ASSESSMENT OF VOICE CHARACTERISTICS THROUGH VIRTUAL MODE

Madhusudhan B. M.

Reg. No: 19SLP019

Dissertation Submitted in Part Fulfillment of Degree of Master of Science

(Speech-Language Pathology)

University of Mysore

ALL INDIA INSTITUTE OF SPEECH AND HEARING

MANASAGANGOTHRI, MYSURU—570006

SEPTEMBER 2021

CERTIFICATE

This is to certify that the dissertation entitled "**Assessment of voice characteristics through virtual mode**" is a bonafide work submitted in part fulfillment for degree of Master of Science (Speech Language Pathology) of the student Registration Number: **19SLP019**. This has been carried out under the guidance of the faculty of this institute and has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Prof. M. Pushpavathi

Director

All India Institute of Speech and Hearing

Manasagangothri, Mysuru - 570006

Mysuru

September, 2021

CERTIFICATE

This is to certify that the dissertation entitled "**Assessment of voice characteristics through virtual mode**" has been prepared under our supervision and guidance. It is also being certified that this dissertation has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Co-guide

Dr. T. Jayakumar

Associate Professor in Speech Sciences

Department of Speech Language Sciences

Guide

Dr. Ajish K Abraham

Professor of Electronics and Acoustics

Department of Electronics

All India Institute of Speech and Hearing

Manasagangothri, Mysuru – 570006

Mysuru

September, 2021

DECLARATION

This is to certify that the dissertation entitled "**Assessment of voice characteristics through virtual mode**" is the result of my own study under the guidance of Dr. Ajish K Abraham, Professor of Electronics and Acoustics**,** Department of Electronics and co-guidance of Dr. T. Jayakumar, Associate Professor in Speech Sciences, Department of Speech Language Sciences, All India Institute of Speech and Hearing, Mysore, and has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysuru **Register No**: **19SLP019**

September, 2021

Those who walk with God always reach their Destination

Dedicated to

My Grandfather & Grandmother

With love

Acknowledgment

I would like to express my sincere thanks to my guide, ***Prof. Ajish K. Abraham,*** *for the constant support and guidance. Without you this study would not have been possible sir. Your patience, ever smiling face, dedication to the work, and guidance in the research motivated me to carry out this study. Thanks for inspiring hope in me, igniting my imagination, and instilling in me – a love of learning.*

I extend my heartfelt gratitude to my co-guide, ***Dr. Jayakumar T,*** *for guiding me. Thank you, sir, for your support, patient listening and valuable feedback. I have never seen a person as patient as you during class, the way you clear our doubts in Neuroscience and ERP Classes have helped me gain lot of knowledge.*

I've got inspired by both the guides on how to handle work stress by not panicking and how supporting a smile will always gives a motivation to do any work.

I'd also want to express my gratitude to our beloved director, ***Dr. M. Pushpavathi,*** *for allowing me to undertake this research. I also like to express my thanks to* ***Dr. Vasanthalakshmi,*** *for her invaluable assistance.*

I would like to thank my best buddies ***(6K), Jeevan, Shoban, Mithun, Pruthvik and Gadappa*** *for all the beautiful memories at AIISH. I really can't imagine my college and hostel life without these kuchikoos. I cherish each and every moment spent with them. Special thanks to* ***Jeeva*** *who supported me constantly throughout all odd times.*

I would like to thank my dear ***Arpi,*** *without your help, wouldn't have completed my dissertation. You stand like a backbone for me always, ready to help anything related to my studies, clearing my doubts, giving valuable suggestion which motivates me to do a better work. Your passion towards research will certainly take you to the next level. I would like to say All the best & You are the Best.*

How can I forget my sweet sisters ***Kavi, Manasa (Kaju), Ri2, Ankita*** *and brother* ***Sumanth*** *from Channapatna for always being my support strength & being there for me whenever needed.*

Thank you, Biriyani ***Sameera (Sami),*** *such a sweet soul is ready to help and motivate me to do the work. You always provide some positive feedbacks and pray for me that almighty will help me in many ways when am in need.*

I personally thank ***Mahadevprasad*** *who gave me advice on the future opportunities in life. I can still remember those words you always say to me,* ***"PASSION FOR WORK is more important than MONEY".*** *And Special thanks to my friends* ***Vishal, Praju, Manoj (Govi), Sagar*** *who have always been with me and have spent some wonderful memories at Tuition and GSSS college.*

*I would like to thank my c**lassmates Akshit, Sunny, Daddu, Dena, Mallu, Anshu, Brenda, Jiji, Jayamma, Archu, Shy, Hegde, Chethan-Chu, Priya, Biraj, Gayathri, Joel, Sumathi, Tanuja,*** *Last but not the least, Renovators & Renovators 2.0, which includes some of the best company I had alongside with me in my entire journey in AIISH.*

*I sincerely thank all the COOKS **Mahadev uncle, Sidda, JagguBai, Devraj anna.** Special mention to **Devraj Uncle** who cared for me no less than a family member as well used to prepare delicious foods. I would really miss those moments.*

*I would like to thank all **staffs of AIISH, my seniors, juniors**, who have helped me directly or indirectly to gain knowledge.*

A sincere thanks to all the participants of my study for the cooperation throughout the data collection. Without your support, this study wouldn't have been completed.

*I thank my family for supporting me in every situation and decisions taken. **Appa, Amma, Shashi, Kiru, Deepu mam, Naveen Anna, Navyaakka, Baby, GundaAnna**, you'll always be in my heart.*

*****

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

Voice plays a major role in oral, verbal communication between humans. When voice, pitch and loudness vary or are inappropriate for a person's age, gender, cultural context, or geographic area, a voice disorder occurs (Aronson & Bless, 2009).

Voice is the most dominant tool to express both emotions and oral language. Dysphonia occurs due to a faulty structure or improper functioning of the larynx, resulting in the change in quality, pitch and loudness of the voice (Smith, 1996). Boone et al. (2010) classified the voice disorder into three types based on the etiological factor. The first type is the organic voice disorder, which may be due to any laryngeal structure deviation that affects the vocal fold vibration. The second type is the neurogenic voice disorder that induces the vocal fold closure due to either paralysis or exhaustion due to neurological diseases. The third type is the hysterical voice disorder induced by psychosocial influences and voice disorders of muscle tension (muscle tension dysphonia), which may occur with excessive laryngeal muscle use.

1.1 Evaluation of voice disorders – conventional methods

The team for evaluation of dysphonia includes the Speech Language Pathologists (SLPs) and Otolaryngologists (ENT). SLP deals with assessment of an individual with Dysphonia to make a proper diagnosis for voice disorder management. The individual quality of voice can be assessed qualitatively by perception and quantitatively by the instrumentation (Hakkesteegt, 2009). European Laryngeal Society (ELS) recommends laryngostroboscopy, perceptual voice assessment, acoustic analysis, aerodynamic measurements, and subjective as well as self-evaluation of voice to assess voice disorders (Dejonckere, 2001).

Conventional methods used in the clinical setting to evaluate voice disorders include both perceptual evaluation and acoustic analysis. Tools for perceptual assessment of voice quality includes the Darley Rating system (Darley, 1969), Grade, Roughness, Breathiness, Asthenia, and Strain (GRBAS) Scale (Hirano, 1981), Buffalo Voice Screening Profile (Wilson, 1987), and the Consensus Auditory Perceptual Evaluation of Voice (CAPE-V) Scales (Kempster, 2009). Objective evaluation of voice quality includes the acoustic analysis which helps to quantify the degree of severity of Dysphonia. It is a reliable measurement of voice quality (Carding, 2009). Acoustic measures include frequency related measures (eg., fundamental frequency, habitual frequency, frequency range etc.), amplitude related measures (eg., habitual intensity, extent, and fluctuation of intensity etc.), perturbation related measures (e.g., jitter, shimmer, etc.) as well as harmonics related measures (e.g., Harmonics to noise ratio etc.). Widely used by various researchers (Dejonckere & Lebaq, 1996; Hirano et al., 1988), Dysphonia Severity Index (Wuyts et al., 2000) and the Acoustic Voice Quality Index (AVQI) evaluates the acoustic parameters of the voice samples of phonation and speech task (Maryn et al., 2010). 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 across the long-term average spectrum (tilt) are the parameters used for AVQI.

AVQI is estimated as

$$AVQI = 2.571*(3.295-0.111*CPPS-0.073*HNR-0.213*SL+ 2.789*ShdB-0.032*Slope + 0.077*Tilt + 0.077*Tilt).$$

1.2 Online evaluation of voice disorders

In the pandemic situation, it is difficult to carry out the conventional method of perceptual evaluation and acoustic analysis of voice, which requires face to face physical interaction with the patient. Alternative methods such as online evaluation of voice may be explored for assessment of the person with Dysphonia in this pandemic situation.

Different applications (apps) such as Zoom, Google Meet, Skype, WebEx, and WhatsApp are available for tele-conferencing for different purposes such as students' online classes, audio visual interaction with friends and family members. Speech Language Pathologists can use these apps for tele-assessment and tele-management of various speech and language disorders. Smartphone are used for establishing communication using these apps. For tele-assessment, the speech samples will be acquired through these Smartphone. Several researchers have analysed recordings from Smartphone to find their suitability for voice assessment.

Lin et al. (2012) measured and compared the voice and speech samples recorded through the iPhone and Personal Computer Memory Card International Association (PCMCIA) laptop recording system. This study concluded that recording using iPhone is suitable for acoustic analysis for speech & voice assessment. Vogel (2014) compared the recordings of Smartphone, landlines, laptops, and hard disc recorders. Results showed that none of the methods of acquisition for capturing voice and speech could provide statistically equivalent values compared to the high-quality recording system. Yun (2015) recorded the voice of 30 individuals within the age range of 21-40 years with three devices such as Samsung galaxy (G90S6) mobile phone, Computerised Speech Lab (CSL - model 4500), and a digital recorder (Model PCM-M10, Sony, Japan). Parameters such as F0, jitter, shimmer, NHR (Noise-Harmonic ratio), and formant derived through Multi-Dimensional Voice Program (MDVP) were analyzed. The author reported no significant difference in the parameters across the three devices used for recording. The author concluded that the Samsung Galaxy Smartphone could be used for the acoustic analysis of voice.

Lee (2018) compared the recording of 180 individuals. The vowel /a/ samples were recorded using CSL-Model 4150B and android Smartphone (Samsung Galaxy Note 5, SM-N920) through the smart recorder app. No significant differences were observed between the devices and the study concluded that Smartphone can be used as a screening tool for voice disorders. Barche et al. (2020) experimented the automatic detection and assessment of voice disorder using the Saarbruecken Voice database. The results show that open smile features are better for the detection and classification of voice disorder. Vogel (2008) compared the voice samples acquired through three recording devices such as hard drive, solid-state recorder, and computer-based (standard laptop). The finding showed that selective acoustic measurements (e.g., F0, noise-to-harmonic ratio, number of pauses) were efficiently extracted using all three methods.

It is evident from the previous research that voice samples recorded with Smartphone are suitable for acoustic analysis. The previous research also reveals that the MDVP and AVQI parameters listed in Section 3.4.1 and 3.4.2 have been used for acoustic analysis of voice. But in all these studies, the voice samples were directly recorded through Smartphone. In the present scenario, it is difficult to have direct recording, the possibilities of collecting samples through online recording need to be explored. No study has been reported in the literature investigating the suitability of samples collected online through videoconference apps, for acoustic analysis of voice.

1.3 Need of the study

In the present pandemic situation, it is difficult to carry out the conventional assessment, which requires face to face interaction of the clinician with the patient. Tele-assessment of voice may be explored as an alternative method for assessment of persons with Dysphonia. Telepractice has been successfully used for voice therapy as well as for prevention of voice disorders (Castillo-Allendes, 2020; Grillo, 2017). Apps such as Zoom, Google Meet, etc. installed in Smartphone are used for tele-conferencing. For tele-assessment, the speech samples need to be acquired through this online mode. Several researchers (Lee et al., 2018; Lin et al., 2012; Vogel et al., 2014; Yun, 2015) have analysed voice samples recorded directly through Smartphone and found that these recordings are suitable for voice assessment. But none of the researchers have investigated whether the voice samples recorded online using Smartphone can be used to assess voice disorders. There is also a need to find out the most suitable among the widely used apps, for online recording.

