CPPS Normative in Hindi Speakers Using a Mixed-Phoneme and an All-Voiced Text With and Without Voicing Detection Algorithm

MOKSH

Registration Number: P01II22S123017

This dissertation is submitted as a part of fulfillment for the degree of Master of Science (Speech-Language Pathology)

University of Mysore

All India Institute of Speech and Hearing, Manasagangothri,

Mysore – 570006

July, 2024

CERTIFICATE

This is to certify that dissertation entitled "**CPPS Normative in Hindi Speakers Using a Mixed-Phoneme and an All-Voiced Text With and Without Voicing Detection Algorithm**" is a bonafide work submitted as a part of the fulfillment for the degree of Master of Science (Speech Language Pathology) of the student registration number: P01II22S123017. It has been carried out under the guidance of a faculty of this institution and has not been submitted earlier to any other university for the award of any diploma or degree.

Mysore

July, 2024

Dr. M. Pushpavathi

Director

All India Institute of Speech and Hearing

Manasagangothri, Mysore – 570006

CERTIFICATE

This is to certify that dissertation entitled "**CPPS Normative in Hindi Speakers Using a Mixed-Phoneme and an All-Voiced Text With and Without Voicing Detection Algorithm**" has been prepared under my 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 diploma or degree.

Mysore

July, 2024

Dr. T. Jayakumar

Professor

Department of Speech Sciences

All India Institute of Speech and Hearing

Manasagangothri, Mysore – 570006

DECLARATION

This is to certify that dissertation entitled "**CPPS Normative in Hindi Speakers Using a Mixed-Phoneme and an All-Voiced Text With and Without Voicing Detection Algorithm**" is a result of my own study done under the guidance of Dr. T. Jayakumar, Professor, Department of Speech Sciences, All India Institute of Speech and Hearing, Mysore. This dissertation has not been submitted earlier to any other university for the award of any diploma or degree.

Mysore **Registration No.: P01II22S123017**

July, 2024

ACKNOWLEDGEMENT

First and foremost, I would like to thank the Director of our institution, **Dr. M Pushpavathi**, for allowing me to carry out this research work.

I would like to thank **Dr. T Jayakumar**, my guide for this dissertation for guiding me through this journey of research. It's worth mentioning that the heights of his knowledge, subject expertise, practical orientation, logical approach is his conduct, calmness, patience and kindness are unmatchable. His quality of patient and uninterrupted listening and valuing others' ideas and thoughts have been truly inspirational.

I would further extend my gratitude to **Dr. Brajesh Priyadarshi** for being an expert in the content validation part of this study.

I further extend my regards and gratitude to **Mr. Srinivas** for helping me with the entire statistical analysis involved in this study. Not only did he guide me with the appropriate tests required in this study, but also extended his helping hand towards me in learning the SPPS software and giving me additional insights about the research and statistics. Thanks to **Ms. Harshada Mali**, PhD scholar at AIISH, for helping me with SPSS in the beginning. You were truly the kick starter for me in the statistical analysis part of this study,

It's also important not to forget **Mr. Jesnu Jose**, who has been no less than a co – guide for me especially in the starting phases of the study. He was always there to help me out with my doubts and would always fill the gap in my guide's absence. Thank you so much sir!

The entire process of this research was impossible without **Ms. Hani Sulthana**, **Ms. Delcitta Fernandes**, and **Ms. Aadarshini N.** (my dearest batchmates), along with

Ms. Kiran and **Ms. Sakshi** from BVP, Pune and **Ms. Tanu** from AYJNISHD, Kolkata, for accepting my request to perceptually rate the voice samples as a part of perceptual screening of voice samples in this study. Thanks a tonne for being a part of this study!

Last, but not the least, I would like to acknowledge all the **participants of this study** for accepting my request to be a part of this study. This study wouldn’t exist without you all. A heartfelt gratitude to all of you!

To anyone who contributed to this study even in the tiniest way, whom I have forgot to mention, please accept my apologies and gratitude alongside for making this study happen!

Moksh

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract

Acoustic analysis of voice has proven to be a cost – effective, easy and reliable method of instrumental voice assessment over the past few decades and Smoothened Cepstral Peak Prominence (CPPS) has been found to be the most sensitive acoustic measure for voice analysis across many research studies. Many authors in the past have recommend the usage of an all – voice sample recording for CPPS analysis that is predicted to yield lesser false positives. Alongside, two different types of algorithms are incorporated in various acoustic analysis softwares for CPPS analysis. The first type uses an Automatic Voicing Detection (AVD) algorithm which extracts only the voiced segments from the sample and computer CPPS values on the extracted voiced sample. Whereas, the second type computes the CPPS on the whole sample. Both algorithms have their own merits and demerits. Since CPPS is one of the most valid and reliable acoustic measures of voice quality, but there is no study in the literature that has found CPPS normative in Hindi speakers using continuous speech sample, a study to establish this was needed to be done. This study aims to establish CPPS normative for different sample stimuli, i.e., sustained vowel phonation samples (/a/, /i/, and /u/), a mixed-sentence stimulus (that contains both voiceless and voiced phonemes) and a voiced-sentence (that contains only voiced phonemes) in young adult Hindi speaking normophonic individuals with and without voicing detection algorithm. 92 participants within the age range of 19 – 30 years were recruited for this study. Samples of all participants were subjected to perceptual and acoustic screening to ascertain normal quality of voice. Final screened sample consisted of 37 females and 36 males. All the screened samples were subjected to CPPS analysis using Praat software. Mean CPPS values for /a/, /i/, /u/, mixed-sentence, and voiced-sentence were reported. Use of AVD yielded higher CPPS values in comparison to when AVD wasn't used. Overall, male samples yielded higher CPPS values than female samples. Amongst sustained phonation samples, /a/ had highest CPPS values followed by /i/ and /u/ in both female and male subgroups. Similarly voiced-sentence yielded higher CPPS values in comparison to mixed-sentence in both female and male subgroups. Given the disadvantages of using Automatic Voicing Detection, and low intertext variability and better sensitivity of CPPS when computed upon an all – voiced text, computation of CPPS using an all – voiced text without the use of automatic voicing detection is suggested for clinical purposes. However, further sensitivity and specificity of the obtained CPPS normative should be calculated by computing CPPS on dysphonic individuals using the stimuli and methods used in this study.

Keywords: AVD, CPPS, mixed-sentence, Praat, voiced-sentence

CHAPTER 1

Introduction

Dysphonia (abnormal voice) is prevalent amongst all age groups of people in the society. Lifetime prevalence of dysphonia is found to be between 18.79% and 21.9% in non-professional voice users across various studies (Liang-Hui et al., 2023; Roy et al., 2005; Sheyona & Devadas, 2022) and as high as 70.1% in professional voice users such as teachers, singers, etc. (Menon et al., 2021). Given such a high incidence and prevalence of voice problems, accurate assessment and diagnostic methods as well as treatment methods are necessary.

Over the time, clinical voice assessment methods have evolved from basic auditory-perceptual assessment of voice where clinicians try to describe voice in terms of its pitch, loudness and quality to more advanced acoustic, aerodynamic, electrophysiologic and visual imaging methods. Guidelines for assessment of voice quality (VQ) given by European Laryngology Society includes baseline VQ anamnesis, videolaryngostroboscopy, patient reported VQ assessment, auditory-perceptual, acoustic and aerodynamic assessment along with instruments associated with voice comorbidities (Lechien et al., 2023). ASHA guidelines on instrumental evaluation of voice recommends use of laryngeal endoscopic imaging, acoustic evaluation and aerodynamic evaluation for clinical assessment of voice (Patel et al., 2018).

