 1988). The objective of diagnostic acoustic voice measurement is to distinguish clinically normal voice from various pathological conditions. It is quite clear the clinical utility of discriminating between the underlying pathological conditions. In the literature, the use of acoustic measurements is to provide objective indices of improvement throughout a therapy program or to compare the relative usefulness of alternative therapy strategies. The basic frequency (F0), jitter, shimmer, HNR and formant frequencies are acoustic parameters commonly used in acoustic analysis applications as well as the most referenced in the literature. To derive indices such as DSI (Dysphonia Severity Index) and AVQI (Acoustic Voice Quality Index), these parameters are calculated. Such indices reflect multiple acoustic parameters yielding a strong acoustic marker of voice. Voice sample such as phonation is frequently used for acoustic analysis. Although speech had been the common mode of communication, phonation alone might not serve the purpose of voice evaluation. In some cases the deviance in the voice quality is hard to identify in the phonation sample. Thus, in addition to phonation, acoustic measurement of speech sample which is the natural mode of voice production is ecologically valid. AVQI is the first measure to incorporate samples of continuous speech to date, considering the sustained vowel samples used in other measurement protocols.

Wuyts et al. (2001) developed DSI, which is one of the leading acoustic measurements incorporated into voice disorder assessment and diagnosis. DSI requires people to phonate at different levels of pitch and loudness. Unlike DSI, AVQI measures the acoustic parameters in a simple phonation and spontaneous speech sample which is relatively a natural mode of communication. AVQI in this regard is advantageous and reflects the acoustic parameters in a natural vocal output. Maryn et al. (2010) developed a multivariant acoustic voice quality index [AVQI]. This index proposes to combine in a single numerical data a sequence of relevant acoustic parameters. The following section describes AVQI in detail.

2.2. Acoustic Voice Quality Index

AVQI is one of the clinical and research utility voice assessment tools. It is a multivariate construct based on linear regression analysis combining several acoustic markers to produce a single score that is reasonably correlated with the auditory perceptual judgment of the overall dysphonia severity (G from GRBAS scale). Maryn et al. (2010) used concatenated and connected vowel samples analyzed by 13 acou stic measurements (based on fundamental frequency, Frequency perturbation, amplitude perturbation, spectral and cepstral analyses). Step by step, multiple linear regression analysis resulted in a six-variable acoustic model. The parameters used for AVQI include Smooth Cepstral Peak Prominence (CPPS), Harmonic-to-noise ratio (HNR), Shimmer local (SL), Shimmer local dB (ShdB), long-term average spectrum slope (slope) and trendline tilt through the long-term average spectrum (tilt). The AVQI is thus constructed as AVQI= 2.571*[3.295-0.111*CPPS-0.073*HNR- 0.213*SL+ 2.789*ShdB-0.032*Slope+ 0.077*Tilt] by the algorithm. A score of 2.95 or below obtained on AVQI identified the sample to be normophonic for Dutch speaker (Maryn et al., 2010). It was reported that higher the AVQI scores more the affected voice quality and vice versa.

2.2.1. Western studies on AVQI

By using AVQI, Kankare et al. (2015) assessed the voice quality of the Finnish speaking population. A total of 50 Finnish native speakers participated with 22 voice patients and 28 healthy individuals. Reading a phonetically balanced text and phonation of vowel /a/ were asked. Three seconds of /a/ phonation and first 23 syllables (8 words) of the text were edited for middle analysis. Five experts assessed the severity of GRBAS-scale dysphonia with Grade, G. Authors found that AVQI indicating ROC as 0.898 was an excellent diagnostic accuracy. The mean dysphonic voice AVQI scores were 3.95 (SD 1.88) and 1.48 (SD 0.67) for healthy voice. The Likelihood ratio (LR) confirmed a good result with the AVQI threshold being 2.23 in diagnostic accuracy. Authors concluded that AVQI is a valid tool for voice quality assessment in Finnish speaking population. Authors added that future studies would be investigated by considering before and after vocal loading and intensive vocal training in a larger study group, in particular with intermediate dysphonia levels and AVQI outcomes in healthy voices.

Núñez-Batalla, Díaz-Fresno, Álvarez-Fernández, Cordero, and Pendás (2017) aimed to quantify dysphonia's overall severity through meta-analysis. The Acoustic Voice Quality Index brings together into a single numerical data with a series of relevant acoustic parameters. This study was participated by a total number of 108 participants. In that 58 people were dysphonic and 60 volunteers were healthy. These people were about 20 –60 years old in the age range. These people were asked to phonate vowel /a/ for about 3-5 seconds and were made to read a passage from the auditory perceptual voice assessment in Spanish. Two specialists used the GRABS and CAPE-V analyzed the samples by auditory–perceptual analysis. The result exposed a significant difference between the two groups as well as also distinguished between the healthy volunteers and the dysphonic individuals. The obtained AVQI was found to be an average of about 7.3 with a standard deviation of 1.07, ranging from 5.3 to 9.8, for the sustained vowels. The AVQI obtained for phrases, with a standard deviation of 0.70 and a range of 8.5 to 11.6, was found to be an average of about 9.7. A longitudinal analysis was also performed in a group of 20 patients to compare the scores of pre- and postoperative AVQI. The AVQI obtained before the operation had an average of about 7.8 and a standard deviation of 0.84, ranging from 6.0 to 9.6, and the AVQI obtained after the operation was about 7.05 with a standard deviation of 1.12 and ranged from 5.3 to 9.3. The study found a good correlation between the overall perception of voice quality and the AVQI value, and there is a significant difference between the normal and dysphonic voices on AVQI ($t(95)=9.5$;

P<.000). The authors concluded that this study shows the clinical utility of the AVQI as a measure of dysphonia severity.

Uloza et al. (2017) aimed at validating and investigating the practicability and robustness of the Lithuanian language (LT) acoustic voice quality index of its diagnostic accuracy. With normal voices (n=46) and with different voice disorders (n=138), a total of 184 native Lithuanian speakers participated. They were asked to read a Lithuanian text and phonate the vowel /a/. The grade (G) from the Grade, Roughness, Breathiness, Asthenia, and Strain (GRBAS) protocol and the overall severity from the Auditory Perceptual Voice Assessment (CAPE-V), Visual Analog Scale (VAS) protocol were rated. To receive an AVQI-Lithuanian score, all concatenated voice samples were acoustically analyzed. There was good consistency between five raters in both auditory and perceptual judgments. It was reported that there was a significant correlation between both auditory-perceptual judgment and AVQI-Lithuanian scores. A reasonable sensitivity = 0.838 and excellent specificity = 0.937 was achieved by the AVQI-LT threshold of 2.97 for the Gmean rating. An AVQI-LT threshold of 3.48 with sensitivity = 0.840 and specificity = 0.922 was determined for the VAS rating. Consequently, the authors concluded that AVQI-LT is measured as a valid and reliable tool for assessing the severity of dysphonia in the population of Lithuania.

The objective of Hosakawa et al. (2017) study was to measure the concurrent validity, responsiveness to change and diagnostic accuracy of the Japanese-speaking population for AVQI voice assessment in Japanese. The study considered a total of 336 voice recordings, including 69 pairs of voice recording (before and after therapeutic interventions). Five experienced raters evaluated the overall voice quality auditory-perceptual evaluation. The simultaneous validity and responsiveness to change and diagnostic accuracy of the AVQI were estimated. It was reported that high correlation coefficients of 0.828 and 0.767, respectively, were indicated on the basis of overall voice quality for the concurrent validity and responsiveness to change. Analysis of receiver operating characteristics (ROC) showed excellent diagnostic accuracy for discrimination of dysphonic and normophonic voices (curve area: 0.905). The best 3.15 AVQI threshold level corresponded to a sensitivity of 72.5 percent and a specificity of 95.2 percent, respectively with the positive and negative probability ratios of 15.1 and 0.29. Thus, the authors concluded that AVQI is used in the Japanese-speaking population as a tool for overall evaluation of voice quality and voice therapy outcomes.

Kim et al. (2018b) investigated the feasibility of its cut-off values and diagnostic accuracy in the Korean population in discriminating between normal and dysphonic voices. A total of approximately 1,524 native Korean subjects were asked to verbalize and maintain the vowel /a/ with 113 normal voices and 1411 voice disorders and read aloud the Korean "Walk" text. A 2-second centre of sustained vowel sound and sentences containing 26 syllables was transformed into edited. The Grade (G) from the GRBAS protocol and the CAPE-V were used to evaluate the voice by five speech language pathologists. In the characteristic curve analysis of the receiver, the cut-off values of AVQI, G, and OS (overall severity) were <3.33, <0.00, <22.00. The characteristic curve analysis of the receiver operation indicated that AVQI has exceptional diagnostic accuracy in distinguishing normal and dysphonic voices, i.e. area under the curve: 0.970-0.977. The authors therefore concluded that AVQI's is a valid tool for assessing overall voice quality and quantifying dysphonia in Korean speaking population

2.2.2. Indian studies

The Acoustic Voice Quality Index was documented by Seshasri (2018) in 80 typically Kannada-speaking children aged 10-12 years. The participants were divided into two groups (10-11 years and 11-12 years) with an interval of one year age internal. The study included recording of vowel /a/ (phonation) and reading task. Used for the reading task was the standardized Kannada passage (Savithri & Jayaram, 2005). The author found no effect on the value of AVQI on age and gender. For 10- 11 year males, the mean AVQI score was 3.66 with a standard deviation of 0.58 and for females, mean AVQI was 3.35 with a standard deviation of 0.60. The mean AVQI score for 11 to 12 year males was 4.09 with a standard deviation of 1.03 and females, it was 3.85 with standard deviations of 0.65. Author also concluded that for older children the AVQI was relatively higher; it was attributed to the onset of mutational changes in these kids. The above study is confined to a paediatric population that enhance AVQI knowledge across various age ranges.