1.4 Aim of the study

To determine the suitability of the voice, recorded online through two widely used video conference apps (App1 and App2) for assessment of voice characteristics, by comparing with the acoustic parameters of voice recorded directly through Smartphone, and to find out the most suitable among the two apps.

1.5 Objectives

- To compare the acoustic parameters (AVQI & MDVP) of the voice samples of individuals with normal voice, recorded online through App1 and App2, with the voice samples recorded directly through mobile phone.
- To compare the acoustic parameters (AVQI & MDVP) of the voice samples of individuals with Dysphonia recorded online through App1 and App2, with the voice samples recorded directly through mobile phone.
- To compare the acoustic parameters (AVQI & MDVP) of the voice samples of individuals with normal voice, recorded online through App1 with the samples recorded through App2.
- To compare the acoustic parameters (AVQI & MDVP) of the voice samples of individuals with Dysphonia, recorded online through App1 with the samples recorded through App2.

Chapter 2

REVIEW OF LITERATURE

2.1 Conventional Methods for Voice Evaluation

The conventional methods used to evaluate voice are majorly divided into two categories: - subjective and objective. Subjective methods of assessment include perceptual voice assessment, self-evaluation of voice, etc. Objective assessment methods are further classified as non-invasive and invasive. Non-invasive methods include acoustic analysis and aerodynamic measurements. Invasive methods include laryngoscopy, video-stroboscopy, etc. A brief review of the conventional methods is given in the following subsections.

2.1.1 Subjective Evaluation

2.1.1.1 Perceptual Evaluation

Perceptual evaluation is the 'gold standard' method for assessment of voice quality which provides the accurate diagnosis of the severity of voice quality. The buffalo-III speech profile (Wilson, 1987) was designed specifically for evaluating pediatric voices. It is a five-point equal-appearing interval rating scale, with ‘1’ indicating "normal" and ‘5’ indicating "very severe" variation. It analyses twelve significant aspects of voice production. The 'GRBAS Scale' (Hirano, 1981) evaluates five aspects of vocal quality and the severity is expressed on a four-point scale ranging from ‘0’ to ‘3’ (De Bodt et al., 1997). Consensus auditory perceptual evaluation of voice (CAPE-V) developed by the American Speech and Hearing Association (ASHA, 2002) is a 100 mm visual analog scale that evaluates six parameters such as overall severity, roughness, breathiness, strain, pitch, and loudness.

2.1.1.2 Self Evaluation

Self-assessment is done using voice handicap index (VHI) (Jacobson et al., 1997). A shortened 10-item version named VHI-10 was subsequently developed for quick assessment of VHI (Rosen et al., 2004).

2.1.2 Objective evaluation

2.1.2.1 Non-Invasive methods

Non-invasive methods include aerodynamic and acoustic measures. Aerodynamic parameters include vital capacity, mean airflow rate, phonation duration, vocal intensity, maximum phonation duration and S/Z ratio. Acoustic parameters evaluated are fundamental frequency-related parameters, intensity, perturbation, noise-related parameters and spectral parameters.

The perceptual evaluation combined with acoustic measurement is considered gold standard for assessment of dysphonia and its severity (Ma & Yiu, 2006). Dysphonia Severity Index is given by the equation:-

DSI=0.13 x MPT + 0.0053 x F0-High – 0.26 x I-Low – 1.18 x Jitter (percent) + 12.4. For perceptually normal voices, the DSI is +5, while for highly dysphonic voices, it is –5. The lower the patient's index, the lower the patient's vocal output. Acoustic Voice Quality Index (AVQI) (Hema et al., 2009; Maryn et al., 2010) combines six parameters using the equation:-

AVQI = 2.571*(3.295 - 0.111*CPPs - 0.073* HNR - 0.213*SL +2.789*Sh dB - 0.032*Slope +0.077*Tilt).

AVQI can be calculated automatically using the script written in Praat for a concatenated sample of sustained vowel and continuous speech. The AVQI score ranges from 0 to 10, indicating average to a profoundly abnormal voice on a scale of severity. The MDVP available in the CSL gives the quantitative measurement of voice quality by calculating 33 parameters on a single vocalization. MDVP is remarkable in its ability to function through a wide variety of pathological voices reliably.

Hema et al. (2009) developed normative data of the MDVP parameters for Indian population. Sixty participants were considered, thirty male and thirty female within the age range of 18 to 25. Results showed significant difference only in perturbation measurements, when the estimated values were compared with built-in normative data of CSL. Vishali (2019) investigated whether the AVQI parameters differed with respect to age and gender in Tamil speakers. No significant difference was observed in AVQI with respect to age and gender.

2.1.2.2 Invasive methods

An invasive method means the introduction of instruments into the body or body cavity. Many instruments are used to visualize the laryngeal state, such as rigid endoscopy, flexible endoscopy, video-stroboscope, video kymography, ultra-high-speed photography, ultrasound glottography and photoglottography. The use of instrumental analysis provides direct access to structures in action and supplements the acoustic and perceptual findings.

2.2 Telepractice

Telepractice is defined as "the application of telecommunications technology to provide remote professional services by connecting clinician to client or clinician to clinician for assessment, intervention and consultation" (American Speech-Language-Hearing Association [ASHA], 2005,)

There are three modes of telepractice namely synchronous, asynchronous, and hybrid: -

- Synchronous (interactive client) - Services with two way audio and video interactions.
- Asynchronous (store-and-forward) - Images or data are collected and distributed (i.e., saved and forwarded) for skilled viewing or understanding.
- Hybrid - Telepractice applications that involve combinations of synchronous, asynchronous and in-person resources.

Before starting the telepractice, the appropriate video conference application such as Zoom, Google-meet, WebEx, etc. need to be selected. The selected application should include features such as webcam sharing, screen sharing, interactive features, etc.

2.2.1 International Guidelines for Telepractice

ASHA suggests using Health Insurance Portability and Accountability Act (HIPAA)-friendly applications such as Go To Meeting, WebEx, etc., and the HIPAA-compliant version of Zoom. Applications such as Facetime, Skype, and Facebook Messenger are not HIPAA-compliant.

HIPAA compliant application should have the features given below (Fetzer & West, 2008)

- Capacity to maintain privacy while sending the emails and documents during screen sharing.
- Capacity to choose the appropriate virtual activities like PBS Kids, ABCya, Starfall, and TeacherspayTeachers (TpT) or Boom Cards, and also to create digital resources for the children and adults to attend the session.
- Facility to make virtual sessions engaging with the child.
- Shall follow the ASHA Code of Ethics (ASHA, 2016a).

2.2.2 Global Scenario of Telepractice

Fong et al. (2021) conducted a study on the implementation of telepractice in Hongkong during covid -19. Totally 135 speech language pathologists participated in the study which investigated the perception, training and knowledge of the telepractice. 83% of the SLP reported that its suitability depends on the type of condition of the patient. Lack of training about the technology usage in telepractice was found to be a major obstacle.

2.2.3 Telepractice in India

Mohan et al. (2017) conducted a study on telepractice in India. 205 members, including speech-language pathologists and audiologists, participated in the study. The participants were providing services in different settings such as clients' homes, elementary school, pre-elementary school, etc., for different age groups varying from infants to older adults. The services delivered include screening, assessment and management follow-up, and the professional consultancy for different communication disorders. 64% of telepractitioners reported client satisfaction with follow-up/monitoring and increased access to resources and 60% of clients were pleased with the regularity of telepractice services.

Yashaswini (2018) conducted a small survey regarding the telepractices in India, in two phases. In the first phase, questionnaires were mailed to all the 17 speech-language pathologists who were tele-practising in Mysore and Bangalore. Majority of the respondents (53.3 percent-86.6 percent) described technical problems as obstacles to telepractice. Furthermore, the service delivery model is different in terms of caregiver orders, documentation, face validity, acceptance, and caregiver accountability.

Gupta et al. (2021) reported that, for implementing telepractice, speech-language pathologists in India were using various software and applications (WebEx, TheraKonnect, VSee, FaceTime, Skype, Google Hangout).

Aggarwal et al. (2020) conducted an online survey in India during the covid-19 using the google form in which 84 SLPs between the ages of 21 and 53 years involved in telepractice participated. Percentage of SLPs using telepractice for treatment of child language disorder, fluency disorder, voice disorders, neurogenic language disorder and neurogenic speech disorders are 75.0%, 45.2%, 31.0%, 29.8%, and 25.0% respectively. The above studies indicate that telepractice would be a viable option for SLPs in India in the coming days.

2.3 Telepractice in voice disorders

According to ASHA, telecommunications is an effective mode of services for speech therapists during pandemic (Weidner & Lowman, 2020). Many studies have reported success stories of tele-rehabilitation of persons with voice disorders. The effectiveness of tele-rehabilitation has been documented in these studies by comparing the pre and post therapy voice parameters (acoustic and perceptual).

Rangarathnam et al. (2015) compared telepractice mode treatment of the voice disorder with conventional face to face method. Participants were divided into two groups (face to face and telepractice), each with seven participants. Pre and post treatment analysis was done using auditory–perceptual, acoustic, aerodynamic, and quality-of-life measures. No significant difference was observed between the face to face and telepractice groups.

Tindall et al. (2008) compared the traditional delivery and telepractice for persons with Parkinson's disease. Results showed no significant difference between the pre-test and a significant difference between pre and post-test. Similar improvements were reported in the face-to-face and telepractice modes.

Mashima et al. (2003) compared the conventional method and telepractice mode of voice therapy. Before therapy, baseline data were collected from all participants. Therapy was given through a real-time audio-video monitoring system. The vocal rehabilitation techniques (Boone, 1982; Boone & McFarlane, 2000) showed no differences in outcome measures between the face to face and online mode.

Majority of the studies reviewed above indicate the efficacy of telepractice mode in treatment of persons with voice disorders. The studies also showed no significant difference between the conventional and telepractice modes in terms of outcome measures. However, no studies were reported on tele-assessment of voice disorders.

2.4 Recording of samples through mobile phone for Assessment of voice

Smartphone are the most easily available devices for recording of voice samples. Curtis et al. (2019) showed that Smartphone can be used for recording the voice samples for assessment and management of speech disorders. Lin et al. (2012) measured and compared the parameters through analysis of voice and speech recorded through the iPhone and Personal Computer Memory Card International Association (PCMCIA) laptop recording system. They concluded that recording using the iPhone is suitable for acoustic analysis for assessment of speech & voice. Vogel et al. (2014) compared the recordings of modern devices such as Smartphone, landlines, laptops, and hard disc recorders. Results showed that none of the methods of acquisition for capturing voice and speech could provide statistically equivalent values compared to the benchmark system.

Lee et al. (2018) compared the recording of 180 individuals. The vowel /a/ samples were recorded using CSL-Model 4150B and android Smartphone (Samsung Galaxy Note 5, SM-N920) through the smart recorder app. No significant differences were observed between the devices and the study concluded that Smartphone can be used as a screening tool for voice disorders.

Grillo (2017) conducted a preliminary study on the VoiceEvalU8 Smartphone application. A total of twenty-one participants participated in the Global Voice Prevention Model (GVPM). Acoustic analyses were done by extracting the AVQI parameters. Findings revealed that F0 and jitter percent show improvements in voice from pre- to post-GVPM, in-person and telepractice settings.

Grillo et al. (2016) recorded voice samples using different Smartphone (iphone 5s, iPhone 6s, Samsung Galaxy S5) and a microphone mounted on the head. Ten participants were asked to phonate /a/ and small phrases at a comfortable pitch and loudness. The findings revealed that there was no significant within-subject heterogeneity between devices and applications. The study concluded that Smartphone are suitable for recording voice samples for regular voice measures that reflect the effects of vocal loading within individuals.