Amongst all the instrumental measures of voice evaluation, acoustic evaluation of voice is an important method due to benefits like the procedure being relatively simple and non-invasive, quick, and low-cost involvement. Amongst various parameters that are representative of voice quality like Low High Spectral Ratio (LHSR), Harmonics to

Noise Ratio (HNR), Noise to Harmonics Ratio (NHR), Jitter, Shimmer, Cepstral Peak Prominence (CPP), Cepstral Peak Prominence Smoothened (CPPS), Degree of Voice Breaks (DVB), etc., CPPS has a growing evidence of support for its usability given its better diagnostic accuracy (Patel et al., 2018).

To obtain CPP/ CPPS, at first, waveform (time domain) undergoes Fourier transform to give a spectrum (frequency domain) which further undergoes another Fourier Transform in log scale to give a cepstrum which is plotted as "quefrency" on the x-axis and "intensity" on y-axis. Then the highest peak in that plot is identified within the range of expected F0. Also, a linear regression analysis is performed to give a trend line. The difference in the amplitude of highest marked peak and the value of ordinate of the trend line below that peak gives CPP. Further an average of cepstrum across time and quefrency yields smoothened cepstrum and similarly calculated difference in amplitude of highest peak and ordinate value of trend line on the point just below the peak yields CPPS (smoothened) (Hillenbrand & Houde, 1996).

There is a good amount of literature that suggests CPPS to be one of the most valid/ reliable measure of voice quality. "CPP/ CPPS is amongst the most robust correlates of breathiness in sustained vowels as well as in continuous speech." (Hillenbrand & Houde, 1996). CPPS is a reliable measure that should be a part of routine instrumental voice analysis (Heman-Ackah et al., 2003). Awan and Roy (2006) indicated that CPP/ CPPS "may be the most significant component" in their "four-factor model for measuring dysphonia severity". CPPS can be viewed as potentially the most accurate acoustic algorithms or single correlates of overall voice quality (Maryn et al., 2009).

American Speech and Hearing Association (ASHA) recommends use of CPPS for measurement of overall quality of the voice signal (Patel et al., 2018).

A group of studies advocate for the use of an all – voiced linguistic stimulus for CPPS computation. Studies have found the voiced sentence “We were away a year ago” to correlate best with auditory perception of dysphonia (Awan et al., 2010; Roy, 2010). Some voice scientists believe that all-voiced structure of the stimulus elicits vocal behavior which in presence of dysphonia inducing etiologies best elicits irregular phonatory physiology (Watts & Awan, 2015). Kim et al. (2021) found that “CPP calculated from voiced speech segments extracted with AVD was less effective at discriminating between speakers with and without a voice disorder than CPP calculated on sustained vowels or a sentence composed primarily of voiced phonemes.”

Although voicing detection algorithms are available to overcome the previously mentioned problems, but the accuracy of the algorithms is still questionable as pointed out in the following studies. Awan et al. (2010) pointed out that failure in accurate voicing detection could cause the differences observed in the CPPS values for sentences with and without unvoiced segments. “AVD may remove aphonic periods from dysphonic participants, resulting in an artificially increased CPP value” (Awan et al., 2009).

Various studies have been done across languages like English, Portuguese, Iranian, Finnish, German, Kannada, Hindi, etc., to establish normative and/ or cutoff (threshold values) for various stimuli (sustained vowels and continuous speech); across various age groups, and using different software programs like Praat, ADSV, Hillenbrand tool, etc.

There is so far only one study done in Hindi speaking individuals which found out the mean CPPS values using sustained vowel phonation /a/ sample (Soni et al., 2023) and no studies have been done using continuous speech/ reading sample which is believed to be a more valid indicator of voice quality.

Need of the Study

Given the disadvantages of using Automatic Voicing Detection, and low intertext variability and better sensitivity of CPPS when computed upon an all – voiced text, computation of CPPS using an all – voiced text without the use of automatic voicing detection becomes a better option to improve diagnostic accuracy. Also limiting the sample to a single sentence without punctuations would prevent lowering of CPPS values due to silent pauses that would normally occur within the reading or conversational sample.

Literature shows that CPPS has great diagnostic value. CPPS computed using an all-voiced text stimulus without the use of Automatic Voicing Detection (AVD) is a more valid measure of voice quality giving lesser of false negative and false positive results and thus higher sensitivity and specificity (Awan et al., 2009, 2010; Kim et al., 2021; Kitayama et al., 2020). Also, no studies have been done in Hindi so far to establish CPPS normative. Thus, it is required to establish CPPS normative in Hindi speakers.

Aim of the Study

The primary aim of this study is to establish CPPS normative in young adult native Hindi speakers using sustained vowel and reading samples (one mixed-phoneme sentence (having both voiced and voiceless phonemes) and another all-voiced sentence (having only voiced phonemes) with and without voicing detection algorithm. Present study further aims to compare CPPS values obtained using different kind of stimuli and using different algorithms (with and without AVD) in the Praat software.

Objectives of the Study

The objectives of the present study are as follows:

1. To establish CPPS normative in young adult native Hindi speakers using sustained vowel and reading samples.
2. To study the effect of Automatic Voicing Detection (AVD) algorithm on CPPS when obtained using sustained vowel and reading samples.
3. To compare CPPS values obtained for males and females.
4. To compare the CPPS values obtained using different stimulus (sustained vowel and reading sample).

We hypothesize that CPPS values computed with an all-voiced sentence is significantly greater than the one computed using a mixed-phoneme text when not using AVD. Our second hypothesis is that CPPS values computed while using AVD are significantly greater than when not using AVD on analyzing a mixed-phoneme sentence. Our third hypothesis is that there is no significant difference in CPPS values when computed with using AVD vs when not using AVD on analyzing an all-voiced sentence.

CHAPTER 2

Literature Review

Human voice is a tool which allows one to communicate his/ her thoughts and feelings using speech as the primary mode of communication. Alongside, we as humans also use vocal acts such as laughing, crying, yelling, etc., to indicate different emotions. Human voice also allows one to sing, mimic, and act which serves as a source of joy and entertainment for the individual himself and others. All the above-mentioned abilities get affected when one's voice gets affected due to any pathology leading to Dysphonia. Dysphonia (abnormal voice) is prevalent amongst all age groups of people in the society as discussed below.

Prevalence of Dysphonia

Roy et al. (2005), in their questionnaire-based interview study comprising of 1326 participants (general population) reported that 29.6% individuals had experienced voice problems at least once in their lifetime while 6.6% individuals reported a currently present voice problem.

Bhattacharyya (2014) reported based upon their analysis of 2012 National Health Interview Survey (NHIS) that nearly 7.7% adults in USA experience voice problems every year. Bhattacharyya (2015) based upon their analysis of 2012 NHIS pediatric voice and language module found out that nearly (1.4% ± 0.1%) of children in United States of America reported voice problems in the span of 1 year.