2.2.3. DSI & AVQI

Latoszek, Ulozaitė- Stanienė, Maryn, Petrauskas and Uloza, (2017) investigated the influence of gender and age on AVQI and DSI. In this study, a total of 123 vocally normal healthy people were evaluated. Mid 3 secs of the sustained vowel /a/ and standardized Lithuanian phrases were edited for AVQI. Participants were asked to phonate vowel /a/ for at least 2 secs for DSI. Second, it took maximum phonation time by asking the participant to maintain their usual pitch and loudness as much as possible. Finally, vocal range profile where at their lowest intensity and highest pitch they were asked to phonate vowel /a/ for at least 2 seconds. Authors reported that there was no gender effect on AVQI and DSI. DSI had, however, a significant age correlation ($P<0.05$), with 5% of the age variance. Authors concluded that AVQI values are not dependent on gender and age in which DSI values are not dependent on gender but correlated slightly with age.

Uloza, Latoszek, Ulozaite-Staniene, Petrauskas, and Maryn (2018) investigated and compared the feasibility and robustness of AVQI and DSI in diagnostic accuracy. A group of 264 subjects with an average age of 43.13 years and a standard deviation of 15.09 years participated in this study. There were about 105 individuals in the normal voice subgroup, 47 being males and 58 being females. There were 159 individuals in the pathological voice subgroup, 72 were males, and 87 were females. They were asked to phonate vowel /a/ for minimum of 5 seconds also read a phonetically balanced text segment. Mid 3 seconds of the sustained phonation was edited for auditory perceptual evaluation. A segment of atleast minimum of 2-s duration of the sustained vowel /a/ has been analyzed for the Dysphonia Severity Index. The mean DSI is + 5 for normal voices perceptually and −5 for severe dysphonic voices. The more negative the DSI of the patient, the worse the vocal quality of the patient. These samples were also rated by five qualified clinical voice specialists who were all native Lithuanians. They used GRBAS scale and CAPE-V for perceptual analysis. The results showed that both auditory-perceptual judgment procedures and acoustic voice measurements were significantly valid at the same time. The DSI threshold (i.e., DSI=3.30) for Gmean was 85.8 % sensitivity and 83.4 % specificity. VASmean also determined the DSI threshold of 3.30 with a reasonable sensitivity of 70.3% and an excellent specificity of 93.9%. In addition, the AVQI threshold (i.e., AVQI=3.31) for Gmean showed a reasonable sensitivity of 78.1

percent and an outstanding 92.0 percent specificity. For VASmean, an AVQI threshold of 3.33 was established with an outstanding 97.0 % sensitivity and a reasonable 81.8 % specificity. They found a moderate negative correlation between DSI and the overall severity of dysphonia and breathability and a weak negative correlation of roughness. The AVQI found high correlations with factors G, R, and B. Since DSI is used as a common objective measure for the assessment of dysphonia severity, there are quite a few limitations as the DSI values may be affected by differences in the age of the patient, gender, ethnic and cultural variations, vocal training, potential differences between software and devices used, and the mental state of the subject during measurement. The same patient's DSI values measured at different times may also vary. Few studies show that the AVQI was sensitive to treatment changes, validating its function as a measure of potential robust and objective voice treatment outcomes. However, the disadvantage of DSI could be the difficult use of DSI with severe dysphonia and aperiodic voice due to the failure to calculate jitter as a basic frequency-based parameter. In addition, this method requires the participation of a professionally skilled being, in spite of the ease of registration and computation of DSI. Consequently, the individual cannot automate and perform this procedure as a voice "self-assessment." It decreases the DSI's potential usability for screening purposes for voice pathology. The DSI is now viewed as an additional quantification of vocal function, while the AVQI is viewed as an additional measure of vocal sound quality. The authors concluded that the DSI and AVQI showed high accuracy in discriminating against normal and dysphonic voices. However, higher validity has been given for AVQI. Consequently, the AVQI seems to be helpful in describing differences in vocal quality status and discriminating between normal and dysphonic voices on the basis of auditory-perceptual judgment. Furthermore, AVQI as "ecologically valid," measurement can potentially be automatic and smooth without involving specific staff, i.e. the individual could be assured of voice quality self-assessment. Therefore, AVQI suggests a reliable voice screening potential.

Barsties, Latoszek, Ulozaitė-Stanienė, Petrauskas, Uloza, and Maryn (2019) evaluated the diagnostic accuracy of both the DSI and AVQI measurements in comparison to the dysphonia classification. With normal voice (N= 105) and various vocal disorders (N= 159), a total of 264 subjects participated in the study. For DSI, subjects were first requested to phonate vowel /a/ for at least 2 secs. Second, MPT was measured after maximum inspiration by asking the subject to keep phonating vowel /a/ long at their usual pitch and loudness as long as possible. Third, at different fundamental frequencies and sound pressure levels, vowel /a/ phonation (gliding) was requested for at least 2 secs. The DSI was calculated using only the lowest intensity and highest frequency. A concatenated voiced segment of the standardized Lithuanian phrase and mid 3 sec of the sustained vowel /a/ phonation is considered for AVQI analysis. The acoustic measurement was obtained from AVQI script version 02.02. DSI has been found to have higher potential in the classification of dysphonia evaluation. The DSI threshold of 3.05 was 94.3% more sensitive and the specificity was 84.3%. Also, the AVQI threshold of 3.31 showed 71.7% reasonable sensitivity and 88% specificity. The authors discussed that DSI is considered a vocal function measure and AVQI is based on the measurement of acoustic voice quality. Thus, by auditory perceptual judgment, AVQI may be more appropriate in discriminating disordered versus non-disordered voice quality. There were also few limitations in the study. First, procedural differences between the measurements that are difficult to compare, i.e. AVQI is the only acoustic measurement that uses continuous speech when DSI measures the vocal function aspects that rely on continuous phonation. Second, single organic types like polyps, edema, and cysts would have irregularities and the noise component would hinder the acoustic measurement. Third, DSI and AVQI use time-domain measurements for their calculation; these can lead to errors in finding both indices, especially for extremely severe heavy voice (e.g., Type III signal). DSI was concluded to have a greater potential to evaluate dysphonia by the authors.

2.2.4. Comparison between CPPS speech tool vs. AVQI v2 vs. AVQI v3

Kim et al. (2018a) compared two versions of AVQI (2.02 and 3.01, which are termed as v2 and v3, respectively) and Praat CPPS using a Korean population voice sample. A total of approximately 2,257 patients with voice disorder participated in this study, who had voice recording prior to the intervention. These individuals were asked to maintain vowel /a/ phonation at their comfortable pitch and loudness and were also asked to read 10 words Korean phrase ' walk ' passage with 25 syllables. The center 2 seconds of the /a/ phonation sample and the 'walk' passage were edited in the Praat software for AVQI analysis. The acoustic components used to measure AVQI are similar on the basis of versions 2.02 and 3.01. Auditory - perceptual evaluation is considered a gold standard for determining severity. In this study, these voice samples were analyzed by two Speech Language Pathologists and a dysphonia professor who had seven years of voice assessment experience. The results revealed that there was a high reliability in differentiating the pathological voice disorders using AVQI (v2 and v3) and Praat CPPS. The AVQIv2 has been found to exceed AVQIv3 and PraatCPPS.

2.2.5. AVQI in Praat vs. Speech tool

AVQI can be calculated using various formulas using different softwares such as Praat and Speech tool. Studies examined the efficacy of using various software to calculate AVQI . Maryn and Weenink (2015) conducted one such study to investigate the proportional relationship; CPPS in Praat corresponded to the original CPPS of SpeechTool. The voice recordings were collected by two sets of subjects. First, voice samples of 22 vocally normal and 228 voice-disordered participants were recorded. Second, voice samples were collected from six vocal normal subjects and 33 voice-disordered subjects. Total 289 (i.e. 28 normo-phonic and 261 dysphonic) subjects, with 193 females aged approximately 10 years and 86 years (mean age 36.8 years and Standard Deviation 18.4 years) and 96 males aged approximately 8 years and 85 years (mean age 43.0 years and Standard Deviation 21.5 years) participated in the study. A mid 3 seconds and first two sentences were edited. Two versions have been compared based on the acoustic marker, the initial version — the original or alfa version developed by Maryn et al. (2010). AVQIalfa is a combination of five acoustic markers in the Praat program plus one acoustic marker in the SpeechTool program. Nevertheless, having to merge two computer programs to reach a only dysphonia severity index reduced AVQIalfa's user-friendliness/feasibility and induced a comparatively difficult method with at least eight steps. Only the Praat program is involved in the beta version, and the number of directions is reduced to just two: (1) making the necessary recordings, and (2) activating the newly modified Praat script that automatically yields an AVQIbeta score. The results revealed that the CPPSPraat data is systematically superior to the CPPSSpeechTool data due to the differences between the two measures in signal processing. The association between AVQIalfa and AVQIbeta revealed that the values of AVQIbeta were very similar to the values of AVQIalfa, which could be said to represent the AVQIalfa almost perfectly (i.e. 96.0 %). The authors concluded that the results of the two methods of CPPS and the two methods of AVQI are very similar, enhancing the clinical feasibility of both methods as measurements of dysphonia severity.

2.2.6. AVQI in Professional Voice Users

D'haeseleer et al. (2017) examined vocal quality, vocal complaints and hazard factors in the development of vocal disorders in theatrical actors. They compared voice quality before and after a theatre performance. Voice samples from 33 actors who had 2 years of experience in the acting field were collected. The mean performance duration was about 96 minutes, and the Standard Deviation was about 25 minutes. The actors ' age ranged from 23 to 69 years (*M*=41.9 years). These actors played a foremost role in the performance of theatre. The artist were asked to produce a sustained vowel /a/ and read a phonetically balanced Dutch text "Papa en Marloes" at regular pitch and loudness. AVQI was calculated and analyzed using 3 seconds of the vowel /a/ section and the first two sentences of the text. The GRBAS scale was used for auditory-perceptual assessment, and the ratings were rated by two voice therapists who have experience in assessing and treating voice disorders. The Voice Handicap Index was used to account vocal symptoms and factors that influenced them. The results revealed that 3.48 equivalent to a mild dysphonia was the mean AVQI in the theatre actors before the performance. 50 % of the actors reported vocal complaints after the performance. 50% never warms the voice before a performance. After a performance, most actors (88.5%) never carry out cool-down exercises. Vocal misuse is present in most actors, such as screaming or throat clearing. The results revealed that the AVQI did not change significantly after the theatre performance. After the performance, the mean AVQI was 3.42 and still was a mild dysphonia. The auditory-perceptual voice evaluation results using the GRBAS scale showed that the values of the two raters were 75 %. The comparison between the perceptual evaluations of pre- and post-theatre performance revealed significant differences in the overall dysphonia grade (P= 0.035). The authors concluded that actors in the theatre have mild dysphonia incidence, vocal complaints, and poor vocal hygiene habits. Future direction highlighted in the article is to study the long-term impact of the of voice theatre performance. Authors also reported that the evaluation protocol warranted videolaryngostroboscopy to investigate the larynx and vocal folds structure and function.