Barche et al. (2020) reported effectiveness of Smartphone in clinical assessment and management. The population tested included those with normal and disordered voice. Results indicated that the Smartphone can be used as instruments for voice assessment outside the clinic.

Petrizzo and Popolo (2020) conducted acoustic analysis on voice samples recorded using Smartphone as well as ipads. Results of analysis of voice parameters suggested that Smartphone and mobile applications can be useful instruments in voice assessment outside the laboratory.

Grillo et al. (2016) measured the acoustic parameters using different Smartphone such as iphone 5 and 6s and Samsung Galaxy S5 with head-mounted condenser microphones. The voice samples were analyzed using software programs such as MDVP, Praat and analysis of dysphonia in speech and voice (ADVS). The finding indicates no significant difference across the devices and concludes that Smartphone can be used to record the voice for estimation of acoustic parameters.

Carson and Ryalls (2018) highlighted the benefits of using Smartphone for an augmentative procedure through acoustic analysis, to assist perceptual evaluation of voice. The major advantages listed are: - they are non-invasive, can be implemented with freeware such as Praat or low-cost phone/ tablet applications such as OperaVOX (On Person Rapid Voice examiner). Findings suggest that the acoustic analysis of voice samples recorded through Smartphone supplements the perceptual evaluation.

Majority of the above studies suggest that Smartphone can be used as devices to record voice samples for voice analysis. Most of the studies conclude that the parameters of voice recorded through Smartphone do not show significant difference when compared with parameters of voice recorded through conventional systems such as digital recorders.

2.5 Recording of samples through online apps for Assessment of voice

McGill and Fiddler (2021) reported that the Zoom platform is widely utilized for telepractice. Zoom was selected based on HIPAA compliance, usability and pricing of software licenses.

Weerathunge et al. (2021) measured the accuracy of acoustic parameters when transmitted through videoconference platforms. Twentynine individuals with dysphonia within the age range of (19-82) participated in the study. Six popular videoconference platforms were considered. After the recording of audio samples, the acoustic analysis was carried out, using Praat acoustic software. The parameters analysed were mean F0, F0 variation (standard deviation and range), SPL variation (standard deviation and range), HNR, L/H ratio, and CPPS. All acoustic parameters except MF0 were affected. They concluded that Microsoft Teams and Zoom had the least impact on acoustic measures. Overall, measures of F0 were least affected by telepractice transmission. In this study, participant’s voice sample was recorded and was reproduced through an external speaker while transmitting through the computer during videoconferencing.

Many researchers investigated the voice samples recorded through different media, such as digital recorder, Smartphone, and laptops to find their suitability to record voice samples for acoustic analysis. However, none of these researchers have investigated whether the voice samples recorded online can be used to assess voice disorders. In the present pandemic situation, it is difficult to record the voice samples directly at the clinic. It has been already established by the previous researchers that recordings with Smartphone can be used for acoustic analysis. Hence, there is a need to investigate whether the voice samples recorded online are suitable for assessing voice disorders. The present study is an investigation in this direction.

Chapter 3

METHOD

3.1 Research design

The present study uses standard group comparison to compare the normal and the dysphonia group to achieve the objectives.

3.2 Participants

A total of 48 literate Kannada speaking adults, within the age range of 20-50 years, were recruited for the study, in two groups. Group I included 31 normal participants, and group II included 17 participants with Dysphonia.

Inclusion and exclusion criteria for selection of group I and II

- Participants having mobile phones within the price range of Rs. 10,000 - 20,000/- with video conferencing apps (App1 and App2) installed in their mobile phones.
- Participants with normal/ corrected vision.

Additional inclusion and exclusion criteria for selection of group I

- Participants with normal speech, voice, language, hearing, intact cognition, with no signs and symptoms of upper respiratory tract infections at the time of recording. No neurological, social, emotional, cognitive or psychiatric disturbances.
- Individuals who have a history of voice disorders, exposure to cigarette/ consumption of alcohol/chemical fumes, or any trauma/ accident/ surgery to the laryngeal system or long-term drug history exposure regardless of medical conditions such as diabetes or hypertension were excluded.

Additional inclusion and exclusion criteria for selection of group II

- Participants who were diagnosed with Dysphonia and without any co-morbid condition.

3.3 Procedure

The researcher virtually connects to the participant (setting at their residence) through App1 and App2 and give clear instructions before the online recording. Oral consent was obtained from the participants and the purpose and procedure of the study was explained. Participants were instructed to be seated comfortably in a noise-free room, keeping the phone at a distance of 10 cm from the mouth. They were requested to perform the following two tasks: - 1. To phonate of /a/ for 5-6 seconds at a pitch and loudness, comfortable to them and 2. To read a standardized passage (Appendix-A) in Kannada (Bengaluru passage). Six lines of text in the passage were shown at a time on the screen. Font size and line spacing of the text in the passage were finalized after ascertaining the readability by three Speech Language Pathologists, before the start of the study. The phonation and the read passage were recorded by the researcher through App1 and App2. The stability in internet connectivity was ensured with speed test by Ookla, (https://www.speedtest.net/) and the data collection was done only when the internet speed was between 60-70 megabytes per second (Mbps). After the online session, the participants were instructed to record the phonation and passage through their mobile phone and send it to the researcher through mail. All the samples were stored in laptop or PC for further analysis.

3.4 Instrumentation and Analysis

13 acoustic parameters were analysed using the MDVP of the CSL (model 4305) for the phonation sample of vowel /a/. For AVQI, both phonation sample /a/ and standardized reading passage were analysed using AVQI script in PRAAT software.

3.4.1 Analysis using MDVP

MDVP evaluates 33 parameters, which can be categorized under eight main types: such as fundamental frequency, frequency and amplitude perturbation (short and long term), voice break related, sub-harmonic related, voice irregularity related, noise related, and tremor related measures. In the current study, only the following 13 parameters were considered to compare voice characteristics between the groups:

1. Average fundamental frequency (MF0) – Average value of all obtained period to period F0 values.
2. Standard deviation of fundamental frequency (STD) – Variation of F0 within the analyzed voice sample.
3. F0 tremor frequency (Fftr) – It is the frequency of most intense low frequency modulated signal.
4. Amplitude tremor frequency (Fatr) – It is the frequency of most intense low frequency amplitude modulated signal.
5. Jitter percentage (Jitt %) – Relative evaluation of period to period (brief term) variability of pitch within the analyzed speech sample.
6. Relative average perturbation (RAP) - Relative evaluation of period to period (concise term) variability of pitch within the analyzed speech sample with smoothening factor of 3 periods.

7. Fundamental frequency variation (vF0) – Relative standard deviation of the period to period calculated fundamental frequency.
8. Shimmer percentage (Shim %) - Relative evaluation of period to period (very short term) variability of peak-to-peak amplitude within the analyzed voice sample.
9. Amplitude perturbation quotient (APQ) - Relative evaluation of period to period (very short term) variability of peak-to-peak amplitude for 11 cycles.
10. Noise to harmonic ratio (NHR) – Average ratio of harmonic energy in the range of 1500-4500 Hz to harmonic energy in the range of 70-4500 Hz.
11. Voice turbulence index (VTI) – It is the ratio of high frequency in-harmonic spectral energy in the range of 70-1600 Hz to spectral harmonic energy in the range of 1600-4500 Hz.
12. F0 tremor intensity index (FTRI) – Average ratio of frequency magnitude of the lowest frequency modulation to total frequency modulation.
13. Amplitude tremor intensity index (ATRI) – Average ratio of the amplitude of the most intense low amplitude modulating component to the analysed voice sample's total amplitude.

3.4.2 Analysis using AVQI

The phonation and reading samples were subjected to AVQI analysis using AVQI script in PRAAT software (v.6.0) (Boersma & Weenink, 2015) on Lenovo G-50 Laptop. The AVQI analysis includes estimation of the following six parameters:-

1. Cepstral peak prominence (CPPS) – The distance between the first rahmonic's peak and the point with equal quefrency on the regression line through the smoothed cepstrum

2. Harmonic to noise ratio (HNR) – The base-10-logarithm of the ratio between the periodic energy and the noise energy, multiplied by 10.
3. Shimmer local (ShimLocal) – The absolute mean difference between the amplitudes of successive periods, divided by the average amplitude.
4. Shimmer dB (SLdB) – The base-10-logarithm of the difference between the amplitudes of successive periods, multiplied by 20.
5. Slope of LTAS – The difference between the energy in 0–1,000 Hz and the energy in 1,000–10,000 Hz of the long-term average spectrum.
6. Tilt of LTAS The difference between the energy in 0–1,000 Hz and the energy in 1,000–10,000 Hz of the trend line through the long-term average spectrum.

3.5 Ethical consideration

The aim and objective of the study were told to all the participants and the method and purpose of the study was explained. Their safety and confidentiality were ensured, and oral consent was taken before conducting the study.

3.6 Data recording

The recording was carried out in a noise-free room, and all the participants were instructed to sit comfortably and asked to join the researcher's meeting using App1 and App2 through mobile phones kept at a distance of 10 cm away from the mouth. The researcher recorded the meeting through online mode. For the 1st task, all the participants were asked to take a long breath and phonate /a/ vowel at a comfortable pitch and loudness level for 5-6 seconds. 3 trials of phonating the vowel /a/ was recorded and the best one selected by the researcher was considered for the analysis. For the 2nd task, participant was asked to read the standard Kannada passage (Bengaluru passage) at a comfortable loudness level. Both the samples were stored in the laptop. After online recording, the text of the passage was sent to the participant through what's app. Participant was informed to record the two tasks through the voice recorder app of their mobile phones and then mail the recorded samples to the researcher. All samples were converted to .wav format, for further analysis.

3.7 Data analysis

Recorded samples of phonation and read text were subjected to the AVQI and MDVP analysis. For the AVQI analysis, 3 seconds slice was taken from the recorded phonations and named as 'SV' (sustained vowel). The read speech samples (89 syllables) were named as 'CS' (continuous speech). Using the Praat software (v.6.1.51), the analysis was done using the AVQI script developed by Maryn et al. (2010). The AVQI scores were obtained on a scale of 0-10. Phonation samples were also subjected to MDVP analysis, and 13 acoustic parameters were extracted.

3.8 Statistical Analysis

The obtained MDVP and AVQI parameters were tabulated and subjected to statistical analysis in Statistical Package for the Social Sciences (SPSS) software (Version 21.0). Descriptive statistics was carried out to calculate the mean, median, and standard deviation for both the groups. Shapiro Wilk test and Kolmogorov-Smirnov test was done to test the normality. As the obtained data was non-normalized, a non-parametric analysis was performed. That is, the Friedman's test was done to compare between Smartphone, App1 and App2.

Chapter 4

RESULTS

The objectives of the present study were to compare the acoustic parameters (AVQI and MDVP) of the voice samples of individuals with normal voice as well as individuals with dysphonia recorded online through two video conferencing platforms (apps), with the voice samples recorded directly through mobile phone. The study also compared the acoustic parameters of those voice samples recorded online through the two apps.

The obtained MDVP and AVQI parameters were tabulated and subjected to statistical analysis. The following analyses were carried out to achieve the objectives of the study:

a) Estimating the mean, median and standard deviation (SD) of the MDVP and AVQI parameters of the voice samples recorded with Smartphone, video conferencing application1(App1) and video conferencing application2 (App2).

b) Test of normality using the Shapiro Wilk test and Kolmogorov-Smirnov test.

c) Comparison of the acoustic parameters of normal voice samples recorded through Smartphone, App1 and App2, using Friedman's test.

d) Comparison of the acoustic parameters of dysphonic voice samples recorded through Smartphone, App1 and App2, using Friedman's test.

e) Comparison of the acoustic parameters of normal voice samples recorded through App1 and App2, using Friedman's test.

f) Comparison of the acoustic parameters of dysphonic voice samples recorded through App1 and App2, using Friedman's test.