A systematic review and meta-analytical study done by Liang-Hui et al. (2023), in which they reviewed 13 articles (published between 2006 – 2019), suggests about 18.79% prevalence of voice problems in older adults. Their findings further suggest a higher prevalence of dysphonia in institutionalized adults (33.03%) than community based older adults (15.2%).

Incidence and prevalence of voice problems has also been studied in special groups basically comprising of occupational voice users. e.g., singers, teachers, actors, SLPs, etc. Menon et al. (2021), in their questionnaire-based interview research conducted across 28 schools and 702 teachers, reported a prevalence of 45.4% for current voice difficulty, 52.8% for voice problems in past 1 year, and 70.1% for any voice problems experienced during their entire teaching career.

Sheyona and Devadas (2022), in a cross-sectional survey of 500 non-professional voice users who were working in various schools/ colleges reported a lifetime prevalence of 21.6% with 4.9% participants reporting presence of voice problem at the time of survey. Females (64%) reported significantly higher voice problems than males (35.8%).

Oliveira et al. (2023), in their systematic review and meta-analytical study that included reviewing and analyzing 73 articles (total sample size of 63,126 voice professionals) reported a prevalence of 44.0%. In auditory perceptual judgement, a greater number of teachers were found to have voice problems than non-teaching voice professionals.

Give such a high prevalence of voice disorders, there exists a need of effective screening and assessment tools for voice analysis and CPPS can serve as an effective parameter for screening voice quality.

Why Use CPPS?

Titze (1995) categorized voice signals into 3 types. Type I signals fundamentally represents the normal voice signals. Whereas, Type II and Type III signals represents voice signals that display more of subharmonic energy and no evident periodicity respectively. Given the above understanding of the Type II and Type III signals, there is no obvious fundamental frequency in either type of voice signal. Given this fact, perturbation analysis is not a reliable measure of Type II and Type III voice signals. Thus, the other indicators of quality of voice, i.e., Jitter, Shimmer, HNR and NHR become less valid and reliable indicators of quality of voice for Type II and Type III signals.

“CPP is amongst the most robust correlates of breathiness in sustained vowels as well as in continuous speech.” (Hillenbrand & Houde, 1996)

“Cepstral measures, and smoothed cepstral peak prominence, in particular, can be viewed as potentially the most accurate acoustic algorithms or single correlates of overall voice quality.” (Maryn et al., 2009)

CPPS when analyzed using a continuous speech sample leads to better performance on measures such as sensitivity, specificity, Positive Predictive Value, and Negative Predictive Value when compared with jitter/ shimmer/ HNR that in turn indicates better diagnostic accuracy of CPPS. “Smoothed cepstral peak prominence are reliable measures that should become routine in instrumental voice analysis.”(Heman-Ackah et al., 2003)

Awan and Roy (2006) indicated that CPP “may be the most significant component” in their “four-factor model for measuring dysphonia severity”.

Maryn et al. (2009) did a meta-analytical study of 25 published research papers which investigated the use of acoustic parameters for voice analysis. They calculated weighted correlation coefficient between perceptual judgements and acoustic markers for a total of 87 acoustic measures for sustained vowels and connected speech sample combined. In relevance to studies on vowels, the authors found out that the following four measures met the criterion of a homogeneous weighted (r) >= 0.60: Pitch amplitude, Pearson r at autocorrelation peak, CPPS, and Spectral flatness of residual signal. W.r.t continuous speech, the same criterion was met be the following three parameters: SNR from Qi, CPP, and CPPS. "Consequently, these six measures are considered to be the most promising measures for the acoustic measurement of overall voice quality, as compared to the remaining 81 measures included in the original meta-analysis." (Maryn et al., 2009)

"If appropriate precautions are not followed and, for example, perturbation analysis is performed on a patient with a profoundly disordered voice producing a type III signal, meaningless numbers will be generated. Such data may be worse than no measures at all and may mislead clinicians trying to design therapy or assess outcomes." (Sataloff, 2017).

"For measuring the overall level of noise in the vocal signal, the recommendation is to use a measure of the vocal cepstral peak prominence." (Patel et al., 2018)

Factors Affecting CPPS Value (Vowel Context, Sample Duration, Vocal Intensity, Fundamental Frequency)

Awan et al. (2012) found out that that there was a significant effect of loudness/ intensity on CPPS (CPPS increased with increase in vocal intensity); age (males having greater CPPS than females); and vowel type (low vowels [/ɑ/ and /æ/] yielded greater CPPS value vs high vowels [/i/ and /u/]).

Sampaio et al. (2020) found out that there was a significant effect of fundamental frequency, sample duration and intensity on CPPS value (Direct relation with F0 and

Intensity and inverse relation with sample duration). Herein sample duration was considered w.r.t extracted vowels from within the connected speech sample.

Brockmann-Bauser et al. (2021) found that higher CPPS values were found out for both patient and control group when vocal intensity was higher, whereas, F0 had a weak correlation with CPPS for both control group and patient group.

With respect to cultural and sociolinguistic effects, Procter and Joshi (2022) found out that CPP values were significantly higher for Standard American English speakers than those with first acquired language (L1) as Spanish or French for the sentence "We were away a year ago" (that is completely voiced).

Advocacy for use of All – Voiced Linguistic Stimulus

Many researchers have found the voiced sentence "We were away a year ago" to correlate best with auditory perception of dysphonia (Awan et al., 2010; Roy, 2010).

The reason why CPPS computation on all – voiced sentences have best correlation with the perceptual evaluation of dysphonia is not very clear. "One theory is that the all-voiced structure of the stimulus somehow elicits vocal behavior which, when dysphonic inducing etiologies are present, will best elicit irregular phonatory physiology" (Watts & Awan, 2015).

Although voicing detection algorithms are available, but the accuracy of the algorithms is still questionable as pointed out in the following studies. Awan et al. (2010) pointed out that failure in accurate voicing detection could cause the differences observed in the CPPS values for sentences with and without unvoiced segments. "Our results suggest that ADSV's voice activity detector may not filter out all the unvoiced frames in an utterance" (Murton et al., 2020). If voicing detection algorithm was accurate enough to cut all the unvoiced segments there would be no significant difference in the CPPS values obtained using different sentences. However, this is not always true. Murton et al. (2020) found the CPPS value to be higher for the all-voiced sentence in comparison to other sentences used from CAPE – V with both Praat and ADSV (given the fact that ADSV uses automatic voicing detection).

With respect to inter-text CPPS variability, Kitayama et al. (2020) found out that there was moderate to large inter-text variability for CPPS between different passages (which decreased using deletion of silent segments) while the inter-text variability was comparable for all voice texts and sustained vowels.

"If frames that were not intended to be voiced (including pauses and voiceless consonants) could be accurately excluded, then utterances with different phonemes could be compared. Aphonic segments, which occur during speech that is intended to have voicing, would be included in the CPP computation and lower the result." (Murton et al., 2020)

Kim et al. (2021) found out that "CPP calculated from voiced speech segments extracted with AVD was less effective at discriminating between speakers with and without a voice disorder than CPP calculated on sustained vowels or a sentence composed primarily of voiced phonemes."