In comparison with non-professional theatre actors, Leyns (2017) investigated the effect of theatre performance on voice quality of professional theatre actors. Subjective and objective measurements have been used to mark risk factors and vocal habits in them. The study was attended by a total of 54 participants whose age ranges between 18 to 48 years (*M*=30.24 years, and *SD*=8.38 years). Of the 54 participants, 27 were professional actors in the theatre and 27 were non-professional actors in the theatre. The age of non-professional actors in the theatre varied between 18 and 42 years. The voice samples were collected with an average duration of 90 minutes before and after a performance. In order to eliminate the vocal load, eight professional dancers were included in the non-professional actors group. Before and after the performance, the participants underwent various voice assessment protocols. The subjects were instructed at the usual pitch and loudness to phonate and read. Both the DSI and AVQI have been calculated. The GRBASI scale has been used for auditory-perceptual evaluations. In order to capture vocal abuse and risk factors, the subjects completed several questionnaires before and after the performance. Except for the acoustic parameters F0, Flow and shimmer local, for aerodynamic measurements and for the LTAS slope, there were no significant differences between pre and post performance. There was no significant change in the overall grade of perceptually evaluated dysphonia. AVQI results showed a pre-and post-mild grade dysphonia overall. DSI results, on the other hand, did not indicate vocal issues. In professional actors, the questionnaires revealed poor vocal habits. The results of this study demonstrated an almost non-existent impact of performance in actors and dancers on the objective and subjective vocal quality. A possible effect of warm-up caused by acting was observed, however. The results are in line with the purpose of both indices: as AVQI purely measures the vocal quality, DSI evaluates the maximum vocal capacity. This is also in support of the study by Maryn and Weenink (2015) who showed that AVQI is a powerful and suitable way of calculating the overall speech dysphonia. The study's limitation is that the recording of the phonation and speech sample in the changing room where there could be considerable amount of noise that would influence the results.

2.2.7. AVQI in children

Reynolds et al. (2012) investigated the diagnostic accuracy of the AVQI in paediatric voice-disordered and normal participants. Previously, the AVQI had a high ecological validity in adult populations. A total of about 67 preterm participants (born under 25 weeks gestation) aged 6 to 15 years and 40 normo-phonic term-born participants aged 5 to 15 years were recruited. These children performed reading and phonation task. The AVQI was found to be an appropriate measure of dysphonia severity in the group of extremely premature infants with a very high incidence of dysphonia and a term born, normo-phonic group. The AVQI values were correlated with GRBAS scores, which represent a strong, positive association between the two measures, showing AVQI to be an accurate indicator of dysphonia severity.

Reynolds, Meldrum, Simmer, Vijayasekaran, and French (2014) used an objective acoustic voice assessment to document the incidence of dysphonia in very premature school-age children, as well as a controlled trial of behavioural voice therapy in this population. A total of 200 participants were considered for the study who were born under 32 week’s gestation. Participants’ age ranged from 5 to 12 years. Phase I included a clinical evaluation conducted by SLP with postgraduate experience including GRBAS, CAPE-V, AVQI, and pVHI. Phase II included an ENT surgeon's video stroboscopic assessment using endoscopic studies. Phase III included intervention where a total of eight behavioural voice therapy sessions, consisting of 45 minutes per session per week, was received by participants. To determine the reactivity of the AVQI to change, a standardized change score analysis was used. A moderate-to-high correlation was found among the scores indicating the AVQI score's responsiveness. AVQI can therefore be used to monitor the therapeutic changes in early intervention.

2.2.8. AVQI in dysphonic population

Maryn, De Bodt and Roy (2010) investigated two experiments in this study, the first experient was to investigate external cross-validity of AVQI and the second experiment was to investigate AVQI's responsiveness to change after treatment. The first experiment involved six vocally normal subjects and 33 voice-disordered subjects. Age range between 16 to 86 years (M= 49.2 years, SD= 20.1 years). The recordings of 33 subjects before and after therapy took part in the second experiment. This group consisted of 22 females and 11 males, with an average age of 40.9 years and a standard deviation of 18.9 years, between 7 and 68 years. All of these 33 subjects received an eclectic treatment program combining the techniques of behavioural voice therapy with an individual combination. Behavioural voice therapy included indirect strategies such as voice hygiene counselling and advice, and healthy voice use and direct strategies include combined speech respiration, resonance, pitch, loudness, voice facilitation, and voice initiation exercises to improve voice and reduce the number and severity of voice-related complaints. With an average of 7.4 sessions, the number of behavioural voice therapy sessions ranged from 1 to 49. These subjects were asked to phonate the vowel /a/ for at least 5 seconds and read a phonetically balanced Dutch text aloud. The Praat software analyzed the mid 3 seconds and the first two sentences. Before voice therapy, one individual had a moderate dysphonia, mean G = 1.6 and an AVQI = 6.31. However, following voice therapy, a normal voice quality was achieved, mean G = 0.0 and an AVQI = 1.83. it is seen that AVQI has reduced after post therapy. Five experienced Speech-Language-Pathologists using the GRBAS scale analyzed these 105 samples further. The results showed moderate reliability between raters. The study obtained an acceptable external validity and thus the authors concluded that the AVQI deserves additional attention as a promising objective measure of treatment outcome that could be integrated as an important part of a multidimensional evaluation of treatment effects.

Ulozaite, Petrauskas, Saferis, and Ulozas (2016) investigated the feasibility and robustness of AVQI-LT for dysphonia quantification, and it‘s correlations with auditory-perceptual judgment. With the average age of about 40 years, a total of about 153 individuals enrolled in this study. In that 107 people were dysphonic and 43 were normal healthy people. They were made to phonate vowel /a/ as well as read the Lithuanian phrase "Turėjo senelė žilą ožel". Only voiced parts were edited for speech and medial 3 seconds of sustained /a/ phonation for acoustic analysis. Five experienced laryngologists, all of whom were native Lithuanians, rated these samples to quantify the vocal deviation for auditory-perceptual analysis. Voice samples were judged based on GRBAS scale and CAPE-V. Authors reported that the auditory perceptual evaluation had almost perfect inter-rater reliability. The validity of both the AVQI-LT and the Auditory Perceptual judgment was highly significant. These results showed significant concurrent validity between both procedures for auditory-perceptual judgment and AVQI-LT. The authors concluded that the robustness of this method of automated voice analysis confirms the high correlation between AVQI-LT and Auditory Perceptual analysis of voice. Therefore, AVQI can allow an important step to make practical, reliable and objective voice evaluation tool suitable for non-experts and voice professionals.

2.2.9. AVQI treatment outcome measures

Kim, Lee, Park, Bae, Lee, and Kwon (2017) investigated the criterion-related simultaneous validity of two standardized auditory-perceptual assessments and the AVQI to measure the severity of dysphonia in patients with vocal cord paralysis (VCP). A total of approximately 210 patients with vocal cord paralysis and 236 people with normal voice were involved in this study. These people were made to phonate the vowel /a/ as well as made read a Korean text aloud. A segment of the sustained vowel was edited in the mid-2-seconds and two sentences containing 26 syllables. Three raters further rated these samples using the two standardized GRBAS and CAPE-V auditory-perceptual evaluations. The results revealed that AVQI, Grade (G) and Overall Severity (OS) values were higher in the Vocal Cord Paralysis group than in the normal voice group. The correlation ranged from 0.904 to 0.926 between AVQI, G, and Overall Severity. In the ROC curve analysis, the cut-off values of AVQI, G, and OS were < 3.79, < 0.00, and < 30.00, respectively, and each analysis had an Area Under Curve of more than 0.89. The authors concluded that the AVQI and auditory evaluation could enhance Vocal Cord Paralysis ' early screening ability as well as help establish an effective diagnosis and treatment plan for vocal cord paralysis and other related dysphonia.

To summarize, AVQI has been developed to measure overall voice quality with objectively with multiparametric acoustic markers and acoustic markers for clinical purposes. It is a validated objective tool to measure overall voice quality and dysphonia severity through two speech tasks such as sustained vowel and continuous speech (Maryn & Roy, 2012). AVQI reflects overall voice quality from at least these two speech tasks in order to be representing daily voice use patterns. The overall quality of the voice is multidimensional and is not linked to a single physical changeable or psycho-acoustic determinant. It is related to phenomena that occur in different signal fields and the AVQI is built on multiple acoustic markers from different fields. The AVQI is useful in clinical voice practice for tracking overall voice quality in time and across treatment. Hence the utility of AVQI needs to be established across different languages and geographical distributions across the world for its greater application.

CHAPTER III

METHOD

3.1. Research design

The present study employed a normative research design to investigate the aims and objectives.

3.2. Participants

A total of 121 participants were recruited for the study in the 20–50 year age range (Mean age =34.02 years; SD=10.98 years). Moreover, based on the age range, these participants were divided into two groups. The first group (Group I) consisted of 64 normo-phonic individuals aged 20–35 years with 32 females and 32 males (Mean age=33.52 years; SD=11.26 years) and the second group (Group II) consisted of 57 normo-phonic individuals aged 36–50 years with 29 females and 28 males (M=43.39 years; SD=5.79). Normo-phonic individuals in both group I and II consisted of 121 participants. The third group consisted of 15 dysphonic individuals aged 20-50 years (M=38.47 years; SD=15.88 years) with 5 females and 10 males. All the participants were native Tamil speakers recruited for the study. Group I and Group II referred to as the control group as in the present study and group III as the clinical group.