4.1 Characteristics of participants

In the present study, we considered two groups of participants. Group I included 31 normal participants with mean age range of 23.03 years (SD =1.741), and group II consisted of 17 persons with Dysphonia with mean age of 37.64 years (SD= 9.816). Details of the group II participants are shown in Table 4.1.

Table 4.1

Details of Group II participants

		Dysphonia		
Participants	Age/Gender	Degree	Severity	Stroboscopic examination
S1	23/M	Moderate	Hoarse voice	Right vocal fold polyp
S2	40/M	Moderate	Hoarse voice	-
S3	50/M	Moderate	Hoarse voice	-
S4	51/M	Severe	Breathy voice	Left vocal fold paralysis
S5	45/M	Moderate	Hoarse voice	-
S6	26/M	Mild	Hoarse voice	-
S7	35/M	Moderate	Hoarse voice	-
S8	41/M	Moderately severe	Hoarse voice	Sulcus vocalis
S9	23/M	Moderate	Hoarse voice	-
S10	30/M	Mild	Hoarse voice	Sulcus vocalis
S11	44/M	Moderate	Hoarse voice	-
S12	40/F	Mild	Hoarse voice	Restricted right vocal fold movement
S13	30/F	Moderate	Breathy voice	Glottic chink
S14	25/F	Moderate	Hoarse voice	-
S15	47/F	Moderate	Hoarse voice	Laryngitis
S16	50/F	Mild	Hoarse voice	-
S17	40/F	Mild	Breathy voice	Restricted right vocal fold movement

Mean age range of participant in Dysphonia group = 37.64 ±9.816

4.2. Comparison of the acoustic parameters of normal voice samples recorded through Smartphone, App1 and App2

The mean, median and standard deviation (SD) values are obtained using the descriptive statistics for the acoustic parameters of voice samples of group I recorded through Smartphone, App1, and App2. These values are tabulated in Table 4.2.

Table 4.2

Mean, SD, Median values of acoustic parameters of voice sample recordings through Smartphone, App1, and App2 in Normal group I

	Smartphone			App1			App2		
Parameters	**Mean**	**SD**	**Median**	**Mean**	**SD**	**Median**	**Mean**	**SD**	**Median**
MF0(Hz)	180.2	51.55	192.5	182.8	49.97	195.2	182.7	53.17	198.4
STD(Hz)	1.944	1.042	1.787	2.529	1.498	1.899	2.357	1.394	2.071
Fftr (Hz)	2.919	2.097	2.878	2.715	1.498	3.252	3.046	2.926	3.125
Fatr (Hz)	2.373	2.163	2.721	2.658	2.369	3.226	2.997	2.418	2.837
Jitt (%)	0.694	0.388	0.573	0.700	0.452	0.566	0.819	0.556	0.635
RAP (%)	0.432	0.229	0.377	0.408	0.273	0.331	0.481	0.333	0.379
vF0 (%)	1.058	0.394	0.966	1.265	0.603	1.096	1.453	1.087	1.057
Shim (%)	3.596	0.939	3.311	4.736	0.948	4.642	6.061	2.753	5.050
APQ (%)	2.801	0.838	2.629	4.311	1.274	4.144	5.342	2.598	4.343
NHR	0.128	0.021	0.132	0.133	0.013	0.134	0.145	0.024	0.142
VTI	0.048	0.025	0.043	0.042	0.020	0.040	0.053	0.026	0.051
Ftri (%)	0.224	0.200	0.186	0.261	0.159	0.246	0.290	0.397	0.228
Atri (%)	2.973	3.226	2.613	4.171	4.900	2.438	5.125	4.607	4.594
CPPS	12.48	1.147	12.20	12.25	0.988	12.32	11.99	0.915	12.21
HNR	14.53	2.120	14.38	14.47	2.473	13.98	14.22	2.553	14.49
ShimLocal	9.621	1.730	9.470	10.19	2.105	10.32	9.907	2.088	10.05
Shimmer dB	0.925	0.146	0.940	0.990	0.162	0.980	0.972	0.170	1.010
SlopeLTAS	-22.61	3.889	-22.24	-20.80	5.073	-21.07	-20.17	4.569	-19.91
TiltLTAS	-11.06	1.853	-10.73	-13.23	1.087	-13.46	-12.96	1.522	-13.33
AVQI	4.574	0.672	4.640	4.272	0.529	4.220	4.412	0.576	4.440

Abbreviations: MFO, Average Fundamental Frequency; STD, Standard Deviation of FO; Fftr, FO Tremor Frequency: Fatr, Amplitude Tremor Frequency; Jitt (%), Jitter Percent; RAP, Relative Average Perturbation; vFO, Fundamental Frequency Variation; Shim (%), Shimmer Percent; APQ, Amplitude Perturbation Quotient; NHR, Noise to Harmonic Ratio; VTI, Voice Turbulence Index; FTRI, FO Tremor Intensity Index; ATRI, Amplitude Tremor Intensity Index. CPPS, Cepstral Peak prominence; HNR, Harmonic to noise ratio; AVQI, Acoustic Voice Quality Index, App1 – video conference application 1, App2 – video conference application 2

The data containing values of acoustic parameters (AVQI and MDVP) were tested for normality using the Shapiro Wilk test and Kolmogorov-Smirnov test. Results showed that the data were not normally distributed ($p<0.05$) for the parameters such as MF0, STD, Fftr, Fatr, Jitt, RAP, vF0, Shim, APQ, NHR, VTI, FTRI, ATRI, CPPS, HNR, ShimLocal, ShimdB, SlopeLTAS, TiltLTAS, and AVQI. Hence a non-parametric test was performed to compare the parameters across the Smartphone, App1, and App2.

To compare the acoustic parameters (MDVP and AVQI) of normal voice sample recorded in Smartphone, App1, and App2 Friedman's two-way analysis of variance was carried out. Results, as shown in the Table 4.3, revealed that there is no significant difference in any of the parameters other than shimmer ($\chi^2(2) = 25.74$, $p<0.05$), APQ ($\chi^2 (2) = 40.25$, $p<0.05$), NHR ($\chi^2 (2) = 12.64$, $p<0.05$), CPPS ($\chi^2 (2) = 7.16$, $p<0.05$) and TiltLTAS ($\chi^2 (2) = 19.93$, $p<0.05$).

Table 4.3

Acoustic parameters with significant difference when voice sample recordings through Smartphone, App1 and App2 in group I were compared using Friedman's Test.

		Group I		
SI. No	**Parameters**		**|z|**	**p value**
1	Shimmer	Smartphone-App1	0.871	0.002
		Smartphone-App2	1.258	0.000
2	APQ	Smartphone-App1	1.161	0.000
		Smartphone-App2	1.548	0.000
3	NHR	App1-App2	0.839	0.003
4	CPPS	Smartphone-App2	0.677	0.023
5	TiltLTAS	Smartphone-App1	1.065	0.000
		Smartphone-App2	0.871	0.002

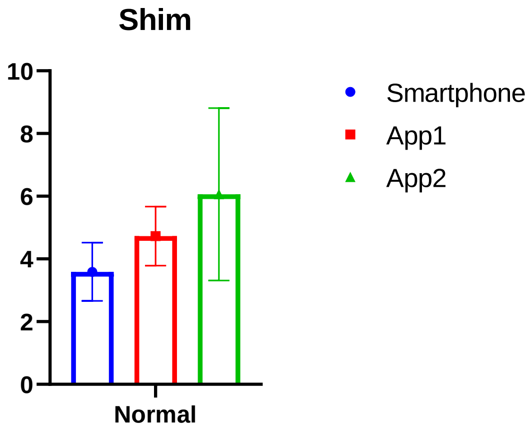
Abbreviations: Shim (%), Shimmer Percent; APQ, Amplitude Perturbation Quotient; NHR, Noise to Harmonic Ratio; CPPS, Cepstral Peak prominence; App1 – video conference application 1, App2 – video conference application 2.

4.2.1 Shimmer

Friedman's test was done to find out whether there is a significant difference in the value of shimmer in the voice samples recorded through Smartphone, App1 and App2 ($\chi^2(2)$ = 25.74, $p<0.05$). As there was a significant difference, pair wise comparison was done between Smartphone and App1 ($|z|=0.871$, $p<0.05$), Smartphone and App2 ($|z|=1.258$, $p<0.05$) and between the App1 and App2 ($|z|=0.387$, $p>0.05$). As indicated in Figure 4.1, it is observed that shimmer value is higher in App2 than App1 and Smartphone.

Figure 4.1

Mean and SD value of Shimmer of voice sample recordings through Smartphone, App1, and App2 in Group I.

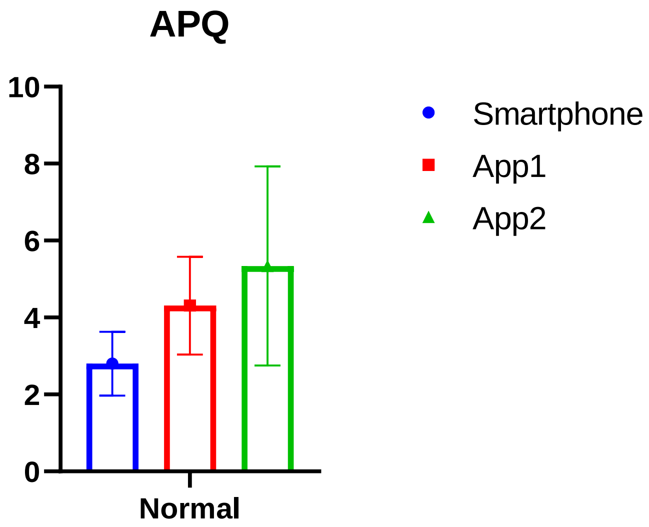


4.2.2 Amplitude perturbation quotient

Friedman's test was done to find out whether there is a significant difference in the value of APQ in the voice samples recorded through Smartphone, App1 and App2 ($\chi^2(2) = 40.25$, $p<0.05$). As there was a significant difference, pair wise comparison was done between Smartphone and App1 ($|z|=1.161$, $p<0.05$), Smartphone and App2 ($|z|=1.548$, $p<0.05$) and between the App1 and App2 ($|z|=0.387$, $p>0.05$). As indicated in Figure 4.2, it is observed that APQ value is higher in App2 than App1and Smartphone.

Figure 4.2

Mean and SD, of APQ of voice sample recordings through Smartphone, App1, and App2 in Group I.

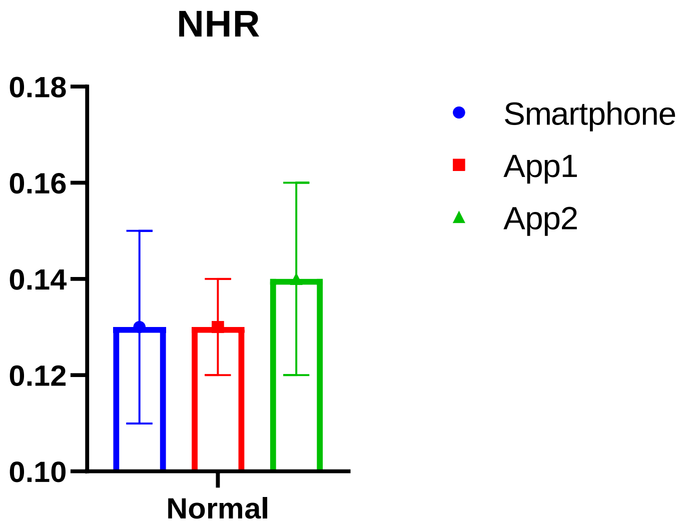


4.2.3. Noise to harmonic ratio

Friedman's test was done to find out whether there is a significant difference in the value of NHR in the voice samples recorded through Smartphone, App1 and App2 ($\chi^2(2) = 12.645$, $p<0.05$). As there was a significant difference, pairwise comparison was done between Smartphone and App1 ($|z|=0.129$, $p>0.05$), Smartphone and App2 ($|z|=0.710$, $p>0.05$) and between the App1 and App2 ($|z|=0.839$, $p<0.05$). As indicated in Figure 4.3, it is observed that NHR value is higher in App2 than App1and Smartphone.