"AVD may remove aphonic periods from dysphonic participants, resulting in an artificially increased CPP value." (Awan et al., 2009)

CPPS Normative Studies Across Languages, Genders and Different Age Groups

Various studies have been done to establish normative and/ or cutoff (threshold values) in different languages (English, Spanish, Korean, Portuguese, Kannada, Iranian, Finnish, Persian, German, and Hindi) for various stimuli (vowels and running speech); across various age groups, and using different software programs like Praat, ADSV, and Hillenbrand speech tool. The list of studies and their findings is mentioned in the Table 2.1. Studies which have investigated 'CPP' have not been included in the list. Only those studies which have investigated 'CPPS' have been included in the list.

Monnappa and Balasubramanium (2015) found an increase in CPPS values (computed using sustained vowel phonation /a/) with age with old aged participants scoring more than middle aged participants followed by young aged participants indicating presence of a well-defined harmonics in old aged individuals. They also reported that females had smaller CPPS amplitudes than males across all age groups which they attributed to poor harmonic structures in females attributed to the prevalence of posterior glottal chink in females.

Hasanvand et al. (2017) also reported lower CPPS values for females in comparison to males for both dysphonic and control group. Mendes et al. (2023) in their study on fado singers reported to have seen lower CPPS values for females than male speakers and young speakers having greater CPPS values than old speakers.

Oliveira Santos et al. (2021) in a perceptuo-acoustic study of voice changes in males and females across five decades of life (from age 30 – 79 years) found no significant effect of gender on CPPS values except for participants in the 3rd decade of life wherein females reported significantly lower CPPS values than males. When studying age-related changes (considering both males and females combined), significant difference was found only in speakers from 3rd and 7th decade of life wherein older adults presented with higher CPPS values.

Spazzapan et al. (2022) in a CPPS analysis of pediatric voices found out that CPPS values were significantly higher for children aged 13 – 18 years old than their younger counterparts. Sex related differences were indicated only for children aged 12 years and above (males having higher CPPS than females).

Sujitha and Pebbili (2022) found out similar sex related difference in young adult Kannada speakers (males reporting higher CPPS values than females). They also reported intertext differences attributed to presence of more/ less voiced phonemes in the text.

Concluding findings from all the age and gender related studies, most of the studies have found higher CPPS values for males than females. Many studies report increase in CPPS values from young to older adulthood and from early childhood to adolescence.

CPPS Normative in Native Hindi Speakers

Within the published literature, only one study has been done by Soni et al. (2023) to find out normative CPPS values in Hindi speakers. They found out mean CPPS values on sustained phonation /a/ samples using Praat software in two age subgroups of 18 – 25 years and 26 – 40 years (Table 2.1). No studies have been carried out on continuous speech sample/ reading sample in Hindi speakers so far.

Table 2.1

CPPS Normative Across Various Languages

				CPPS (+/-SD) (dB)	
Authors	**Language**	**Age Range**	**Software**	**Vowel /a/**	**Continuous speech**
Heman-Ackah et al., 2014	English	-	Hillenbrand	-	4.77 (0.97)
Brinca et al., 2014	Portuguese	19 - 66 (females)	Praat	-	-
Reddy et al., 2014	English	17 - 30 (untrained singers)	Hillenbrand	6.80 (0.53)	3.63 (0.48)
		20 - 30 (non - singers)	Hillenbrand	3.48 (2.31)	2.20 (0.46)
Balasubramanium et al., 2015	-	18 - 30 (Carnatic singers)	Hillenbrand	8.42	-
		(Non-singers)		5.18	-
Monnappa & Balasubramanium, 2015	-	18 - 40 (Males)	Hillenbrand	9.84 (1.53)	-
		18 - 40 (Females)		8.34 (1.45)	-
		40 - 60 (Males)		6.38 (1.64)	-
		40 - 60 (Females)		7.75 (1.48)	-
		> 60 (Males)		10.36 (1.63)	-
		> 60 (Females)		8.73 (1.13)	-
Sauder et al., 2017	English	18 - 85	ADSV	-	5.89 (1.00)
			Praat	-	20.11 (1.27)
Castellana et al., 2018	-	21 - 58	Hillenbrand	18.20 (1.30)	-
Núñez-Batalla et al., 2019 (**Cutoff value**)	Spanish	20 - 60 (Female)	Praat	16.00	7.90 - 11.30
		20 - 60 (Male)		16.40	7.80 - 10.90
Phadke et al., 2020	Finnish	42.6 (8.9) (Females)	Praat	13.90 (1.90)	10.50 (1.20)
Murton et al., 2020 **(Cutoff value)**	English	22 - 59	Praat	14.45	9.30
			ADSV	11.46	6.11
Hassan et al., 2021	English	20 - 50	Praat	23.95 (3.69)	20.72 (1.72)
Oliveira Santos et al., 2021	-	30 - 39 (Males)	Praat	16.42 (1.64)	-
		40 - 49 (Males		16.31 (2.75)	-
		50 - 59 (Males)		6.84 (2.23)	-
		60 - 69 (Males		17.60 (3.19)	-
		70 - 79 (Males)		17.08 (3.37)	-

		30 - 39 (Females)		15.03 (2.29)	-
		40 - 49 (Females)		15.14 (2.03)	-
		50 - 59 (Females)		16.62 (2.45)	-
		60 - 69 (Females)		16.65 (3.09)	-
		70 - 79 (Females)		17.21 (2.44)	-
Spazzapan et al., 2022	Portuguese	5 - 7 (Males)	Praat	15.16 (1.53)	-
		8 - 9 (Males)		14.86 (1.81)	-
		10 - 11 (Males)		15.20 (2.02)	-
		12 (Males)		16.13 (1.76)	-
		13 - 15 (Males)		17.59 (2.45)	-
		16 - 18 (Males)		18.59 (2.02)	-
		5 - 7 (Females)		14.53 (1.92)	-
		8 - 9 (Females)		13.99 (1.81)	-
		10 - 11 (Females)		14.33 (2.10)	-
		12 (Females)		14.55 (1.96)	-
		13 - 15 (Females)		14.74 (2.30)	-
		16 - 18 (Females)		14.93 (2.04)	-
Sujitha & Pebbili, 2022	Kannada	20 - 30 (Males)	Hillenbrand	8.64 (1.05)	5.65 (0.79) 5.22 (0.60)
		30 - 40 (Males)		7.73 (1.09)	4.93 (0.73) 4.49 (0.53)
		20 - 30 (Females)		7.37 (0.93)	5.70 (0.85) 5.25 (0.61)
		20 - 30 (Females)		7.63 (1.27)	6.07 (1.12) 5.37 (0.98)
Kim et al., 2022	Korean	56.3 (11.9)	Praat	16.50 (2.80)	13.40 (2.30)
Saeedi et al., 2022	Persian	37 (11)	Praat	16.22 (2.43)	11.54 (1.56)
Barsties v. Latoszek et al., 2023 **(Threshold)**	German	26.79 (7.06)	VoxPlot	15.02 (Hoarseness), 14.47 (Roughness)	-
Buckley et al., 2023	English	18 - 91 (Males)	Praat	17.52 (2.90)	8.92 (1.26)
		18 - 91 (Females)		16.17 (2.56)	9.17 (1.34)
Mendes et al., 2023	Portuguese	18 - 66 (Females)	Praat	17.15 (2.12)	9.40 (0.90)
		20 - 70 (Males)		19.63 (2.32)	8.56 (1.06)
Soni et al., 2023	Hindi	18 – 25 (Females)	Praat	14.39 (2.27)	-
		26 – 40 (Females)		16.08 (2.04)	
		18 – 25 (Males)		16.28 (2.52)	
		26 – 40 (Males)		17.62 (2.28)	

CHAPTER 3

Methods

Study Design

Normative and Comparative study.