3.2.1. Inclusionary criteria

The participants who satisfied the following criteria were considered for the study;

- Individuals with Tamil as their native language
- Should have good proficiency in reading and writing Tamil language (LEAP- Q) (Ramya & Goswami, 2007)

- Should have learned Tamil as the primary language till 10th standard
- Should be exposed to the language (Tamil) for at least 15 years
- Should be a resident of the region for at least 15 years
- Should have normal voice quality at the time of the testing (perceptual evaluation)

3.2.2. Exclusionary criteria

- The study excluded participants with active or history of vocal-tract-related infections or history of chronic obstructive pulmonary disorder, asthma, or any other lung infections.
- The study excluded participants with complaints on speech, language or hearing loss or any associated communication disorders or neurological impairment.
- Participants with active or history of alcohol consumptions, smoking or tobacco intake was excluded from the study.

3.3. Stimuli

A phonation sample of vowel /a:/ for a minimum of 5 to 6 seconds and a continuous speech/reading sample in Tamil was recorded. The standardized Tamil passage (Savithri & Jayaram, 2005) was used for the reading task (continuous speech). The participants will be asked to read nine sentences, the first three and the last three sentences were excluded, and the middle three sentences were considered for analysis, which contains 57 syllables.

3.4. Ethical consideration

The enrolled participants were clearly explained regarding the aim and objectives of the study. They were ensured about their safety and confidentiality during their participation in the study. Prior to the study, the participants gave their written informed consent (see Annexure I).

3.5. Procedure

The recording was carried out in a quiet room/ noise free place within the premises of participants working setup. The participants were asked to sit comfortably on a chair. A distance of 10 cm was maintained between mouth to the microphone to avoid breathing noise. Sennheiser CX 275s earphones with the microphone were used for recording purpose and the voice/speech were directly recorded in the Praat software (6.0.28 version) using a Lenovo Laptop.

First, the participants were made to phonate a sustained vowel /a:/ for a minimum of 5 to 6 seconds. Second, the participants were asked to read the standardized reading passage continuously in the Tamil language. Both the samples were recorded at the participants' comfortable pitch and loudness level. To avoid reading errors, each subject was made to familiarize with the reading material before actual recording. The recordings were obtained from the participants preferably before noon to eliminate the effect of vocal loading on the samples. Three trials of both phonations and continuous speech (reading) were obtained from them. Best one of the three trails were considered for the analysis.

3.6. Data Analysis

The recorded samples were saved in .wav file format and renamed with convenient codes. Praat software analyzed the samples. AVQI (v02.02) was measured using the Maryn and Weenik (2015) algorithm. The script of AVQI has an algorithm which contains the subsequent formula, AVQI = 9.072-0.245*CPPS-0.161*HNR-

0.470*SL+6.158*ShdB-0.071*Slope+0.170*Tilt was used. To elicit AVQI, the script given by Maryn & Weenik (2015), was copied into a text file and saved as .text. Finally, the saved .wav files were opened in the Praat software. Along with these files, the script file was also opened by clicking the Praat option-open Praat script- select the '.text.' Script and the AVQI values and its constituent parameters were extracted after the algorithm 'runs' in the software. The following measures were documented.

a) Acoustic Voice Quality Index (AVQI)

b) Smoothened Cepstral Peak Prominence (CPPs)

c) Harmonic to Noise Ratio (HNR)

d) Shimmerlocal (SL)

e) ShimmerdB (shdB)

f) Slope of Long term average spectrum (LTAS)

g) The tilt of the regression line through the Long-term average spectrum (Tilt)

3.7. Statistical analyses

The following statistical tests were performed to analyze the data obtained.

- Shapiro Wilks test of normality was done to check whether the data distributed normally.
- Descriptive statistics were carried out to document the mean and standard deviation values for reference measures for AVQI and its constituent parameters for individuals with normal voice quality and individuals with dysphonia.
- The difference between control (normo-phonics in group I & II) and clinical group (group III) on AVQI value and its constituent parameters was examined using independent 2 sample t-test.

- The effect of group and gender on AVQI scores were analysed using MANOVA test.

CHAPTER IV

RESULTS

The study aimed at investigating the AVQI in native Tamil speakers with the objectives of establishing the normative data for AVQI, and examining the effect of age and gender on AVQI. A total of 121 normo-phonic and 15 dysphonic individuals participated in the study. The AVQI and its constituent parameters namely, CPPS, HNR, Shimmer Local, Shimmer dB, LTAS, and Tilt of LTAS were recorded and considered for analyses. Age and gender served as independent variables whereas the AVQI and its constituent parameters were considered to be the dependent variables in the present study. Using SPSS version 21, descriptive and inferential statistical analyses were performed on these measures. For each of the measures, descriptive measures such as mean (M) and standard deviation (SD) have been calculated. The normality test of Shapiro-Wilk was carried out on the data indicating a normal data distribution ($p>.05$). Therefore, parametric tests such as Multivariate analysis of variance (MANOVA) and t-tests were performed on the data for further analyses. The details of statistical analyses are provided below.

4.1. Comparison of AVQI value between group I and group II

The mean (M) and standard deviation (SD) values of AVQI recorded from 121 normo-phonic individuals were noted. In the age range of 20 and 35 years, Group I had 32 males and 32 females among 121 participants. Group II was made up of 28 males and 29 females between 36 and 50 years of age. The details of descriptive statistics of the AVQI in normo-phonic individuals are tabulated in table 4.1. The mean AVQI was found to be 2.76 (*SD*=0.77) across all the participants considered (20-50 years).

Table 4.1

Mean (M) and standard deviation (SD) values of AVQI in normo-phonic individuals

Groups	Gender	M	SD
Group I	Males	3.05	0.75
	Females	2.68	0.65
	Total	2.87	0.72
Group II	Males	2.82	0.77
	Females	2.50	0.83
	Total	2.65	0.81

The AVQI was examined across two age groups (group I & group II) from table 4.1, the total mean AVQI was found to be higher for group I (*M*=2.87, *SD*=0.72) compared to group II (*M*=2.65, *SD*=0.81). The total SD AVQI was higher in group II compared to group I. Within groups, the mean AVQI value was higher in males compared to females in both the groups. Also, the SD AVQI value was higher in males for group I and it was higher in females for group II.

4.2. Comparison of constituent parameters of AVQI between group I and II

Further, the mean (M) and standard deviation (SD) values of constituent parameters of AVQI were examined. The details of descriptive statistics of the constituent parameters of AVQI such as CPPS, HNR, Shimmer Local, Shimmer dB, LTAS, and Tilt of LTAS in normo-phonic individuals (group I and II) are tabulated in table 4.2

Table 4.2

Mean (M) and standard deviation (SD) values of the constituent parameters of AVQI in normo-phonic individuals (group I & group II)

Groups	Gender	CPPS		HNR		Shimmer Local		Shimmer dB		LTAS		Tilt of LTAS	
		M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
I	Males	14.31	1.20	15.89	1.89	7.66	1.67	0.75	0.09	19.97	3.01	14.22	0.50
	Females	13.95	1.13	17.93	1.68	6.76	1.63	0.79	0.10	17.98	3.88	14.31	0.4
	Total	14.13	1.17	16.91	2.05	7.15	1.73	0.72	0.10	18.88	3.97	14.26	0.49
II	Males	14.32	1.10	15.85	1.86	8.67	1.28	0.79	0.09	20.36	2.29	14.37	0.59
	Females	14.20	1.43	17.81	1.92	7.07	2.03	0.70	0.13	17.30	3.85	14.26	0.42
	Total	14.26	1.27	16.84	2.11	7.86	1.87	0.74	0.12	18.81	3.51	14.31	0.51

On observing the mean values of the constituent parameters in table 4.2, the total mean CPPS, Shimmer Local, Shimmer dB and Tilt of LTAS were found to be higher in group II compared to group I individuals. The total mean HNR and LTAS were observed to be higher in group I individuals compared to group II individuals.

Further, Independent sample t-test was performed to compare group I and group II on AVQI and its constituent parameters. The results demonstrated a significant difference between group I and group II in Shimmer Local ($t(119) = -2.15$, $p < 0.05$). That is, Shimmer Local was significantly higher in group II compared to group I.

4.3. Within group comparison for gender difference

Within the groups, males had slightly higher mean CPPS, Shimmer Local, Shimmer dB, and LTAS compared to females in both the groups. Females had higher mean HNR compared to males in both the groups. Males and females obtained relatively similar Tilt of LTAS values in both the age groups. The SD values of CPPS, HNR, Shimmer Local, Shimmer dB, and Tilt of LTAS was found to be higher in group II compared to group I. The SD values of LTAS was higher in group I compared to group II.

Independent sample t-test which was carried out to find the significant difference between males and females participants on the AVQI and its constituent parameters that revealed a significant difference between gender on AVQI ($t(119) = -2.55$, $p < 0.05$), HNR ($t(119) = 6.02$, $p < 0.001$), Shimmer Local ($t(119) = -4.04$, $p < 0.001$), Shimmer dB ($t(119) = -3.69$, $p < 0.001$), and LTAS ($t(119) = -4.28$, $p < 0.001$). There was no gender difference found on CPPS and tilt of LTAS parameters.

Table 4.3

Within group comparison for gender difference

Parameters	Males	Females
AVQI	2.92	2.58
CPPS	14.31	14.07
HNR	15.87	17.87
Shimmer dB	0.77	0.74
LTAS	20.16	17.64
Tilt of LTAS	14.29	14.28

On observing the mean values of within group comparison of gender difference in table 4.3, the total mean AVQI, Shimmer dB and LTAS were found to be significantly higher in males compared to females. In CPPS, and Tilt of LTAS males have higher value compared to females but it's not statistically significantly. The total mean HNR was observed to have significantly higher value in females compared to males.