Figure 4.3

Mean and SD, of NHR of voice sample recordings through Smartphone, App1, and App2 in Group I.

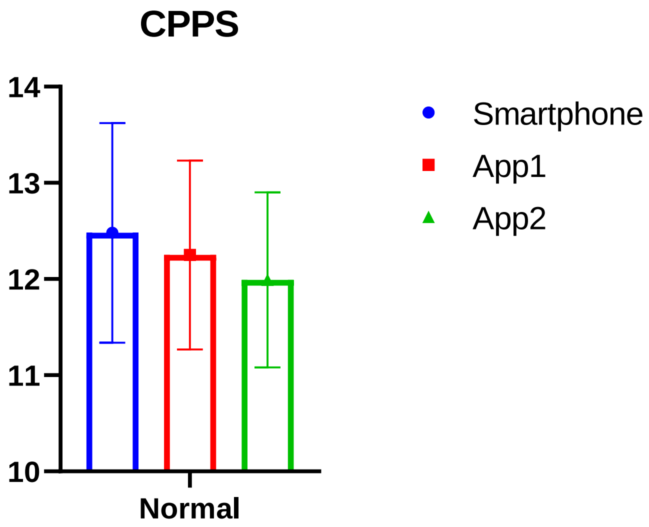


4.2.4 Cepstral peak prominence

Friedman's test was done to find out whether there is a significant difference in the value of CPPS in the voice samples recorded through Smartphone, App1 and App2 ($\chi^2(2) = 7.161$, $p<0.05$). As there was a significant difference, pairwise comparison was done between Smartphone and App1 ($|z|=0.387$, $p>0.05$), Smartphone and App2 ($|z|=0.677$, $p>0.05$) and between the App1 and App2 ($|z|=0.387$, $p>0.05$). As indicated in Figure 4.4, it is observed that CPPS value is higher in Smartphone than App1 and App2.

Figure 4.4

Mean and SD, of CPPS of voice sample recordings through Smartphone, App1, and App2 in Group I.

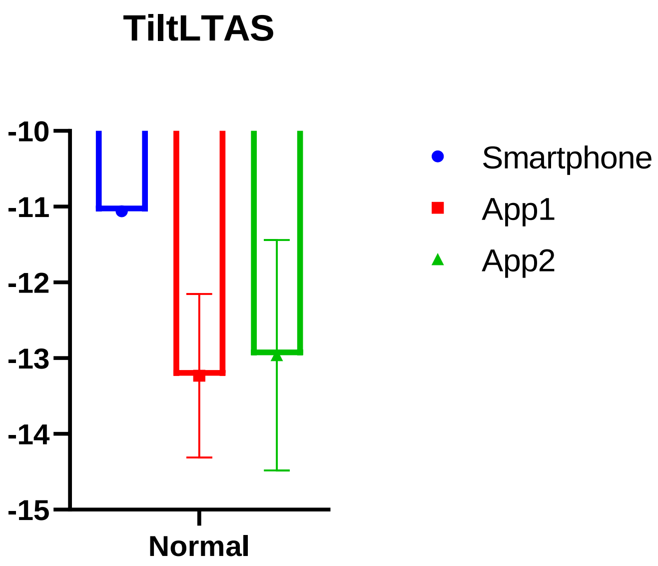


4.2.5 Tilt of Long Term Average Spectrum

Friedman's test was done to find out whether there is a significant difference in the value of Tilt of LTAS in the voice samples recorded through Smartphone, App1 and App2 ($\chi^2(2)$ = 19.935, p<0.05). As there was a significant difference, pair wise comparison was done between Smartphone and App1 (|z|=1.065, p<0.05), Smartphone and App2 (|z|=0.871, p<0.05) and between the App1 and App2 (|z|=0.194, p>0.05). As indicated in Figure 4.5, it is observed that Tilt of LTAS value is higher in App1 than App2 and Smartphone.

Figure 4.5

Mean and SD, of Tilt LTAS of voice sample recordings through Smartphone, App1, and App2 in Group I.



4.3 Comparison of the acoustic parameters of dysphonic voice samples recorded through Smartphone, App1 and App2

The mean, median and standard deviation of the acoustic parameters in Smartphone, App1, and App2 of the voice samples of persons with dysphonia (Group II) are tabulated in Table 4.4.

Table 4.4

Mean, SD, Median values *of acoustic parameters of voice sample recordings through Smartphone, App1, App2 in Group II.*

	Smartphone			**App1**			**App2**		
Parameters	**Mean**	**SD**	**Median**	**Mean**	**SD**	**Median**	**Mean**	**SD**	**Median**
MF0 (Hz)	200.0	66.40	189.3	199.8	59.99	177.4	192.3	52.90	171.2
STD (Hz)	11.44	19.27	4.541	12.43	12.91	7.925	8.786	11.73	4.283
Fftr (Hz)	3.218	2.590	3.175	4.253	2.523	3.279	3.863	2.528	3.252
Fatr (Hz)	3.108	2.492	2.500	2.597	2.575	2.523	3.091	2.888	3.030
Jitt (%)	1.893	1.700	1.321	1.711	1.292	1.071	2.115	1.161	1.788
RAP (%)	1.156	1.080	0.798	0.998	0.745	0.650	1.250	0.682	1.061
vF0 (%)	5.206	7.901	2.158	4.169	5.506	1.991	6.928	7.981	3.914
Shim (%)	7.170	3.932	6.705	7.559	3.795	6.267	8.851	2.656	7.790
APQ (%)	5.402	2.227	5.162	6.382	3.115	5.565	7.936	2.835	7.218
NHR	0.198	0.126	0.152	0.157	0.057	0.141	0.174	0.053	0.158
VTI	0.047	0.024	0.047	0.048	0.022	0.045	0.044	0.019	0.045
Ftri (%)	0.890	2.369	0.315	0.593	0.488	0.360	0.636	0.654	0.559
Atri (%)	4.694	4.747	3.885	6.033	6.882	4.181	5.935	7.755	1.805
CPPS	11.19	1.552	11.07	11.52	1.202	11.82	11.39	1.449	11.51
HNR	14.37	2.187	14.07	13.83	2.427	13.94	13.98	2.964	13.02
ShimLocal	10.01	1.865	9.870	10.01	2.971	10.57	11.03	2.309	11.80
ShimmerdB	0.980	0.156	1.020	1.038	0.131	1.010	1.045	0.188	1.110
SlopeLTAS	-21.77	3.182	-21.08	-20.88	3.767	-21.10	-20.40	3.692	-20.44
TiltLTAS	-11.98	1.691	-12.51	-13.02	1.515	-13.39	-13.07	1.388	-13.66
AVQI	4.747	0.775	4.910	4.600	0.681	4.730	4.510	0.728	4.680

Abbreviations: MFO, Average Fundamental Frequency; STD, Standard Deviation of FO; Fftr, FO Tremor Frequency: Fatr, Amplitude Tremor Frequency; Jitt (%), Jitter Percent; RAP, Relative Average Perturbation; vFO, Fundamental Frequency Variation; Shim (%), Shimmer Percent; APQ, Amplitude Perturbation Quotient; NHR, Noise to Harmonic Ratio; VTI, Voice Turbulence Index; FTRI, FO Tremor Intensity Index; ATRI, Amplitude Tremor Intensity Index. CPPS, Cepstral Peak prominence; HNR, Harmonic to noise ratio; AVQI, Acoustic Voice Quality Index, App1 – video conference application 1, App2 – video conference application 2.

The data containing values of acoustic parameters (AVQI and MDVP) were tested for normality using the Shapiro Wilk test and Kolmogorov-Smirnov test. Distribution of the data were not normal ($p<0.05$) for the parameters such as MF0, STD, Fftr, Fatr, Jitt, RAP, vF0, Shim, APQ, NHR, VTI, FTRI, ATRI, CPPS, HNR, ShimLocal, ShimdB, SlopeLTAS, TiltLTAS, and AVQI. Hence a non-parametric test was performed to compare the parameters across the Smartphone, App1, and App2.

To compare the acoustic parameters (MDVP and AVQI) of Dysphonia voice sample recorded in Smartphone, App1, and App2, Friedman's two-way analysis of variance was carried out. Results, as shown in the Table 4.5, revealed that there is no significant difference in any of the parameters.

Table 4.5.

Comparison of acoustic parameters of samples recorded through Smartphone and App1 & App2 values in Group II.

	Dysphonia	
Parameters	**χ^2 value**	***p* value**
MF0 (Hz)	0.567	0.753
STD (Hz)	3.077	0.215
Fftr (Hz)	2.133	0.344
Fatr (Hz)	0.327	0.849
Jitt (%)	3.656	0.161
RAP (%)	3.469	0.177
vF0 (%)	3.323	0.190
Shim (%)	2.338	0.311
APQ (%)	2.469	0.291
NHR	3.469	0.177
VTI	0.406	0.816
Ftri (%)	1.477	0.478
Atri (%)	0.250	0.882
CPPS	1.529	0.465
HNR	2.716	0.257
ShimLocal	4.353	0.113
ShimmerdB	4.455	0.108
SlopeLTAS	4.149	0.126
TiltLTAS	3.254	0.197
AVQI	2.000	0.368

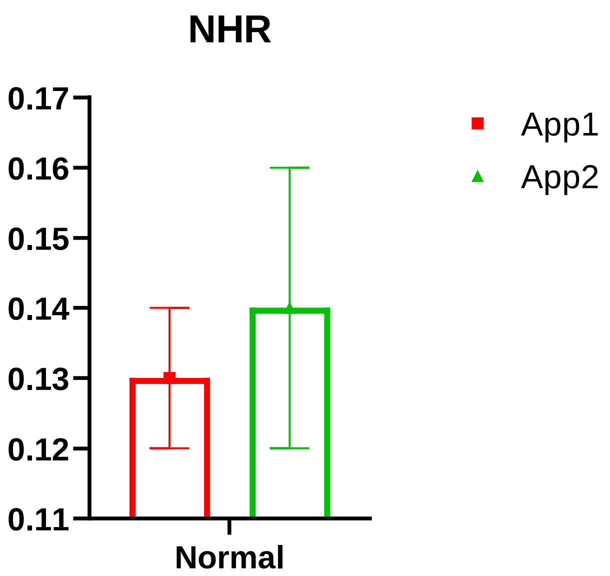
Abbreviations: MFO, Average Fundamental Frequency; STD, Standard Deviation of FO; Fftr, FO Tremor Frequency: Fatr, Amplitude Tremor Frequency; Jitt (%), Jitter Percent; RAP, Relative Average Perturbation; vF0, Fundamental Frequency Variation; Shim (%), Shimmer Percent; APQ, Amplitude Perturbation Quotient; NHR, Noise to Harmonic Ratio; VTI, Voice Turbulence Index; FTRI, F0 Tremor Intensity Index; ATRI, Amplitude Tremor Intensity Index. CPPS, Cepstral Peak prominence; HNR, Harmonic to noise ratio; AVQI, Acoustic Voice Quality Index, App1 – video conference application 1, App2 – video conference application 2.

4.4 Comparison of the acoustic parameters of normal voice samples recorded through App1 and App2

To compare the acoustic parameters (MDVP and AVQI) of normal voice sample recorded in App1, and App2, Friedman's two-way analysis of variance was carried out. Results, as revealed that there is no significant difference in any of the parameters other than NHR (χ^2 (2) = 12.64, p<0.05) Friedman's test was done to find out whether there is a significant difference in the value of NHR in the voice samples recorded through App1 and App2 (χ^2(2) = 12.645, p<0.05). As there was a significant difference, pairwise comparison was done between App1 and App2 (|z|=0.839, p<0.05). As indicated in Figure 4.6, it is observed that NHR value is higher in App2 than App1.