Participants

A total of 92 young adult native Hindi speakers in the age range of 18 to 30 years were recruited for the study. Amongst the 92 participants, 46 were males and 46 were females. Exclusion criteria consisted of any voice and/ or respiratory problems/ complaints; history of cold and cough within past 1 month; history of chronic smoking or alcohol/ drug abuse; and history of any neurological and/ or speech and language problem.

Development and Validation of the Stimuli

Two type of reading stimuli were used to conduct this study. One sentence having a combination of voiceless and voiced phonemes of the Hindi language termed as "mixed-sentence" (Figure 3.1) and other sentence having only voiced phonemes termed as "voiced-sentence" (Figure 3.2). These two sentences were developed by the author and they were subjected to validation by a Linguist who is also a native Hindi speaker and has more than 20 years of experience in academics and Clinical Linguistics for a list of features. All the features were rated on a 5-point Likert type rating scale (Appendix). The features for which sentences were rated were as follows: "the sentence has only voice phonemes" (applicable only for voice-sentence); "all phonemes and words used pertain to Hindi language"; "the words used in the sentence are commonly used by native Hindi speakers in routine conversations"; "the overall morphosyntactic structure of the sentence is simple"; "the sentence is grammatically correct"; and "the sentence is coherent and cohesive".

Figure 3.1

Mixed-sentence Text

गले में ध्वनि पैदा करने वाला एक यंत्र होता है जो आवाज़ बनाता है।
Total no. of syllables: 23
Total no. of words: 14

Figure 3.2

Voiced-sentence Text

गले में मौजूद आवाज़ बनाने वाले अंग ने आवाज़ बनाई और मैं बोलने लगा।
Total no. of syllables: 25
Total no. of words: 14

Equipment

For the purpose of voice recording, Boya BY – M1 Lavalier Condenser Microphone (Omnidirectional) having a frequency range of 65Hz to 18kHz was used. The microphone was head-mounted with 'Rode – Lav headset mount'. The microphone meets all the standards as suggested by Svec and Granqvist (2010). The voice samples were recorded and analyzed in the Praat software (Version 6.3.12) in Asus Vivobook 14 laptop.

Voice Sample Recording Setup

All the recordings were done is a quiet room. The participants were made to sit in an upright posture on a chair. The headset with microphone was placed on the participants' head with the microphone being 4 to 10 cm away from mouth at 45º to 90º azimuth (Svec & Granqvist, 2010). No pop filter was used over the microphone in order to maintain the integrity of the voice signal.

Recording Parameters

Single channel recordings were made with a sampling frequency of 16kHz (Adjustable in Praat recording window) and a microphone gain of 10dB (Adjustable in Windows Realtek Audio Console). These parameters obey the guidelines given by Svec and Granqvist (2010).

Tasks

Firstly, sustained vowel phonation productions were elicited from every participant for each of the following vowels /a/, /i/, and /u/. The participants were instructed to maintain sustained vowel productions for a minimum of five seconds.

Following vowel production, two reading sentences were elicited from every participant. First, using voiced-sentence; and second using mixed-sentence. The order of elicitation of these productions was counterbalanced. Both the sentences were given to participants prior to the final recording to familiarize them with the sentences, thus to avoid any pauses due to hesitation/ unfamiliarity of the sentences. The participants were instructed to read each sentence in one go without taking any pause within the sentence.

The participants were instructed to produce the above-mentioned productions (vowels and sentences) at a habitual level of loudness and in the habitual pitch.

Analysis

The recorded vowel samples were trimmed to middle five seconds of the production. The silent gap on either side of the recorded sentence was also removed. Only these trimmed portions were subjected to further analysis.

Samples from all participants were subjected to a perceptual screening and an acoustic screening to ascertain that the participants considered have a normal voice quality.

Perceptual Screening

Perceptual screening was done by three raters (Master's students of SLP) in which they had to rate each reading sample (mixed-sentence) for the Grade (G) parameter on the GRBAS scale. The perceptual raters were blinded to the participants' characteristics. Only those participants whose voice was rated '0' for the Grade (G) parameter on GRBAS by at least two out of three raters were considered for further analysis.

Acoustic Screening

All sustained vowel phonation /a/ samples were subjected to acoustic screening for the parameters of jitter, shimmer, and HNR. The reference values (Table 3.1) were taken from Hippargekar et al. (2022). Those participants were excluded from further analysis whose either of jitter, shimmer, or HNR value lied out of mean (+ 2 SD) for jitter and shimmer; and mean (–2 SD) for HNR.

Table 3.1

Reference Mean (SD) Values for Jitter, Shimmer, and HNR (Hippargekar et al., 2022)

Parameter	Females	Males
Jitter Local (%)	0.3 (0.12)	0.37 (0.15)
Shimmer Local (%)	3.15 (0.76)	3.31 (1.56)
HNR (dB)	21.60 (1.71)	22.18 (1.71)

Those participants' samples, who passed both acoustic and perceptual screening were subjected to CPPS analysis (both with and without automatic voicing detection) using the Praat plugin given by Heller Murray et al. (2022). This plugin returns CPPS values in decibels (dB) for selected samples both using automatic voicing detection and without using automatic voicing detection. The CPPS settings used by the algorithm are identical to Murton et al. (2020), Watts et al. (2017), and Brockmann-Bauser et al. (2021). Power cepstrogram is obtained using 60Hz pitch floor; maximum frequency of 5000Hz; time step of 2ms; and pre-emphasis at 50Hz. Further, this algorithm extracts out CPPS by setting subtraction of spectral tilt before smoothening to "no"; time averaging window to 0.01 seconds; quefrency averaging window to 0.001 seconds; peak search pitch range to 330Hz; tolerance of 0.05; interpolation as "parabolic"; tilt line quefrency range set from 0.001 seconds to 0 seconds; line type as "straight"; and fit method as "robust".

Statistical Analysis

All the statistical analysis were carried out in SPSS statistical analysis software (Version 27).

Descriptive statistics were calculated for the following parameters separately for male and female group: F0, Jitter, Shimmer, HNR, CPPS (with AVD); and CPPS (without AVD).

Shapiro Wilk's test of Normality was run for the entire data set (separately for each stimuli and algorithm). The entire data set was found to be normally distributed thus indicating use of parametric tests for the inferential statistics.

Paired sample t–test was run to infer significance of difference between the CPPS values obtained with AVD and without AVD for all the recorded stimuli for both male and female subgroups.

Repeated measures ANOVA along with post-hoc Boneferroni test was carried out to infer the significance of interstimulus difference separately for the CPPS values computed with AVD and those computed without AVD for both male and female subgroups.

Independent t–test was run to infer gender comparison for CPPS values computed across all five stimuli and both the algorithms.

CHAPTER 4

Results

The present study aimed to establish CPPS normative for different sample stimuli, i.e., sustained vowel phonation samples (/a/, /i/, and /u/), a mixed-sentence stimulus (that contains both voiceless and voiced phonemes) and a voiced-sentence (that contains only voiced phonemes) in young adult Hindi speaking normophonic individuals with and without voicing detection algorithm.