Two-way MANOVA was performed to examine the effect of age and gender on the AVQI value and its constituent parameters. Results revealed no significant main effect of age on the AVQI value and its constituent parameters except for Shimmer Local ($F(1,117)=6.27$, $p<0.05$). Group II had significantly higher Shimmer Local compared to Group I. There was a significant main effect of gender on AVQI value ($F(1,117)=6.41$, $p<0.05$), HNR ($F(1,117)=35.48$, $p<0.001$), Shimmer Local ($F(1,117)=17.74$, $p<0.001$), Shimmer dB ($F(1,117)=14.02$, $p<0.001$), and LTAS ($F(1,117)=15.72$, $p<0.001$). The mean AVQI, Shimmer Local, Shimmer dB and LTAS was found to be significantly higher for males compared to females. HNR was found to be significantly higher in females compared to males. However, there was no significant main effect of gender on

CPPS and Tilt of LTAS measures. There was no significant interaction effect between age and gender on the AVQI value and its constituent parameters.

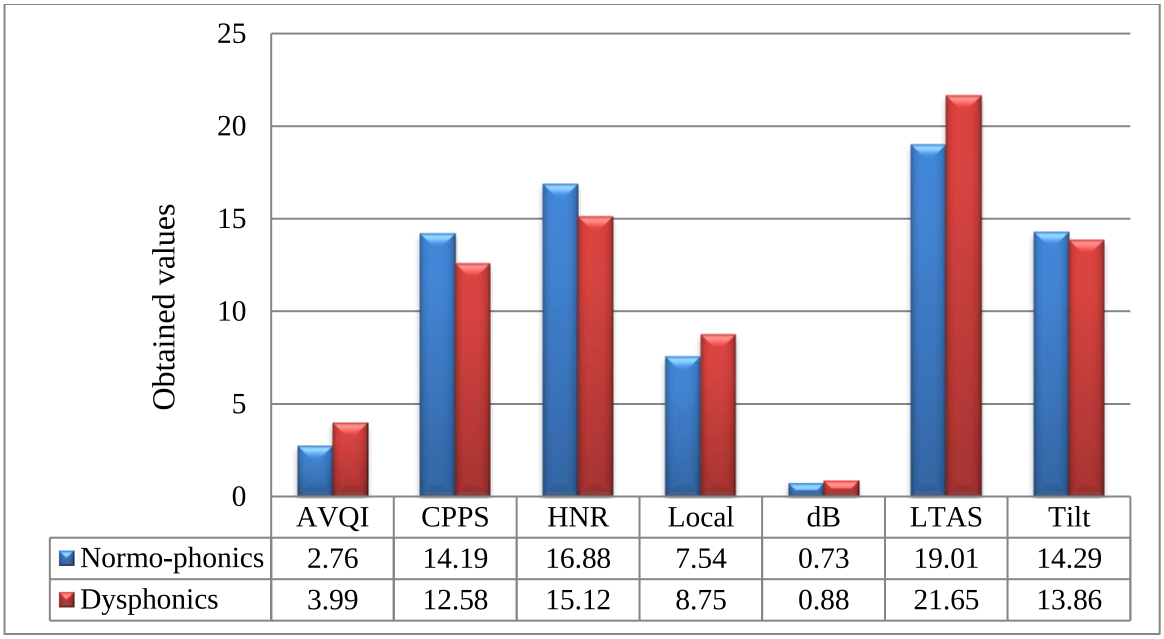
4.4. Comparison between normo-phonics and dysphonics on AVQI and its constituent parameters.

The AVQI and its constituent parameters were analyzed in dysphonic individuals (*N*=15) i.e. group III. The mean (M) and standard deviation (SD) values of AVQI measures are tabulated in table 4.4. Also, the differences in AVQI and its constituent parameters between normo-phonic and dysphonic individuals are presented in figure 4.1.

Table 4.4

Mean (M) and standard deviation (SD) values of AVQI and its constituent parameters in dysphonic individuals (group III)

Measures	M	SD
AVQI	3.99	1.01
CPPS	12.58	2.15
HNR	15.12	2.08
Local	8.75	2.26
dB	0.88	0.15
LTAS	21.65	4.79
Tilt	13.86	0.65



	AVQI	CPPS	HNR	Local	dB	LTAS	Tilt
Normo-phonics	2.76	14.19	16.88	7.54	0.73	19.01	14.29
Dysphonics	3.99	12.58	15.12	8.75	0.88	21.65	13.86

Figure 4.1: AVQI and its constituent parameters between normo-phonic and dysphonic individuals

Independent sample t-test was performed to find out whether the AVQI and its constituent parameters are statistically different between normo-phonic and dysphonic individuals though the number of individuals in each group is different. The results revealed a statistically significant difference between group I, II and group III in AVQI and its constituent parameters which is shown in table 4.4.

The results of Independent sample t-test for group comparison (group I,II, Vs III) depicting ‘t’ and ‘p’ value are represented in table 4.5, shows that AVQI and its constituent parameters are significantly different between normo-phonics and dysphonics.

Table 4.5

Results of Independent sample t-test

Acoustic measures	**t**	**p**
AVQI	-5.58	0.000
CPPS	4.39	0.000
HNR	3.09	0.002
Shimmer Local	-2.51	0.013
Shimmer dB	-4.32	0.000
LTAS	-2.74	0.007
Tilt of LTAS	3.01	0.003

4.3. Reference measures of mean AVQI and its constituent parameters

Table 4.6 gives the norm reference data of AVQI parameters for younger (20-35 years) and relatively older (36-50 years) individuals separately for males and females in Tamil language. This preliminary data would serve as standard reference data collectively. Based on the statistical test results, the reference measures of AVQI and its constituent parameters in native Tamil speaking adults are given in table 4.6.

Table 4.6

Reference measures of mean AVQI and its constituent parameters in native Tamil speaking phono-normal adults

	20-35 years		**36-50 years**	
	Males	**Females**	**Males**	**Females**
AVQI	3.04(0.74)	2.68(0.65)	2.81(0.76)	2.49(0.82)
CPPS	14.31(1.20)	13.95(1.13)	14.32(1.10)	14.20(1.43)
HNR	15.89(1.89)	17.93(1.68)	15.85(1.86)	17.81(1.92)
Shimmer Local	7.66(1.37)	6.76(1.63)	8.67(1.28)	7.17(1.87)
Shimmer dB	0.75(0.09)	0.69(0.10)	0.79(0.09)	0.70(0.13)
LTAS	19.97(3.01)	17.98(3.88)	20.36(2.29)	17.79(3.14)
Tilt of LTAS	14.22(0.50)	14.31(0.48)	14.37(0.59)	14.26(0.42)

*Values in parenthesis are standard deviation

CHAPTER V

DISCUSSION

The primary objective of the present study is to establish standard reference date for Acoustic Voice Quality Index in Tamil speaking normo-phonic adults in the age range of 20 to 50 years. The secondary objective of the study is to investigate the effect of age and gender on AVQI and its constituent parameters. Also, the study further compares the obtained AVQI value with few dysphonic individuals for validity (field testing). The results obtained are discussed in the further sections.

5.1. AVQI in Tamil language

The mean value of AVQI obtained for native Tamil speakers in the age range of 20 – 50 years was 2.76 (±0.76). This is in concordance with the values reported in the literature for AVQI in Indian languages for example, 3.00 in Malayalam and 3.03 in Kannada (Benoy, 2017). This is also in support with the AVQI values computed in different western languages i.e., 2.70 in German (Barsties & Maryn, 2012); 3.25 in English (Maryn et al, 2014); 3.07 in French (Maryn et al., 2014); 2.80 in Dutch (Barsties & Maryn, 2015); 2.97 in Lithuanian (Uloza et al., 2017); 3.12 in Japanese (Hosokawa et al., 2017); 3.33 in Korean (Kim et al., 2018b). This indicates that AVQI in Tamil language is well within the range of established normative findings across different languages and is observed to be least affected by the language characteristics. The AVQI obtained in the present study is 2.76 which differ from 3.00 in Malayalam and 3.03 in Kannada (Benoy, 2017). The variations in the AVQI values across these three Indian languages may be attributed to the methodological differences, linguistic background and cultural adaptations. Only fifty participants each involved in Malayalam and Kannada language of Benoy's (2017) study. Whereas, the present study employed a total of 121 phono-normal individuals. The present also employed study narrowed age group (20-50 years) when compared other studies. In a study done by Maryn, Corthals, Van Cauwenberge Roy, & De Bodt, (2010), they had taken a wide age group of about 8 to 85 years; Maryn, De Bodt, Barsties, & Roy (2014) and also, they had taken 10 to 77 years old individuals.

5.1.1. AVQI across age and gender.

As physical, psychological and social factors change, age and gender affect the voice (Sataloff, 2005). There was no significant effect of age found on AVQI measure in the present study. This is in support with Benoy (2017) study where in there was no change reported in the AVQI measures across age (20-50 years). Latoszek et al. (2017) also reported no effect of age on AVQI scores. This means that AVQI is more independent from age-related changes in vocal fold anatomy or physiology. There was a significant difference was observed on AVQI values between males and females in the present study. Males have higher AVQI values (2.92) than females (2.58). These findings are contradicting with the previous studies by Maryn et al. (2010); Reynolds et al. (2012); Barsties and Maryn (2012); Hosokawa et al. (2017); Benoy (2017) and Latoszek et al. (2017). This shows that AVQI is sensitive to the gender related difference in vocal physiology and vocal anatomy. While the present study showed a significant gender difference on AVQI, the earlier studies did not show gender difference. The earlier studies which did not show a significant gender difference on AVQI included male and female participants in an unequal ratio (103 males and 164 females in Hosokawa et al., 2017; 783 males and 741 females in Kim et al., 2018). This may probably be attributed to gender the difference not noticed in the western studies when compared to the present study.