Figure 4.6

Mean and SD, of NHR of voice sample recordings through App1, and App2 in Group I.



4.5 Comparison of the acoustic parameters of dysphonic voice samples recorded through App1 and App2

To compare the acoustic parameters (MDVP and AVQI) of Dysphonia voice sample recorded through App1 and App2, Friedman's two-way analysis of variance was carried out. Results, as shown in the Table 4.6, revealed that none of the parameters were having significant difference in App1 and App2.

Table 4.6.

Comparison of acoustic parameters of samples recorded through App1 and App2 in Group II.

	Dysphonia	
Parameters	**χ^2 value**	***p* value**
MF0 (Hz)	0.567	0.753
STD (Hz)	3.077	0.215
Fftr (Hz)	2.133	0.344
Fatr (Hz)	0.327	0.849
Jitt (%)	3.656	0.161
RAP (%)	3.469	0.177
vF0 (%)	3.323	0.190
Shim (%)	2.338	0.311
APQ (%)	2.469	0.291
NHR	3.469	0.177
VTI	0.406	0.816
Ftri (%)	1.477	0.478
Atri (%)	0.250	0.882
CPPS	1.529	0.465
HNR	2.716	0.257
ShimLocal	4.353	0.113
ShimmerdB	4.455	0.108
SlopeLTAS	4.149	0.126
TiltLTAS	3.254	0.197
AVQI	2.000	0.368

Abbreviations: MFO, Average Fundamental Frequency; STD, Standard Deviation of FO; Fftr, FO Tremor Frequency: Fatr, Amplitude Tremor Frequency; Jitt (%), Jitter Percent; RAP, Relative Average Perturbation; vFO, Fundamental Frequency Variation; Shim (%), Shimmer Percent; APQ, Amplitude Perturbation Quotient; NHR, Noise to Harmonic Ratio; VTI, Voice Turbulence Index; FTRI, FO Tremor Intensity Index; ATRI, Amplitude Tremor Intensity Index. CPPS, Cepstral Peak prominence; HNR, Harmonic to noise ratio; AVQI, Acoustic Voice Quality Index, App1 – video conference application 1, App2 – video conference application 2.

Chapter 5

DISCUSSION

We attempted to answer the following questions in the present study:

1. Is there any difference in the acoustic parameters (AVQI & MDVP) of the voice samples of individuals with normal voice, recorded online through App1 and App2, with the voice samples recorded directly through mobile phone?
2. Is there any difference in the acoustic parameters (AVQI & MDVP) of the voice samples of individuals with Dysphonia, recorded online through App1 and App2, with the voice samples recorded directly through mobile phone?
3. Is there any difference in the acoustic parameters (AVQI & MDVP) of the voice samples of individuals with normal voice, recorded online through App1 with the samples recorded through App2?
4. Is there any difference in the acoustic parameters (AVQI & MDVP) of the voice samples of individuals with Dysphonia, recorded online through App1with the samples recorded through App2?
5. Are the voice samples recorded through App1 and App2 suitable for voice evaluation? Among the two Apps, which one is the most suitable?

5.1 Selection of participants and their characteristics

A total of 48 literate, Kannada speaking adults within the age range of 20-50 years were recruited for the study in two groups. Group I included 31 normal participants with a mean age range of 23.03 years (SD =1.741, Male=15 Female= 16), and group II consisted of 17 persons with Dysphonia with a mean age of 37.64 years (SD= 9.816, Male=11 and Female= 6). All the participants have mobile phones within the price range of Rs.10,000-20,000/- with App1 and App2 installed in their mobile phones. Participants with normal speech (including voice) and language, hearing, and intact cognition were included in group I. Participants in group II were individuals who were diagnosed with Dysphonia and without any co-morbid condition. Through the speed test by Ookla, (https://www.speedtest.net/), it was ensured that the online recording was done only when the internet speed was between 60-70 megabytes per second.

5.2 Difference between the acoustic parameters (AVQI & MDVP) of the voice samples of individuals with normal voice, recorded online through App1 and App2, and the voice samples recorded directly through mobile phone

When the acoustic parameters (MDVP and AVQI) of the voice samples of individuals with normal voice, recorded online through App1 and App2, were compared with the voice samples recorded directly through mobile phone, no significant difference was observed in any of the parameters other than shimmer, APQ, NHR, CPPS, TiltLTAS.

As the shimmer and APQ values increase, voice quality decreases (Muta et al., 1988). A significant difference was observed between the samples recorded through mobile phone and App1 and mobile phone and App2 for the parameters such as Shimmer and APQ. Previous studies (Winholtz & Titze, 1998) have reported that amplitude perturbation is more sensitive to instrumental or ambient noise compared to frequency-related measures. Shimmer values are adversely affected by background noise across devices (Carson et al., 2003; Lebacq et al., 2017). As these samples were recorded online through App1 and App2, it may be possible that the ambient noise must also have influenced the recording. Similar findings are supported by (Maryn et al., 2017; Weerathunge et al., 2021)

Noise to harmonic ratio (NHR) is the ratio of a periodic noise signal to periodic harmonic signal. Higher NHR values are indicative of higher noise component in the signal, which indicates abnormality in voice (Titze, 1995). A significant difference was observed in noise to harmonic ratio between the samples recorded through mobile phone and App2. In contrast, no significant difference was observed between recording of mobile phones and App1. Lebacq et al., 2017 opined that signal processing occurs in some mobile phones, affecting noise to the harmonic ratio when formant changes are introduced. In samples recorded with App2, these formant changes must have occurred which might have disturbed the NHR (Weerathunge et al., 2021).

Cepstral peak prominence measures the degree of harmonic organization or configuration in the spectrum. Decrease in CPPS show abnormality in voice (Hillenbrand et al., 1994; Hillenbrand & Houde, 1996). A significant difference was also observed in cepstral peak prominence between the samples recorded through mobile phone and App2, whereas no significant difference between the mobile phone and App1. CPPS was found to be consistently lower in App2 compared to mobile phones. Lower the value, the chance of abnormality in the voice. Higher the value of CPPS indicates better voice quality (Hillenbrand et al., 1994; Hillenbrand & Houde, 1996).

Tilt of LTAS is the ratio of energy in the low frequency and energy in the high-frequency. A steeper LTAS indicates that the difference between energies of low and high-frequency. Thus a greater spectral tilt reduces voice quality as reported by Löfqvist and Mandersson (1987). Significant differences were observed in the tilt of LTAS between the samples recorded through mobile phones and App1 as well as between the samples recorded through mobile phones and App2. The Tilt of LTAS parameter shows the highest relative random error that will reduce the mean relative across the mobile phones (Schaeffler et al., 2019). This supported our study because the mean values of Tilt of LTAS in samples recorded through App1 and App2 were high compared to mobile phones.

5.3 Difference between the acoustic parameters (AVQI & MDVP) of the voice samples of individuals with dysphonia, recorded online through App1 and App2, and the voice samples recorded directly through mobile phone

Lee et al., 2018; Lin et al., 2012; Vogel et al., 2014 and Yun, 2015 have analysed voice samples recorded directly through mobile phones and found that these recordings are suitable for voice analysis. The present study did not show any significant difference between the acoustic parameters (AVQI & MDVP) of the voice samples of individuals with dysphonia, recorded online through App1 and App2, and the voice samples recorded directly through mobile phone**.** Samples recorded through App1 & App2 can be used for voice analysis.

Figure 5.1

Waveform of voice sample for phonation /a/ recorded through App2 in Group II.

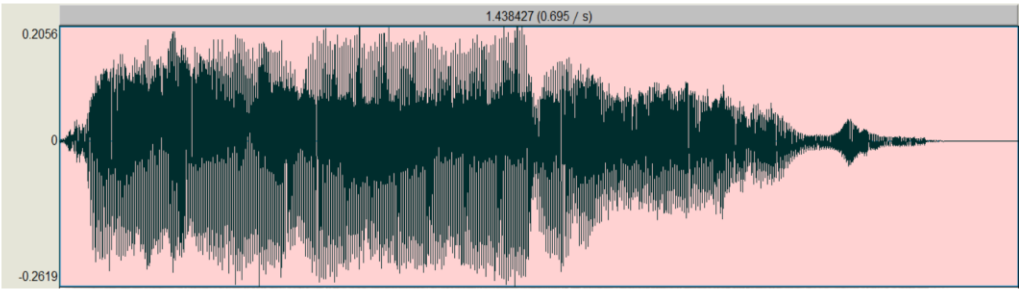
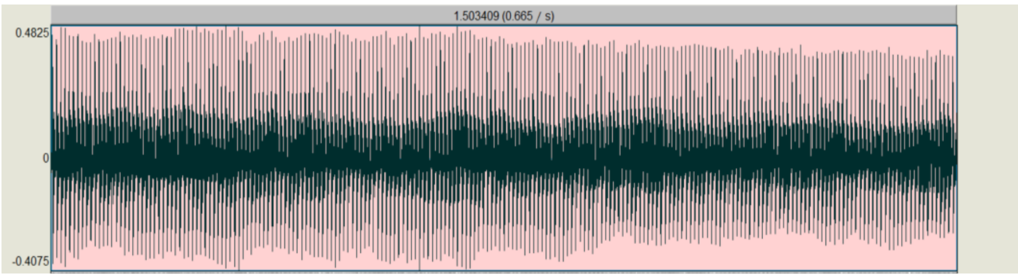


Figure 5.2

Waveform of voice sample for phonation /a/ recorded through App2 in Group I.



In dysphonic voice, abnormal glottal resistance during phonation (e.g., vocal fold paralysis, vocal nodules, polyps) results in inadequate vocal fold approximation, adding a significant amount of noise to the signal. This will further lead to more aperiodic signals so that noise and perturbation will increase in hoarse voice (Isshiki et al., 1966). As shown in Figure 5.1, more of aperiodic variability of the signal is observed in Dysphonia group and signal compression is observed after 1.5 seconds. In Figure 5.2, there was more periodic signals and less variability in the signal compression in the normal group. Comparison of the mean and SD values between the Smartphone, App1 and App2 shows little variation across the parameters such as shimmer, APQ, NHR, CPPS, TiltLTAS. However, we could not observe a significant difference in dysphonia group (Group II) whereas in normal group (Group I) a significant difference was observed in these parameters.

5.4 Difference in the acoustic parameters (AVQI & MDVP) of the voice samples of individuals with normal voice, recorded online through App1 and App2

No significant difference was observed in any of the parameters other than NHR, when the voice samples of individuals with normal voice recorded through App1 were compared with the samples recorded through App2. App2 NHR is higher than App1. Higher value of NHR indicates the lower voice quality (Titze, 1995).

5.5 Difference in the acoustic parameters (AVQI & MDVP) of the voice samples of individuals with Dysphonia, recorded online through App1 and App2

The voice sample recorded through telepractice undergoes noise suppression that detects sustained sounds and reduces intensity (Gunawan et al., 2014; Jagadeesan et al., 2006; Weerathunge et al., 2021). This could affect the acoustic measurement, and more of aperiodic signal noise perturbation present in dysphonia group, will lead to hoarseness voice (Isshiki et al., 1966). By comparing Mean and SD value, little variation was observed between App1 and App2 in dysphonia group whereas significant difference was observed for NHR in normal group. As NHR value increases the voice quality decreases (Titze, 1995).

5.6 Suitability of the voice samples recorded through App1 or App2 for voice evaluation

Table 5.1

Parameters of significant difference between the recording media in normal and dysphonic voice samples

Sl. No.	Normal voice samples			Dysphonic voice samples		
	Between mobile phone recording and App1	Between mobile phone recording and App2	Between App1 and App2	Between mobile phone recording and App1	Between mobile phone recording and App2	Between App1 and App2
1.	Shimmer	Shimmer	NHR	-	-	-
2.	APQ	APQ	-	-	-	-
3.	TiltLTAS	TiltLTAS	-	-	-	-
4.	-	CPPS	-	-	-	-

-No significant difference across any parameter

Several researchers (Lee et al., 2018; Lin et al., 2012; Vogel et al., 2014; Yun, 2015) have analysed voice samples recorded directly through mobile phones and found that these recordings are suitable for voice assessment. As shown in Table 5.1, for persons with dysphonia, there is no significant difference between the samples recorded through App1 or App2 when compared with mobile phone recordings. Hence, it can be concluded that samples recorded with App1 and App2 are suitable for voice evaluation.