The results of the present study are discussed under the following subheadings:

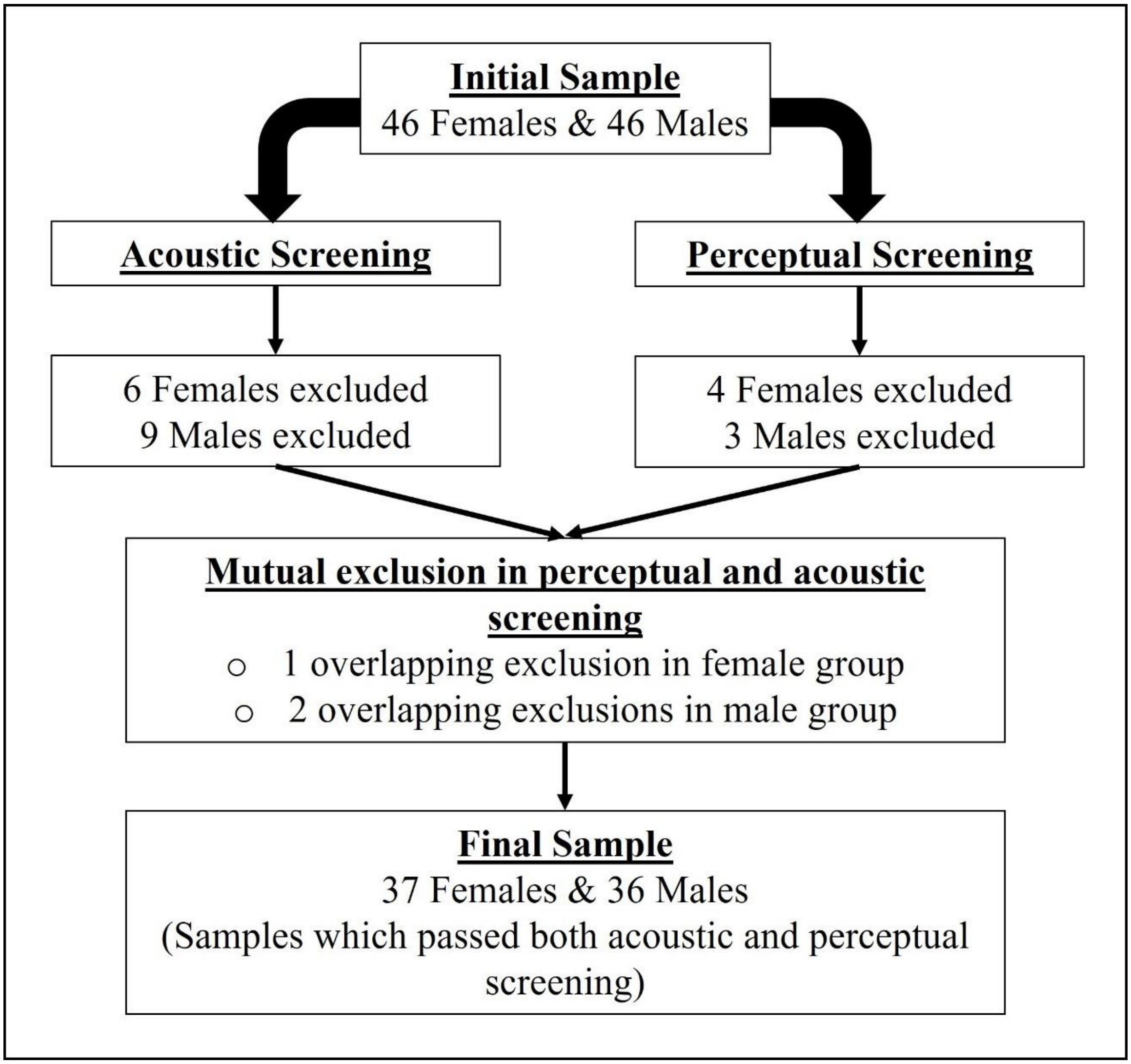
1. Perceptual and acoustic screening
2. Descriptive statistics
3. Gender comparison
4. Effect of using AVD
5. Interstimulus comparison

Perceptual and Acoustic Screening

Four females and three males were excluded from the study based on the perceptual screening criteria and six females (one overlapping with perceptual exclusion) and nine males (two overlapping with perceptual exclusion) were excluded based on the acoustic screening criteria. Final sample considered for CPPS analysis consisted of 37 females and 36 males (Figure 4.1).

Figure 4.1

Acoustic and Perceptual Screening Results



Descriptive Statistics

Mean and standard deviation values of CPPS computed without AVD are given in Table 4.1 and Figure 4.2. Mean and standard deviation values of CPPS computed with AVD are given in Table 4.1 and Figure 4.3. In general, vowels yielded higher CPPS values compared to the sentences except for vowel /u/, which yielded lower CPPS values than sentences in some cases in females.

Table 4.1

CPPS Values (dB) Computed Without AVD

Parameter	Females (n = 37)		Males (n = 36)	
	Mean	SD	Mean	SD
Vowel /a/	15.13	1.59	17.74	1.36
Vowel /i/	13.28	1.77	16.40	2.03
Vowel /u/	11.63	1.05	14.53	1.68
Mixed-Sentence	11.23	1.22	12.21	1.09
Voiced-Sentence	12.50	1.58	13.09	1.20