5.1.2. Constituent parameters of AVQI in Tamil language

5.1.2.1. Across age

In Group II individuals, the CPPS was found to be higher than in group I individuals, but the difference was not statistically significant. CPPS is a robust measure of the severity of dysphonia. It measures the periodicity of the voice signal and the degree of harmonicity. Monnappa, and Balasubramanium (2015) have reported that cepstral measures increase with aging indicating that the hormonal structure of voice is affected. The result of the present study indicates an increase in the mean CPPS for group II individuals which could probably suggest the presence of age related vocal changes.

HNR measures the relative quantity of the vocal signal to additive noise. This additive noise arises from turbulent airflow during voice production (measured on phonation) generated at the level of vocal folds. The resulting noise is reflected in the spectrum as a higher level of noise. Aperiodic vocal fold vibration may also result in the noise. The ratio of periodicity to aperiodicity is therefore HNR. High HNR is connected to sonorous and harmonic voice produced through the vocal folds due to increased airflow from the lungs. On the other hands low HNR denotes relatively weak voice. In the present study HNR is higher in group I individuals who were relatively young compared to group II individuals (Ferrand, 2002; Teixeira, Oliveira, & Lopes, 2013). However this difference was not statistically significant.

Shimmer local is the average absolute difference, divided by the average amplitude, between the amplitudes of two consecutive periods. The higher the shimmer local the more perturberated is the voice (Freitas, Pestana, Almeida, & Ferreira, 2015). The present study indicates a statistically significant higher Shimmer local for group II (older group) compared to group I (younger group). The age related voice changes has resulted instability in the loudness of voice quality leading to perturberated voice which is reflected in shimmer local measures. The physiological laryngeal changes in aging voices are associated with higher Shimmer local (Brockmann, Storck. Carding & Drinnan, 2008; Hodge, Colton, & Kelly, 2001).

Shimmer dB quantifies the average absolute difference between two consecutive periods of the base 10 logarithm and is called ShdB. Shimmer local is expressed in percentage where as Shimmer dB expressed in dB. 3.810% (Shimmer local) is a threshold for vocal pathology and 0.350 dB (Shimmer dB) is set as a threshold for vocal pathology. The shimmer changes reflect the vocal folds with decreased glottal resistance and mass lesion and are associated with breathiness. Although the difference was not statistically significant, the present study observed slightly higher Shimmer dB in group II compared to group I. Although, there was a significant difference between two groups for Shimmer local, shimmer dB did not show such difference between two age groups. There was no significant difference in Shimmer dB, it may be because the age range of the participants considered in the group II was 36-50 years.

The LTAS which displays the average sound level reflects the glottal and vocal tract characteristics. However LTAS is known to be affected by vocal loudness i.e. increased vocal loudness might result in larger LTAS (Nordenberg & Sunburg, 2004).

Therefore, LTAS also depends on the sub glottal pressure. The LTAS and tilt of LTAS was found to be higher in group II individuals compared group I. However, LTAS & Tilt of LTAS difference between two groups was not statistically significant.

4.1.2.1. Across gender

Males had higher CPPS, Shimmer Local, Shimmer dB, LTAS, and Tilt of LTAS compared to females where as HNR was found to be higher in females than males. A significant gender effect was observed on HNR, Shimmer Local, Shimmer dB, and LTAS parameters. Higher mean CPPS, Shimmer Local, Shimmer dB, LTAS, and Tilt of LTAS values for males have been reported by Benoy (2017) in the Indian context. However, CPPS and Tilt of LTAS did not vary significantly between genders in the present study. Goy, Fermandes, Pichora-Fuller, and Van Lieshout (2013) revealed that females have higher value of HNR compared to that of males due to physical and structural variations present across the genders. The present study is in consensus with the above study where the HNR which measures the noise in the voice signal is higher for females. This could be due to the presence of posterior chink in the female vocal folds. However, this is in contrast with Yumoto, Gould, and Bare (1982) where they found males to have higher HNR compared to females.

According to Nordemberg and Sundberg (2004), LTAS reflects the contribution of glottal source and vocal tact for the quality of voice. LTAS was found to be higher in males compared to females in the present study. This may be due to stronger glottal source and longer vocal tract in males compared to females which have led to a higher LTAS values in males. Further, Shimmer Local and Shimmer dB also were significantly higher in males compared to females. Female voices are generally high pitched and aspirated than male voices. This adds on to the noise component of the female voices which reflects insufficient sub glottal pressure. Subglottal pressure determines the loudness of the voice. Hence females have reduced Shimmer (Local & dB) values compared to male voices.

4.2 Comparison of AVQI and its constituent parameters between normo-phonics and dysphonics

AVQI is a sensitive tool to identify pathological voice. The present study compared the mean AVQI between normo-phonics and dysphonics. The mean AVQI was observed to be higher in dysphonics (3.99) than normo-phonics (2.79). This is in agreement with existing literature on AVQI including, Maryn et al. (2010), Reynolds et al. (2012), Hosokawa et al. (2017), Uloza et al. (2017), and Benoy (2017). The mean CPPS was found to be higher in normo-phonics compared to dysphonics. This is in support with study done by Heman-Ackah et al. (2014). CPPS value has a limitation based on the pause and unvoiced segments in the sample recorded for the acoustic analysis (Kim et al., 2018). Since the dysphonic voices do not have continuous voiced segments, CPPS in the dysphonic group is reduced in the present study.

Similarly, the mean HNR and Tilt of LTAS were found to be significantly higher in normo-phonics compared to dysphonics. HNR characterises the ratio between the periodic component and the additional noise coming through the vocal folds. Higher HNR indicates a regular vocal fold periodicity. Hence the dysphonic group in the present study showed a lower HNR indicating vocal fold aperiodicity (Felippe, Grillo, & Greche, 2006). Higher mean values of HNR, and Tilt of LTAS in normo-phonic group reflects a good voice quality.

Shimmer Local, Shimmer dB and LTAS were found to be significantly higher in dyshonics compared to normo-phonics. Most of the dysphonic voices relate to hoarseness, breathiness and roughness (Teixeira, Oliveria, & Lopes, 2013). This affects the sub glottal pressure, glottal resistances and loudness in the voice quality which gets reflected in the Shimmer parameters (Local & dB). Hence, Shimmer dB and Shimmer Local parameters are higher in dysphonic group. Since LTAS reflects the glottal or source and vocal tract characteristics, the dysphonic voice is certain to have these characteristics affected leading to higher LTAS in dysphonics compared to normo-phonics (Lofqvist & Mandersson, 1987).

CHAPTER VI

SUMMARY AND CONCLUSIONS

Traditionally, voice quality assessment is based on measuring sustained vowels. Acoustic and perceptual evaluation of continuous speech sample should also be considered in order to improve ecological validity. Due to its non-invasiveness, and relatively low cost, acoustic voice analysis has gradually received attention as an objective voice quality measure over the last decades. However, acoustic voice analysis should ideally include both sustained vowels and continuous speech in order to ensure ecological validity of such measurements. On the other hand, some linguistic variations may be sensitive to voice assessment methods, which consider a combination of sustained vowels and continuous speech. Therefore, the implementation of speech samples might induce linguistic differences that should be identified and accounted for in process of voice analysis and voice assessment.

AVQI developed by Maryn et al. (2010), which uses concatenated speech and phonation samples. It's a multiparametric measure derived by the weighted combination parameters such as CPPS, HNR, Shimmer local, Shimmer dB, Ltas and Tilt of LTAS. The AVQI possesses concurrent validity, concurrent crossvalidity, diagnostic precision, and is sensitive to change. Furthermore, AVQI can be computed using freely available computer programs as in Praat and Speech Tool; it has the additional advantage of easy access and affordability. Therefore, the AVQI seems to be beneficial in describing differences in voice quality status and discriminating normal and dysphonic voices. AVQI norms were developed in western countries languages i.e., 2.70 in German

(Barsties & Maryn, 2012); 3.25 in English (Maryn et al, 2014); 3.07 in French (Maryn et al., 2014); 2.80 in Dutch (Barsties & Maryn, 2015); 2.97 in Lithuanian (Uloza et al., 2017); 3.12 in Japanese (Hosokawa et al., 2017); 3.33 in Korean (Kim et al., 2018b). Further, the usefulness of AVQI has also played a major role in dysphonic population where the effect of therapeutic management could be monitored prior and post therapy. AVQI is used as an objective measurement of voice for paediatric population who born less than 25 weeks gestation and who underwent endotracheal intubation immediately after birth. In recent studies AVQI was also probe into professional voice users – theatre artist. Such studies in professional voice users helped to identify the risk factors for developing vocal pathology and thus highlighted the importance of vocal hygiene habits.