5.7 Better App for voice evaluation

In the normal group, as shown in Table 5.2 (Mean, SD and Median), parameters such as Shimmer, APQ, NHR, and Tilt of LTAS was observed to be higher in App2. The shimmer and APQ values are increased as voice quality decreases as reported by Muta et al. (1988). NHR was higher in App2, higher the values indicate noise component in the signal, which indicates an abnormality in voice (Titze 1995). A tilt of LTAS is greater in App2 which reduces voice quality reduces as reported by Löfqvist & Mandersson, (1987). CPPS is lower in App2, lower values in CPPS show abnormality in voice (Hillenbrand et al., 1994; Hillenbrand & Houde, 1996). In normal group Shimmer, APQ, NHR and Tilt of LTAS was higher in App2 and CPPS value is lower in App2. App2 shows similar result in the dysphonic group, so we can conclude that App1 is better for recording voice for voice evaluation than App2.

Table 5.2

Mean, SD, Median values *of acoustic parameters of voice sample recordings through Smartphone, App1, andApp2 in normal group I*

		App1			**App2**	
Parameters	**Mean**	**SD**	**Median**	**Mean**	**SD**	**Median**
Shim (%)	4.736	0.948	4.642	6.061	2.753	5.050
APQ (%)	4.311	1.274	4.144	5.342	2.598	4.343
NHR	0.133	0.013	0.134	0.145	0.024	0.142
CPPS	12.25	0.988	12.32	11.99	0.915	12.21
TiltLTAS	-13.23	1.087	-13.46	-12.96	1.522	-13.33

Abbreviations: Shim (%), Shimmer Percent; APQ, Amplitude Perturbation Quotient; NHR, Noise to Harmonic Ratio; CPPS, Cepstral Peak prominence; App1 – video conference application 1, App2 – video conference application 2.

Chapter 6

SUMMARY AND CONCLUSIONS

Aim of the present study was to determine whether the voice samples, recorded online through two widely used video conferencing apps (App1 and App2), are suitable for assessment of voice characteristics. To find the suitability, the acoustic parameters of the voice samples recorded directly through mobile phones were compared with those of voice samples recorded online using the two apps.

A total of 48 literate Kannada speaking adults, within the age range of 20-50 years, participated in the study in two groups. Group I included 31 normal participants (15 Male and 16 Female) with a mean age range of 23.03 years (SD =1.741), and group II consisted of 17 persons with dysphonia (11 Male and 6 Female) with a mean age of 37.64 years (SD= 9.816). The researcher used two of the most widely used video conferencing apps, i.e., App1 and App2.

Voice samples of all participants were recorded for sustained phonation of /a/ and a standardized passage in Kannada. 20 acoustic parameters of the recorded voice samples were analysed using the MDVP of the CSL-model 4305 and AVQI.

6.1 Important results of the study

The important findings of the study are summarized below: -

- For persons with Dysphonia, no significant difference was found in any of the parameters between voice samples recorded using Smartphone and App1, between voice samples recorded using Smartphone and App2 and between voice samples recorded using App1 and App2. Hence, both App1 and App2 are suitable for recording voice samples for voice analysis.

- App1 was found to be slightly better than App2 for recording voice samples for voice analysis
- For normal participants, significant difference was observed in parameters such as shimmer, APQ, NHR, CPPS, and TiltLTAS between voice samples of Smartphone and App1 as well as between voice samples recorded using Smartphone and App2.
- For normal participants, for the parameter Noise to Harmonic Ratio (NHR), significant difference was observed between the voice samples recorded using App1 and App2
- For normal participants, for the parameter Cepstral Peak Prominence (CPPS), a significant difference was found between voice samples recorded using Smartphone and App2.

6.2 Implications of the study

- The study has established that voice samples recorded using video conferencing apps, App1 and App2 are suitable for acoustic analysis. Hence, voice samples recorded online using App1 and App2 can be used for assessment of voice disorders through acoustic analysis. This opens up the possibility of tele-assessment of voice disorders.

6.3 Limitations of the present study

- In the present study, comparison between the voice samples recorded through mobile phones and apps were done on the basis of acoustic parameters. Perceptual evaluation was not done for comparison.

- The subjects used their own mobile phones for direct recording as well as for connecting to the researcher through App1 and App2. Even though the price range of the mobile phones used were specified, the difference in the features between the mobile phones were not considered while comparing the samples.

- Voice samples recorded through mobile phones were taken as the reference for comparison. This was based on the earlier studies (Lee et al., 2018; Lin et al., 2012; Vogel et al., 2014; Yun, 2015) which proved that samples recorded through mobile phones are suitable for acoustic analysis. Moreover, in the present pandemic situation, it was not practically feasible to record the voice samples of the participants directly using standard recording devices such as digital recorders.

6.4 Future recommendations

- Future studies can record the voice samples directly using standard recording devices and use them as the reference for comparison.

- The perceptual assessment may also be done along with acoustic analysis for validation.
- The variability in the values of acoustic parameters of the voice samples recorded through different models and types of mobile phones may be investigated.

6.5 Significance of the results of the study

The significance of the results of the present study should be seen in the following context. Telepractice was not effectively utilized for assessment of voice disorders, as it was doubtful whether the voice samples recorded online are suitable for analysis leading to assessment. The present study has established the suitability of using voice samples recorded through video conferencing apps for acoustic analysis and thus opened up the possibility of utilizing telepractice for assessment of voice disorders. This will be extremely useful in the present pandemic situation, where the client is unable to visit the clinic for assessment. Moreover, persons with voice disorders will be able to enjoy other benefits of telepractice such as reduced time and assured availability of clinician.

REFERENCES

Aggarwal, K., Patel, R., & Ravi, R. (2020). Uptake of telepractice among speech-language therapists following COVID-19 pandemic in India. *Speech, Language and Hearing*. https://doi.org/10.1080/2050571X.2020.1812034

American Speech-Language-Hearing Association. (2004). Preferred Practice Patterns for the Profession of Speech-Language Pathology [Preferred Practice Patterns]. *Available from Www.Asha.Org/Policy. Index*. https://doi.org/10.1044/policy.PP2004-00191

American Speech-Language-Hearing Association. (2016a). *Code of ethics* [Ethics]. Available from www.asha.org/policy/.

American Speech-Language-Hearing Association. (2016b). *Scope of practice in speech-language pathology* [Scope of Practice]. Available from www.asha.org/policy/.

American Speech-Language-Hearing Association. (2018). *Scope of practice in audiology* [Scope of Practice]. Available from www.asha.org/policy/.

Aranson, A.E., & Bless, D (2009). *Clinical Voice Disorders* (4^{th} ed.). New York: Thieme

Barche, P., Gurugubelli, K., & Vuppala, A. K. (2020). Towards automatic assessment of voice disorders: A clinical approach. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 2020-Octob*, 2537–2541. https://doi.org/10.21437/Interspeech.2020-2160

Boersma P, Weenink D. *Praat: Doing phonetics by computer.* 2015. [Computer program]. Version 6004. Retrieved from http://www.praat.org.

Boone, D. R., McFarlane, S. C., Von Berg, S. L., &Zraick, R. I. (2005). The voice and voice therapy.

Brennan, D., Tindall, L., Theodoros, D., Brown, J., Campbell, M., Christiana, D., Smith, D., Cason, J., & Lee, A. (2010). A Blueprint for Telerehabilitation Guidelines. *International Journal of Telerehabilitation*, 31–34. https://doi.org/10.5195/ijt.2010.6063

Carding, P. N., Wilson, J. A., Mackenzie, K., & Deary, I. J. (2021). Measuring voice outcomes: state of the science review. *Search.Proquest.Com*. https://doi.org/10.1017/S0022215109005398

Carding, P. N., Wilson, J. A., MacKenzie, K., & Deary, I. J. (2009). Measuring voice outcomes: State of the science review. In *Journal of Laryngology and Otology* (Vol. 123, Issue 8, pp. 823–829). https://doi.org/10.1017/S0022215109005398

Carson, C. P., Ingrisano, D. R. S., & Eggleston, K. D. (2003). The effect of noise on computer-aided measures of voice: A comparison of CSpeechSP and the multi-dimensional voice program software using the CSL 4300B module and multi-speech for windows. *Journal of Voice*, *17*(1), 12–20. https://doi.org/10.1016/S0892-1997(03)00031-6

Carson, C. K., & Ryalls, J. (2018). A New Era in Acoustic Analysis : Use of Smartphone and Readily Accessible Software / Applications for Voice Assessment. *JSM Communication Disorders*, *2*(1), Aricle no. 1006.

Castillo-Allendes, A., Contreras-Ruston, F., Cantor-Cutiva, L. C., Codino, J., Guzman, M., Malebran, C., Manzano, C., Pavez, A., Vaiano, T., Wilder, F., & Behlau, M. (2020). Voice Therapy in the Context of the COVID-19 Pandemic: Guidelines for Clinical Practice. In *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2020.08.001

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. https://doi.org/10.1044/jshr.1202.246

De Bodt, M. S., Wuyts, F. L., Van De Heyning, P. H., & Croux, C. (1997). Test-retest study of the GRBAS scale: Influence of experience and professional background on perceptual rating of voice quality. *Journal of Voice*, *11*(1), 74–80. https://doi.org/10.1016/S0892-1997(97)80026-4

Dejonckere, P. H., & Lebacq, J. (1996). Acoustic, perceptual, aerodynamic and anatomical correlations in voice pathology. *ORL*, *58*(6), 326–332. https://doi.org/10.1159/000276864

Dejonckere, Philippe H., Bradley, P., Clemente, P., Cornut, G., Crevier-Buchman, L., Friedrich, G., Van De Heyning, P., 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: Guideline elaborated by the Committee on Phoniatrics of the European Laryngological Society (ELS). *European Archives of Oto-Rhino-Laryngology*, *258*(2), 77–82. https://doi.org/10.1007/s004050000299

Fetzer, D. T., & West, O. C. (2008). The HIPAA Privacy Rule and Protected Health Information. *Academic Radiology*, *15*(3), 390–395. https://doi.org/10.1016/j.acra.2007.11.008

Fong, R., Tsai, C. F., & Yiu, O. Y. (2021). The Implementation of Telepractice in Speech Language Pathology in Hong Kong during the COVID-19 Pandemic. *Telemedicine and E-Health*, *27*(1), 30–38. https://doi.org/10.1089/tmj.2020.0223

Grillo, E. U. (2017). An Online Telepractice Model for the Prevention of Voice Disorders in Vocally Healthy Student Teachers Evaluated by a Smartphone Application. *Perspectives of the ASHA Special Interest Groups*, *2*(3), 63–78. https://doi.org/10.1044/persp2.sig3.63

Grillo, E. U., Brosious, J. N., Sorrell, S. L., & Anand, S. (2016). Influence of Smartphone and Software on Acoustic Voice Measures. *International Journal of Telerehabilitation*, *8*(2), 9–14. https://doi.org/10.5195/ijt.2016.6202

Gunawan, D., Dickins, G., … P. H.-U. P. A. 14, & 2014, undefined. (2014). Spectral and spatial modification of noise captured during teleconferencing. *Google Patents*. https://patents.google.com/patent/US20140278380A1/en

Gupta, S. K., Navya, A., Swetha, G., Gupta, P., & Gopikishore, P. (2021). Telepractice in Speech-Language Pathology in India. In *Telerehabilitation in Communication Disorders and Mental Health* (pp. 16–53). https://doi.org/10.4135/9789353885915.n2

Hakkesteegt, M. M. (2009). *Dysphonia Severity Index and Voice Handicap Index*. www.ipskampdrukkers.nl

Hema, N., Mahesh, S., & Pushpavathi, M. (2009). Normative data for Multi-Dimensional Voice Program (MDVP) for adults-A computerized voice analysis system. *Journal of All India Institute of Speech and Hearing*, *28*(1), 1-7.