Figure 4.2

CPPS Values (dB) Computed Without AVD

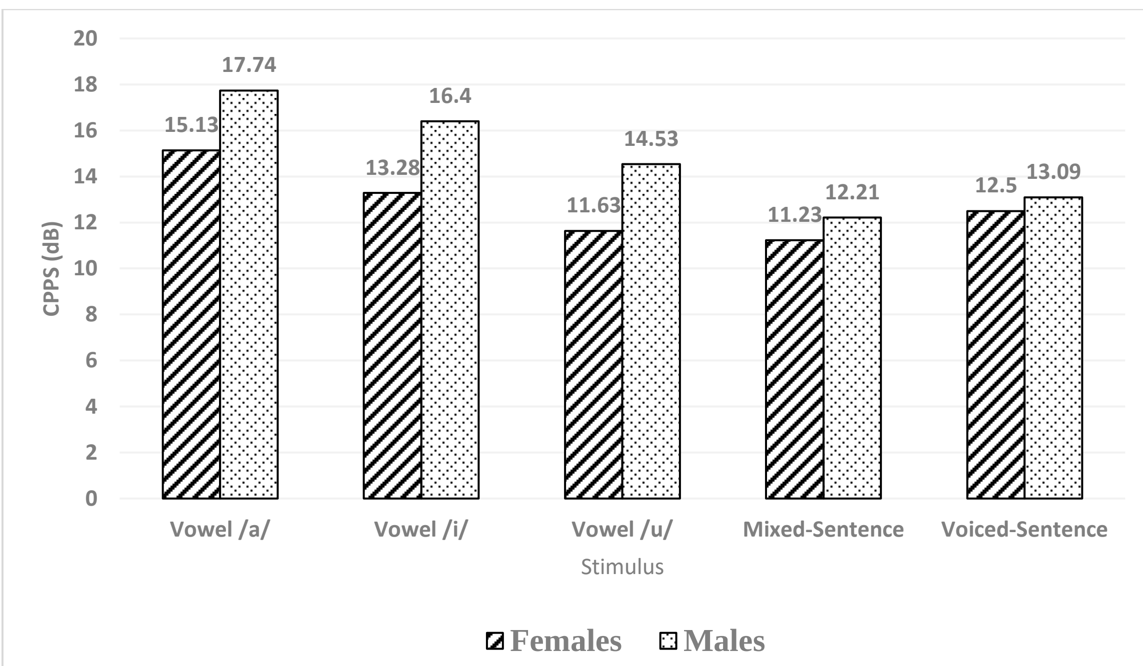


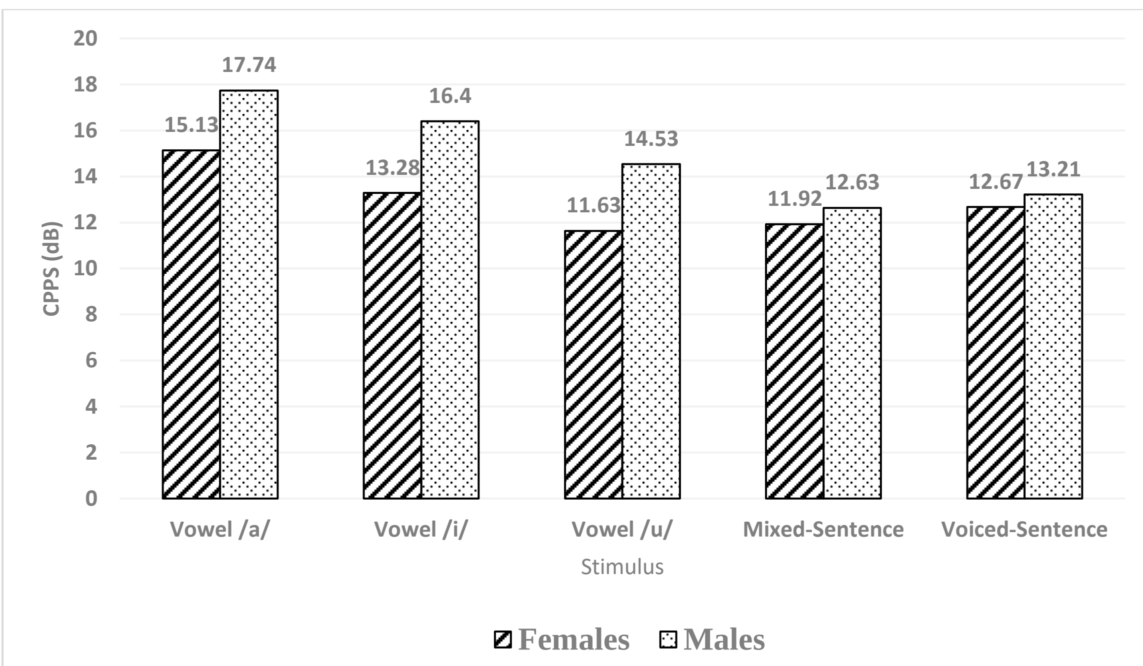
Table 4.2

CPPS Values (dB) Computed With AVD

Parameter	Females (n = 37)		Males (n = 36)	
	Mean	SD	Mean	SD
Vowel /a/	15.13	1.59	17.74	1.36
Vowel /i/	13.28	1.77	16.40	2.03
Vowel /u/	11.63	1.05	14.53	1.68
Mixed-Sentence	11.92	1.25	12.63	1.18
Voiced-Sentence	12.67	1.54	13.21	1.16

Figure 4.3

CPPS Values (dB) Computed With AVD



Mean and standard deviation values of other acoustic measures computed on sustained vowel phonation /a/, like F0, jitter, shimmer, and HNR are mentioned in Table 4.3.

Table 4.3

Acoustic Parameters Computed on Sustained Vowel Phonation /a/

Parameter	Females (n = 37)		Males (n = 36)	
	Mean	SD	Mean	SD
F0 (Hz)	214.45	19.89	128.84	13.65
Jitter local (%)	0.32	0.09	0.33	0.08
Shimmer local (%)	2.75	0.86	2.55	0.66
HNR (dB)	22.86	1.74	21.53	1.93

Gender Comparison

CPPS values computed for all stimulus across both the algorithms were subjected to gender comparison using independent t – test (Table 4.4). A significant difference ($p < 0.05$) in CPPS values was found between both gender groups for /a/, /i/, /u/, and mixed sentence for both the algorithm types. For all the four type of stimulus mentioned before, males were found to have significantly higher CPPS values for both algorithms. However, no significant gender difference ($p > 0.05$) was found for voiced-sentence for both the algorithm types. Graphical representation of gender comparison for both the algorithm types is presented in Figure 4.4 and Figure 4.5.

Table 4.4

Comparison of CPPS Values Obtained Across Stimuli and Algorithms Between Males and Females

Stimulus	Algorithm	t–value	p-value
Vowel /a/	Without AVD	7.535	<0.001*
	With AVD	7.521	<0.001*
Vowel /i/	Without AVD	6.997	<0.001*
	With AVD	6.996	<0.001*
Vowel /u/	Without AVD	8.896	<0.001*
	With AVD	8.891	<0.001*
Mixed-Sentence	Without AVD	3.617	<0.001*
	With AVD	2.516	0.014*
Voiced-Sentence	Without AVD	1.802	0.076
	With AVD	1.700	0.094

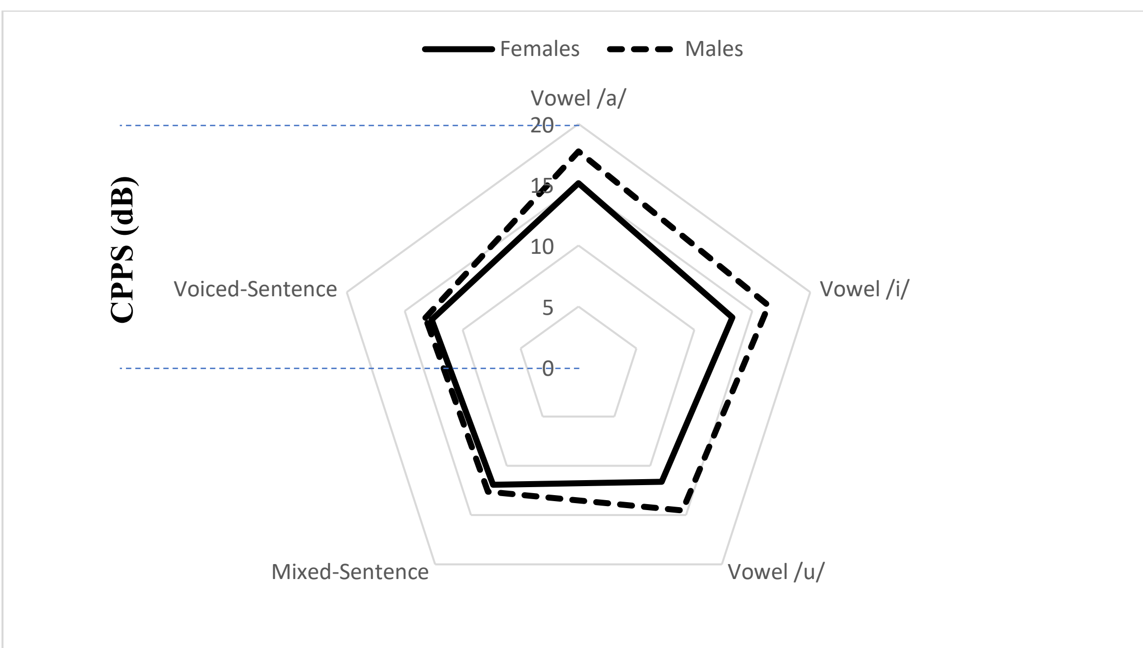
* p-value < 0.05 is indicative of significant difference