In this regard, the present study aimed at establishing reference AVQI for Tamil language. Further, the effect of age and gender on AVQI and its constituent parameters were investigated. The study considered Tamil speaking adults in the age range of 20-50 years. A total of 136 individuals participated in the study of whom 121 individuals had normal voice quality and 15 individuals had dysphonia. These 121 participants were in the age range 20–50 years (*M*=34.02 years; *SD*=10.98 years). Further, based on the age range, these participants were divided into two groups. The first group (Group I) consisted of 64 normo-phonic individuals aged 20–35 years with 32 females and 32 males (*M*=33.52 years; *SD*=11.26 years) and the second group (Group II) consisted of 57 normo-phonic individuals aged 36–50 years with 29 females and 28 males (*M*=43.39 years; *SD*=5.79). The third group consisted of 15 dysphonic individuals aged 20-50 years (*M*=38.47 years; *SD*=15.88 years) with 5 females and 10 males. A phonation sample of vowel /a/ for a minimum of 5 to 6 seconds and a continuous speech sample were recorded from all the participants. The standardized Tamil passage (Savithri & Jayaram, 2005) was used for the reading task (continuous speech). The participants were asked to read nine sentences, the first three and the last three sentences were excluded, and the middle three sentences were considered, which contains 57 syllables. The recording was carried out in a quiet room/noise free place within the premises of their working setup. The participants were asked to sit comfortably on a chair. A distance of 10 cm was maintained between mouth to the microphone to avoid breathing noise. Sennheiser CX 275s earphones with the microphone were used, which was recorded directly in the Praat software (6.0.28 version) using a Lenovo Laptop. The recorded samples were saved in .wav file format and renamed with convenient codes. Praat software analyzed the samples (6.0.28). AVQI (v02.02) was measured using the Maryn and Weenik (2015) algorithm. The script of AVQI has an algorithm which contains the subsequent formula, AVQI = 9.072-0.245*CPPS-0.161*HNR-0.470*SL+6.158*ShdB-0.071*Slope+0.170*Tilt was used. To elicit AVQI, the script given by Maryn et al. (2010), was copied into a text file and saved as .text. Finally, the saved .wav files were opened in the Praat software. Along with these files, the script file was also opened by clicking the Praat option-open Praat script- select the '.text.' Script and the AVQI values and its constituent parameters were extracted after the algorithm 'runs' in the software. The following measures were documented.

h) Acoustic Voice Quality Index (AVQI)

i) Smoothened Cepstral Peak Prominence (CPPs)

j) Harmonic to Noise Ratio (HNR)

k) Shimmerlocal (SL)

l) ShimmerdB (shdB)

m) Slope of Long term average spectrum (LTAS)

n) The tilt of the regression line through the Long-term average spectrum (Tilt)

The result revealed that the normative value of AVQI for the Tamil speaking population is 2.76 (*SD*=0.77). This value is in agreement with AVQI values reported across the world for several other languages. The means AVQI value for group I (20-35 years) was 2.87 (*SD*=0.72) and group II (36-50 years) was 2.65 (*SD*=0.81). AVQI did not show any difference across age but evidenced gender differences. This implies that AVQI is independent of age related physiological voice changes. However, AVQI was found to be sensitive to the gender based difference in structural and functional aspects of laryngeal structures. In addition to AVQI, its constituent parameters were also analysed. The CPPS and HNR was higher in group I individuals compared group II where as Shimmer Local, Shimmer dB and LTAS was higher in group II compared to group I. Similarly, the constituent parameters were compared between genders. The Shimmer Local, Shimmer dB and LTAS were significantly higher in males compared to females. Females showed a significant higher HNR compared to males. These differences across age and between genders are attributed to the physiological and anatomical changes of vocal fold and vocal tracts. The AVQI and its constituent parameters were compared between normo-phonics and dysphonic individuals. Since AVQI is known to be a sensitive in identifying dysphonic voice, its clinical implication is supported and strengthened by the finding of the present study. The present study provided AVQI value of 2.76 for normo-phonics and 3.99 for dysphonics. The AVQI constituent parameters like Shimmer Local, Shimmer dB and LTAS were significantly higher in dysphonics and parameters like CPPS, HNR and Tilt of LTAS were significantly lower in dysphonics.

The obtained results between normo-phonics and dysphonics needs to be verified by considering large number of voice patients. Thus, the present study adds on to the literature the reference measures of AVQI for Indian population.

Implications of the study

1. The outcome of the present study would serve as norm reference data of AVQI values in Tamil speaking population.
2. The outcome of the present study enlightens the understanding of the influence of age and gender on AVQI value and its component parameters. Hence, there was a significant higher Shimmer Local in middle aged individuals compared to younger individuals. There was a significant influence of gender on AVQI, and its constituent parameters like HNR, Shimmer local, Shimmer dB and LTAS. Also, there was no gender difference found for CPPS and Tilt of LTAS parameters which needs to be confirmed with larger samples.
3. The result of the present study would serve as a baseline for dysphonic patients who undergo therapeutic management where therapy progress can be monitored.

Limitation of the study

1. The sample size of dysphonics group was relatively less (*N*=15) compared to normo-phonics group (*N*=121).

Future directions

1. The present study developed reference measures only for individuals within the age range of 20-50 years, future studies can investigate the AVQI measures in other age groups such as in children and in geriatric population.

2. AVQI measure can be profiled for individuals who are undergoing treatment, before and after intervention.
3. AVQI can be examined in various professional voice users and establish clinical norms.
4. AVQI can be investigated in other regional Indian languages to see whether different language has any effect on AVQI.

References

Barsties, V., Latoszek, B., Ulozaitė-Stanienė, N., Petrauskas, T., Uloza, V., & Maryn, Y. (2019). Diagnostic Accuracy of Dysphonia Classification of DSI and AVQI. *The Laryngoscope*, *129*(3), 692-698.

Benoy, J. J. (2017). *Acoustic voice quality index (AVQI) and perceptual measures in the Indian population*. Master's Dissertation submitted to the University of Mysore, Mysore.

Brockmann, M., Storck, C., Carding, P. N., & Drinnan, M. J. (2008). Voice loudness and gender effects on jitter and shimmer in healthy adults. *Journal of Speech, Language, and Hearing Research*, 51(1),1152-1160.

Carding, P. N., Wilson, J. A., MacKenzie, K., & Deary, I. J. (2009). Measuring voice outcomes: state of the science review. *The Journal of Laryngology & Otology*, *123*(8), 823-829.

Casper, J. K., & Leonard, R. (2006). *Understanding voice problems: A physiological perspective for diagnosis and treatment*. Lippincott Williams & Wilkins.

Christmann, M. K., Brancalioni, A. R., Freitas, C. R. D., Vargas, D. Z., Keske-Soares, M., Mezzomo, C. L., & Mota, H. B. (2015). Use of the program MDVP in different contexts: a literature review. *Revista CEFAC*, *17*(4), 1341-1349.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969). Differential diagnostic patterns of dysarthria. *Journal of speech and hearing research*, *12*(2), 246-269.

Deary, I. J., Wilson, J. A., Carding, P. N., & MacKenzie, K. (2003). VoiSS: a patient-derived voice symptom scale. *Journal of Psychosomatic Research*, *54*(5), 483-489.

Dejonckere, P. H., & Lebacq, J. (1996). Acoustic, perceptual, aerodynamic and anatomical correlations in voice pathology. O*torhinolaryngology*, *58*(6), 326-332.

Dejonckere, P. H., Bradley, P., Clemente, P., Cornut, G., Crevier-Buchman, L., Friedrich, G., Heyning, P. V. D., Remacle, M., & Woisard, V. (2001). A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. *European Archives of Oto-rhino-laryngology*, *258*(2), 77-82.

D'haeseleer, E., Meerschman, I., Claeys, S., Leyns, C., Daelman, J., & Van Lierde, K. (2017). Vocal quality in theater actors. *Journal of Voice*, 31(4), 510-e7.

Felippe, A. C. N. D., Grillo, M. H. M. M., & Grechi, T. H. (2006). Standardization of acoustic measures for normal voice patterns. *Revista Brasileira de Otorrinolaringologia*, *72*(5), 659-664.

Ferrand, C. T. (2002). Harmonics-to-noise ratio: an index of vocal aging. *Journal of voice*, *16*(4), 480-487.

Freitas, S. V., Pestana, P. M., Almeida, V., & Ferreira, A. (2015). Integrating voice evaluation: correlation between acoustic and audio-perceptual measures. *Journal of Voice*, *29*(3), 390-e1-e7.

Goy, H., Fernandes, D. N., Pichora-Fuller, M. K., & van Lieshout, P. (2013). Normative voice data for younger and older adults. *Journal of Voice*, *27*(5), 545-555.

Hadjitodorov, S., & Mitev, P. (2002). A computer system for acoustic analysis of pathological voices and laryngeal diseases screening. *Medical Engineering & Physics*, *24*(6), 419-429.

Hakkesteegt, M. M., Brocaar, M. P., Wieringa, M. H., & Feenstra, L. (2008). The relationship between perceptual evaluation and objective multiparametric evaluation of dysphonia severity. *Journal of Voice*, *22*(2), 138-145.

Heman-Ackah, Y. D., Sataloff, R. T., Laureyns, G., Lurie, D., Michael, D. D., Heuer, R., & Lyons, K. (2014). Quantifying the cepstral peak prominence, a measure of dysphonia. *Journal of Voice*, *28*(6), 783-788.

Hirano, M. (1981). Clinical examination of voice. *Disorders of Human Communication*, *5*, 1-99.

Hirano, M., Hibi, S., Yoshida, T., Hirade, Y., Kasuya, H., & Kikuchi, Y. (1988). Acoustic analysis of pathological voice: some results of clinical application. *Acta oto-laryngologica*, *105*(5-6), 432-438.

Hogikyan, N. D., & Sethuraman, G. (1999). Validation of an instrument to measure voice-related quality of life (V-RQOL). *Journal of Voice*, *13*(4), 557-569.

Hosokawa, K., Barsties, B., Iwahashi, T., Iwahashi, M., Kato, C., Iwaki, S., Sasi, H., Miyuachi, A., Matsushiro, N., Inohara, H., & Ogawa, M. (2017). Validation of the acoustic voice quality index in the Japanese language. *Journal of Voice*, *31*(2), 260-e1 to 260 e8.

Jacobson, B. H., Johnson, A., Grywalski, C., Silbergleit, A., Jacobson, G., Benninger, M. S., & Newman, C. W. (1997). The voice handicap index (VHI) development and validation. *American Journal of Speech-Language Pathology*, *6*(3), 66-70.

Jayakumar, T., & Savithri, S. R. (2012). Effect of geographical and ethnic variation on Dysphonia Severity Index: a study of the Indian population. *Journal of Voice*, *26*(1), e11-e16.

Jiang, J. J., & Titze, I. R. (1994). Measurement of vocal fold intraglottal pressure and impact stress. *Journal of Voice, 8*(2), 132-144.

Kankare, E., B. Barsties, Y. Maryn, I. Ilomäki, A. M. Laukkanen, J. Tyrmi, L. Rantala et al. (2015). "A preliminary study of the Acoustic Voice Quality Index in Finnish speaking population." In 11th *Pan European Voice Conference*; 2015 *August 31th—September 4th; Florence, p. 218.* Firenze University Press, Italy.