Hillenbrand, J., Cleveland, R. A., & Erickson, R. L. (1994). Acoustic correlates of breathy vocal quality. *Journal of Speech and Hearing Research*, *37*(4), 769–778. https://doi.org/10.1044/jshr.3704.769

Hillenbrand, J., & Houde, R. A. (1996). Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *Journal of Speech, Language, and Hearing Research*, *39*(2), 311–321. https://doi.org/10.1044/jshr.3902.311

Hirano, M. (1981). Clinical examination of voice. Wien; New York: Springer

Hirano, M., & McCormick, K. R. (1988). Clinical Examination of Voice by Minoru Hirano. *The Journal of the Acoustical Society of America*, *80*(4), 1273–1273. https://doi.org/10.1121/1.393788

Hirano M. (1981). Psycho-acoustic evaluation of voice. In *Clinical examination of voice disorders of human communication* (pp. 81–84). https://ci.nii.ac.jp/naid/10011685687/

Isshiki, N., Yanagihara, N., & Morimoto, M. (1966). Approach to the objective diagnosis of hoarseness. *Folia Phoniatrica et Logopaedica*, *18*(6), 393–400. https://doi.org/10.1159/000263069

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.

Jagadeesan, R., LK Surazski - US Patent 7, 012,901, & 2006, undefined. (2006). Devices, software and methods for generating aggregate comfort noise in teleconferencing over VoIP networks. *Google Patents*. https://patents.google.com/patent/US7012901B2/en

Kempster, G. B., Gerratt, B. R., Abbott, K. V., Barkmeier-Kraemer, J., & 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. https://doi.org/10.1044/1058-0360(2008/08-0017)

Kojima, T., Fujimura, S., Hori, R., Okanoue, Y., Shoji, K., & Inoue, M. (2019). An Innovative Voice Analyzer "VA" Smart Phone Program for Quantitative Analysis of Voice Quality. *Journal of Voice*, *33*(5), 642–648. https://doi.org/10.1016/j.jvoice.2018.01.026

Lebacq, J., Schoentgen, J., Cantarella, G., Bruss, F. T., Manfredi, C., & DeJonckere, P. (2017). Maximal Ambient Noise Levels and Type of Voice Material Required for Valid Use of Smartphone in Clinical Voice Research. *Journal of Voice*, *31*(5), 550–556. https://doi.org/10.1016/j.jvoice.2017.02.017

Lee, S. J., Lee, K. Y., & Choi, H. S. (2018). Clinical usefulness of voice recordings using a Smartphone as a screening tool for voice disorders. *Communication Sciences and Disorders*, *23*(4), 1065–1077. https://doi.org/10.12963/CSD.18540

Lin, E., Hornibrook, J., & Ormond, T. (2012). Evaluating iPhone recordings for acoustic voice assessment. *Folia Phoniatrica et Logopaedica*, *64*(3), 122–130. https://doi.org/10.1159/000335874

Löfqvist, A., & Mandersson, B. (1987). Long-time average spectrum of speech and voice analysis. *Folia Phoniatrica et Logopaedica*, *39*(5), 221–229. https://doi.org/10.1159/000265863

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. https://doi.org/10.1016/j.jvoice.2008.12.014

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. https://doi.org/10.1016/j.jcomdis.2009.12.004

Maryn, Y., Ysenbaert, F., Zarowski, A., & Vanspauwen, R. (2017). Mobile Communication Devices, Ambient Noise, and Acoustic Voice Measures. *Journal of Voice, 31*(2), 248.e11-248.e23. https://doi.org/10.1016/j.jvoice.2016.07.023

Mashima, P. A., Birkmire-Peters, D. P., Syms, M. J., Holtel, M. R., Burgess, L. P. A., & Peters, L. J. (2003). Telehealth: Voice Therapy Using Telecommunications Technology. *American Journal of Speech-Language Pathology, 12*(4), 432–439. https://doi.org/10.1044/1058-0360(2003/089)

McGill, M., & Fiddler, K. (2021). A User's Guide for Understanding and Addressing Telepractice Technology Challenges via ZOOM. *Perspectives of the ASHA*

Special Interest Groups, 1–6. https://doi.org/10.1044/2021_persp-20-00100

Mohan, H. S., Anjum, A., & Rao, P. K. S. (2017). A Survey of Telepractice in Speech-Language Pathology and Audiology in India. *International Journal of Telerehabilitation*, *9*(2), 69–80. https://doi.org/10.5195/ijt.2017.6233

Muta, H., Baer, T., Wagatsuma, K., Muraoka, T., & Fukuda, H. (1988). A pitch-synchronous analysis of hoarseness in running speech. *Journal of the Acoustical Society of America*, *84*(4), 1292–1301. https://doi.org/10.1121/1.396628

Navya, A., Swetha, G., Gupta, P., & Gopikishore, P. (2021). Telepractice in Speech-Language Pathology in India. In *Telerehabilitation in Communication Disorders and Mental Health* (pp. 16–53). https://doi.org/10.4135/9789353885915.n2

Pathology, A. H. C. on the S. of P. in S.-L. (2007). *Scope of practice in speech-language pathology [Scope of Practice]*. https://doi.org/10.1044/policy.SP2016-00343

Petrizzo, D., & Popolo, P. S. (2020). Smartphone Use in Clinical Voice Recording and Acoustic Analysis: A Literature Review. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2019.10.006

Rangarathnam, B., McCullough, G. H., Pickett, H., Zraick, R. I., Tulunay-Ugur, O., & McCullough, K. C. (2015). Telepractice versus in-person delivery of voice therapy for primary muscle tension dysphonia. *American Journal of Speech-Language Pathology, 24*(3), 386–399. https://doi.org/10.1044/2015_AJSLP-14-0017

Rosen, C. A., Lee, A. S., Osborne, J., Zullo, T., & Murry, T. (2004). Development and validation of the voice handicap index-10. In *Laryngoscope* (Vol. 114, Issue 9 I, pp. 1549–1556). https://doi.org/10.1097/00005537-200409000-00009

Schaeffler, F., Jannetts, S., & Beck, J. (2019). Reliability of clinical voice parameters captured with Smartphone - Measurements of added noise and spectral tilt. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 2019,* 2523–2527. https://doi.org/10.21437/Interspeech.2019-2910

Scope of practice in audiology. Ad Hoc Committee on scope of Practice in Audiology. (1996). *ASHA. Supplement*, *38*(2 Suppl 16), 12–15. www.asha.org/policy.

Scope of Practice in Speech-Language Pathology. (2002). *Communication Disorders Quarterly*, *23*(2), 77–83. https://doi.org/10.1177/152574010202300203

SIG 18 Perspectives Vol. 4, No. 1, March 2014. (2014). *Perspectives on Telepractice*, *4*(1). https://doi.org/10.1044/teles4.1.1

Smith, E., Verdolini, K., Gray, S., Nichols, S., Lemke, J., Barkmeier, J., Dove, H., & Hoffman, H. (1996). Effect of voice disorders on quality of life. *Journal of Medical Speech-Language Pathology*, *4*(4), 223–244. https://doi.org/10.1016/s0194-5998(05)80764-8

Tindall, L. R., Huebner, R. A., Stemple, J. C., & Kleinert, H. L. (2008). Videophone-delivered voice therapy: A comparative analysis of outcomes to traditional delivery for adults with Parkinson’s disease. *Telemedicine and E-Health, 14*(10), 1070–1077. https://doi.org/10.1089/tmj.2008.0040

Titze, I. R. (1995). *Workshop on acoustic voice analysis: Summary statement*. National Center for Voice and Speech.

Vishali (2019). Acoustic voice quality index (AVQI) in Tamil language. An Unpublished Masters' Dissertation, Submitted to University of Mysore, Mysore.

Vogel, A., methods, P. M.-B. research, & 2008, undefined. (2008). Comparison of voice acquisition methodologies in speech research. *Springer*, *40*(4), 982–987. https://doi.org/10.3758/BRM.40.4.982

Vogel, A. P., Rosen, K. M., Morgan, A. T., & Reilly, S. (2014). Comparability of modern recording devices for speech analysis: Smartphone, landline, laptop, and hard disc recorder. *Folia Phoniatrica et Logopaedica*, *66*(6), 244–250. https://doi.org/10.1159/000368227

voice, M. H.-C. examination of, & 1981, undefined. (n.d.). Psycho-acoustic evaluation of voice. *Ci.Nii.Ac.Jp*. Retrieved May 20, 2021, from https://ci.nii.ac.jp/naid/10011685687/

Weerathunge, H. R., Segina, R. K., Tracy, L., & Stepp, C. E. (2021). Accuracy of Acoustic Measures of Voice via Telepractice Videoconferencing Platforms. *Journal of Speech, Language, and Hearing Research*, 1–14. https://doi.org/10.1044/2021_jslhr-20-00625

Weidner, K., & Lowman, J. (2020). Telepractice for Adult Speech-Language Pathology Services: A Systematic Review. *Perspectives of the ASHA Special Interest Groups*, *5*(1), 326–338. https://doi.org/10.1044/2019_persp-19-00146

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

Winholtz, W. S., & Titze, I. R. (1998). Suitability of minidisc (MD) recordings for voice perturbation analysis. *Journal of Voice*, *12*(2), 138–142. https://doi.org/10.1016/S0892-1997(98)80032-5

Wuyts, F. L., De Bodt, M. S., Molenberghs, G., Remacle, M., Heylen, L., Brussels, B. M., Kristiane, B., Lierde, V., Raes, J., & Van De Heyning, P. H. (2000). The Dysphonia Severity Index: An Objective Measure of Vocal Quality Based on a Multiparameter Approach. In *Journal of Speech, Language, and Hearing Research* • (Vol. 43, Issue 3). American Speech-Language-Hearing Association. https://doi.org/10.1044/jslhr.4303.796

Yashaswini, R., and P. K. S. R. (2018). Tele speech-language pathology and audiology in india - a short report. *Journal of the International Society for Telemedicine and EHealth*, *January*, 1–8. https://doi.org/10.29086/JISfTeH.6.e19

Yun, M.-H., Lee, J.-H., Lee, S.-H., & Jin, S.-M. (2015). Feasibility of Galaxy Smartphone Recording as Portable Recorder for Acoustic Analysis of Voice. *Journal of The Korean Society of Laryngology, Phoniatrics and Logopedics*, *26*(2), 104–111. https://doi.org/10.22469/jkslp.2015.26.2.104

Appendix

KANNADA PASSAGE

ಬೆಂಗಳೂರು ನಮ್ಮ ರಾಜ್ಯದ ಒಂದು ದೊಡ್ಡ ಊರು.

ಈ ಊರನ್ನು ನಮ್ಮ ರಾಜ್ಯದ 'ಬೊಂಬಾಯಿ' ಎನ್ನುವರು.

ಇಂಡಿಯಾದ ದೊಡ್ಡ ನಗರಗಳಲ್ಲಿ ಇದೂ ಒಂದು.

ಈ ಊರನ್ನು ನೋಡಲು ಜನರು ಬೇರೆಬೇರೆ ಊರುಗಳಿಂದ ಬರುವರು.

ಇದಲ್ಲದೆ ನಮ್ಮ ರಾಜ್ಯದಲ್ಲಿರುವ ಬೇಲೂರು, ಜೋಗ, ನಂದಿ,

ಇವುಗಳನ್ನು ನೋಡಲು ಜನರು ಬರುವರು.

ಈ ನಾಡಿನಲ್ಲಿ ರೇಷ್ಮೆಯನ್ನು ಬೆಳೆಯುವರು.