Kempster, G. B., Gerratt, B. R., Veralolini, A. K., Barkmeier-Kraemer, J. B., & Hillman, R. E. (2009). Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol. *American Journal of Speech-Language Pathology*, *18*(2), 124-132.

Kim, G. H., Lee, Y. W., Bae, I. H., Park, H. J., Lee, B. J., & Kwon, S. B. (2018). Comparison of Two Versions of the Acoustic Voice Quality Index for Quantification of Dysphonia Severity. *Journal of Voice* (in press).

Kim, G. H., Lee, Y. W., Bae, I. H., Park, H. J., Wang, S. G., & Kwon, S. B. (2018). Validation of the acoustic voice quality index in the Korean language. *Journal of Voice*, 1-9.

Kim, G. H., Lee, Y. W., Park, H. J., Bae, I. H., Lee, B. J., & Kwon, S. B. (2017). Auditory-Perceptual and Acoustic Evaluation in Measuring Dysphonia Severity of Vocal Cord Paralysis. *Journal of the Korean Society of Laryngology Phoniatrics and Logopedics*, 28(2), 106-111.

Kreiman, J., & Sidtis, D. (2011). *Foundations of voice studies: An interdisciplinary approach to voice production and perception*. John Wiley & Sons.

Kreiman, J., Gerratt, B. R., Garellek, M., Samlan, R., & Zhang, Z. (2014). Toward a unified theory of voice production and perception. *Loquens*, *1*(1).

Krishnamurthi, B. (2018). *Encyclopedia Britannica*. Retrieved on 11.08.2018 from https://www.britannica.com/topic/Tamil-language.

Latoszek, B. B. V., Ulozaitė-Stanienė, N., Maryn, Y., Petrauskas, T., & Uloza, V. (2017). The Influence of Gender and Age on the Acoustic Voice Quality Index and Dysphonia Severity Index: A Normative Study. *Journal of Voice* (in press), doi:10.1016/jvoice2017.11011.

Laver, J., Hiller, S., & Beck, J. M. (1992). Acoustic waveform perturbations and voice disorders. *Journal of Voice*, *6*(2), 115-126.

Lee, L., Stemple, J. C., Glaze, L., & Kelchner, L. N. (2004). Quick screen for voice and supplementary documents for identifying pediatric voice disorders. *Language, Speech, and Hearing Services in Schools*, 34(4):308-49.

Leyns, J. D. C. (2017). *Vocal Quality In Theatre Actors.* Doctoral dissertation submitted to Ghent University.

Löfqvist, A., & Mandersson, B. (1987). Long-time average spectrum of speech and voice analysis. *Folia phoniatrica et logopaedica*, *39*(5), 221-229.

Maryn, Y., & Weenink, D. (2015). Objective dysphonia measures in the program Praat: smoothed cepstral peak prominence and acoustic voice quality index. *Journal of Voice*, 29(1), 35-43.

Maryn, Y., Corthals, P., Van Cauwenberge, P., Roy, N., & De Bodt, M. (2010). Toward improved ecological validity in the acoustic measurement of overall voice quality:

combining continuous speech and sustained vowels. *Journal of voice*, *24*(5), 540-555.

Maryn, Y., De Bodt, M., & Roy, N. (2010). The Acoustic Voice Quality Index: toward improved treatment outcomes assessment in voice disorders. *Journal of communication disorders*, 43(3), 161-174.

Maryn, Y., De Bodt, M., Barsties, B., & Roy, N. (2014). The value of the Acoustic Voice Quality Index as a measure of dysphonia severity in subjects speaking different languages. *European Archives of Oto-Rhino-Laryngology*, *271*(6), 1609-1619.

Maryn, Y., Kim, H. T., & Kim, J. (2016). Auditory-perceptual and acoustic methods in measuring dysphonia severity of Korean speech. *Journal of Voice*, *30*(5), 587-594.

Mehta, D. D., & Hillman, R. E. (2008). Voice assessment: updates on perceptual, acoustic, aerodynamic, and endoscopic imaging methods. *Current Opinion In Otolaryngology & Head And Neck Surgery*, *16*(3), 211.

Mehta, D. D., Deliyski, D. D., Quatieri, T. F., & Hillman, R. E. (2011). Automated measurement of vocal fold vibratory asymmetry from high-speed videoendoscopy recordings. *Journal of Speech, Language, and Hearing Research*, 54(1):45-54.

Monnappa, D., & Balasubramanium, R. K. (2015). Cepstral analysis of voice in healthy aged individuals. *Journal of Laryngology and Voice*, *5*(2), 34.

Nordenberg, M., & Sundberg, J. (2004). Effect on LTAS of vocal loudness variation. *Logopedics Phoniatrics Vocology*, *29*(4), 183-191.

Núñez-Batalla, F., Díaz-Fresno, E., Álvarez-Fernández, A., Cordero, G. M., & Pendás, J. L. L. (2017). Application of the acoustic voice quality index for objective measurement of dysphonia severity. *Acta Otorrinolaringologica*, *68*(4), 204-211.

Rabinov, C. R., Kreiman, J., Gerratt, B.R., & Bielamowiez, S. (1995). Comparing reliability of perceptual ratings of roughness and acoustic measures of jitter. *Journal of Speech, Language, and Hearing Research*, *38*(1), 26-32.

Rameshkumar, E., & Rosmi, T. K. (2017). Prevalence of age, gender and pathological conditions of vocal cords leading to hoarseness of voice in a tertiary care hospital. *International Journal of Advances in Medicine*, *3*(2), 345-348.

Ramya, M., & Goswami, S. P. (2009). *Language Proficiency Questionnaire: An adaptation of LEAP-Q in Indian context.* Master's Dissertation submitted to University of Mysore, Mysuru.

Reynolds, V., Buckland, A., Bailey, J., Lipscombe, J., Nathan, E., Vijayasekaran, S., Kelly, R., Maryn, Y., & French, N. (2012). Objective assessment of pediatric voice disorders with the acoustic voice quality index. *Journal of Voice*, 26(5), 672-e1-672.e7.

Reynolds, V., Meldrum, S., Simmer, K., Vijayasekaran, S., & French, N. P. (2014). Dysphonia in preterm children: assessing incidence and response to treatment. *Contemporary Clinical Trials*, 37(2), 170-175.

Sataloff, R. T. (2017). *Professional Voice: The Science and Art of Clinical Care, 3-Volume Set*. Plural Publishing.

Seshasri, D. (2018). *Acoustic voice quality index in Kannada speaking children in the age range of 10 to 12 years*. Masters Dissertation submitted to the University of Mysore, Mysuru.

Sundberg J. (1987). *The Science of the Singing Voice* (Northern Illinois University, DeKalb), Chaps. 2 and 4. Northern Illinois University Press ISBN: 978-0-87580-542-9.

Teixeira, J. P., Oliveira, C., & Lopes, C. (2013). Vocal acoustic analysis–jitter, shimmer and HNR parameters. *Procedia Technology*, *9*, 1112-1122.

Timmermans, B., De Bodt, M. S., Wuyts, F. L., & Van de Heyning, P. H. (2004). Training outcome in future professional voice users after 18 months of voice training. *Folia phoniatrica et logopaedica*, *56*(2), 120-129.

Uloza, V., Latoszek, B. B. V., Ulozaite-Staniene, N., Petrauskas, T., & Maryn, Y. (2018). A comparison of Dysphonia Severity Index and Acoustic Voice Quality Index measures in differentiating normal and dysphonic voices. *European Archives of Oto-Rhino-Laryngology*, *275*(4), 949-958.

Uloza, V., Petrauskas, T., Padervinskis, E., Ulozaitė, N., Barsties, B., & Maryn, Y. (2017). Validation of the acoustic voice quality index in the Lithuanian language. *Journal of Voice*, *31*(2), 257.e1 to 257.e11.

Ulozaite, N., Petrauskas, T., Saferis, V., & Ulozas, V. (2016). Correlations between automated dysphonia quantification and perceptual voice evaluation. *Biomedical Engineering*, *20*(1).

Wilson, D. K. (1987). Children’s voice problems. *Voice Problems of children*, *3rd ed*. Philadelphia: Williams & Wilkins, 1-15.

Wolfe, V., Cornell, R., & Fitch, J. (1995). Acoustic prediction of severity in commonly occurring voice prolems. *Journal of Speech, Language and Hearing Research*, 38(2), 297-303

Wuyts, F. L., De Bodt, M. S., Molenberghs, G., Remacle, M., Heylen, L., Millet, B., Van Lierde K, Raes J & Van de Heyning, P. H. (2000). The dysphonia severity index: an objective measure of vocal quality based on a multiparameter approach. *Journal of Speech, Language, and Hearing Research*, *43*(3), 796-809.

Yumoto, E., Gould, W. J., & Baer, T. (1982). Harmonics-to-noise ratio as an index of the degree of hoarseness. *The Journal of the Acoustical Society of America*, *71*(6), 1544-1550.

Zhang, Z. (2016). Mechanics of human voice production and control. *The Journal of the Acoustical Society of America*, *140*(4), 2614-2635.

ANNEXURE-I

ALL INDIA INSTITUTE OF SPEECH AND HEARING: MYSORE-6

Title of the Dissertation: **Acoustic Voice Quality Index (AVQI-T) in Tamil Language**

Informed Consent Form

I have been informed about the aim, objectives and the procedure of the study. Both the reading as well as phonation sample will be audio recorded as a part of the study. A Sennheiser cx 275s earphone will be used to record the samples that will be directly connected to the laptop. The phonation samples will be recorded for about 6 to 7 seconds with three trials and a reading sample (in Tamil) will be recorded for about 3 trials. The recorded sample will be kept confidential for ensuring the anonymity. The outcome of the study will benefit the adult Tamil speaking (Male and Female) individuals to rule out the presence of voice problems by developing a normative database on Acoustic Voice Quality Index (AVQI) in Tamil language.

I, ________________________________ the undersigned give my written consent to be participant of this research study.

Signature of the participant

(Name & Address)

Contact No:

Signature of the Investigator

(Name & Designation)

Ms. Vishali.P

II M.Sc (SLP)