Figure 4.4

Gender Comparison for CPPS Computed Without AVD

Figure 4.5

Gender Comparison for CPPS Computed With AVD



Effect of Using AVD

Table 4.5 presents a comparison between CPPS values obtained with AVD and those obtained without using AVD. Results of t – test show that there was no significant difference ($p > 0.05$) obtained on using AVD for either of the vowels. Whereas, CPPS values computed with AVD were significantly higher ($p < 0.05$) than the CPPS values computed without AVD for both the sentences, i.e., the mixed-sentence and the voiced-sentence. However the effect of AVD was smaller on the voiced-sentence in comparison to the mixed-sentence.

Table 4.5

Comparison Between CPPS Values Computed Without AVD Vs With AVD

Stimulus	Gender	t–value	p-value	Effect Size (Cohen's d)
Vowel /a/	Males	1.339	0.189	0.02
	Females	0.433	0.668	0.01
Vowel /i/	Males	0.269	0.790	0.01
	Females	0.239	0.812	0.01
Vowel /u/	Males	0.723	0.475	0.01
	Females	0.502	0.619	0.01
Mixed-Sentence	Males	10.565	<0.001*	0.24
	Females	13.053	<0.001*	0.32
Voiced-Sentence	Males	3.749	<0.001*	0.20
	Females	4.764	<0.001*	0.22

* p-value < 0.05 is indicative of significant difference

Interstimulus Comparison

CPPS values obtained using different stimulus were compared with each other separately for those computed with AVD and those computed without using AVD. The vowels /a/, /i/, and /u/ were compared with each other and the sentences, i.e., the mixed-sentence and the voiced-sentence were compared with each other. Table 4.7 presents this interstimulus comparison when CPPS was computed without AVD and Table 4.8 presents this interstimulus comparison when CPPS was computed with AVD. Repeated measures ANOVA (Table 4.6) along with post hoc Boneferroni test revealed significant difference ($p < 0.05$) in all the interstimulus comparisons for all the pairs.

CPPS values were found to be highest for vowel /a/, followed by /i/ and then /u/. Also, CPPS values for the voiced-sentence were found to be higher than CPPS values for the mixed-sentence for both with AVD and without AVD algorithms. The results were consistent for both male and female subgroups.

Table 4.6

Interstimulus Comparison Calculated Using Repeated Measure ANOVA

Algorithm	Gender	df	F	p-value
CPPS computation without AVD	Males	4	121.135	<0.001*
	Females	4	88.342	<0.001*
CPPS computation with AVD	Males	4	108.792	<0.001*
	Females	4	69.770	<0.001*

* p-value < 0.05 is indicative of significant difference

Table 4.7

Interstimulus Comparison in CPPS Computed Without AVD

Stimulus Compared	Gender	Mean Difference	p-value
Vowel /a/ – Vowel /i/	Males	1.334	<0.001*
	Females	1.850	<0.001*
Vowel /a/ – Vowel /u/	Males	3.206	<0.001*
	Females	3.501	<0.001*
Vowel /i/ – Vowel /u/	Males	1.862	<0.001*
	Females	1.651	<0.001*
Mixed-Sentence – Voiced-Sentence	Males	0.876	<0.001*
	Females	1.267	<0.001*

* p-value < 0.05 is indicative of significant difference

Table 4.8

Interstimulus Comparison in CPPS Computed With AVD

Stimulus Compared	Gender	Mean Difference	p-value
Vowel /a/ – Vowel /i/	Males	1.341	<0.001*
	Females	1.851	<0.001*
Vowel /a/ – Vowel /u/	Males	3.203	<0.001*
	Females	3.501	<0.001*
Vowel /i/ – Vowel /u/	Males	1.863	<0.001*
	Females	1.650	<0.001*
Mixed-Sentence – Voiced-Sentence	Males	0.582	<0.001*
	Females	0.754	<0.001*

* p-value < 0.05 is indicative of significant difference

CHAPTER 5

Discussion

The present study aimed to establish CPPS normative for different sample stimuli, i.e., sustained vowel phonation samples, a mixed-sentence stimulus and a voiced-sentence in young adult Hindi speaking normophonic individuals with and without voicing detection algorithm. Other objectives of the study included:

1. To study the effect of Automatic Voicing Detection (AVD) algorithm on CPPS when obtained using different kind of stimulus.
2. To compare CPPS values obtained for males and females.
3. To compare the CPPS values obtained using different stimulus.

Effect of AVD on CPPS

The AVD algorithm removes all the voiceless parts in a recorded sample, i.e., silent gaps, pauses between words and/ or sentences, voiceless phonemes and voiceless parts in the phonemes, e.g., the closure duration in voiceless plosives etc. These voiceless parts otherwise cause to CPPS to reduce which is prevented by the AVD algorithm.

The results of this study showed that use of AVD had no effect on CPPS values computed on vowels. This is supported by the fact that vowels are entirely voiced productions and thus have no voiceless segments in essence. Thus, AVD algorithm won't make any difference because of absence of voiceless segments.

Use of AVD caused CPPS values to increase for both the sentences, i.e., the mixed-sentence as well as the voiced-sentence. In the case of mixed-sentence, AVD was expected to cause an increase in CPPS due to deletion of voiceless phonemes in the utterance. But ideally, in the case of voiced-sentence, this difference wasn't expected as all the phonemes were voiced. Visual comparison of the spectrograms in Praat software of the “voiced-sentence” samples analyzed without AVD and the samples in which only the voiced-segments were extracted by the AVD algorithm suggested that the change in CPPS values was due to deletion of gaps present between words within the sentence. This gap varied from one speaker to other depending on their rate of speech. Several measures were taken the reduce this effect. Sentences were framed in a way to avoid any natural punctuations within the sentence. Also, participants were familiarized with the sentences and were instructed in a way to minimize any gaps within the sentences (refer to methods section).

It is noteworthy that mean difference of CPPS values caused by AVD in the mixed-sentence (Females: 0.69dB; Males: 0.42 dB) was nearly three to four times to that in the voiced-sentence (Females: 0.17 dB; Males: 0.12 dB). Thus the effect of CPPS in much smaller in the case of voiced-sentence in comparison to mixed-sentence.

Also efficacy of AVD can be questioned on a finding that the voiced-sentence yielded significantly higher CPPS values than the mixed-sentence, which shouldn’t have been the case if at all the AVD was efficient enough to filter out all the voiceless segments. On the other hand this difference can also be accounted to inter-text variability.

Effect of Gender on CPPS

Results of this study have showed that male samples yielded significantly higher CPPS values than female samples for all the stimuli except for the voiced-sentence in which the difference was insignificant. The results were consistent for both the algorithms that is with AVD and without AVD. Similar findings were reported by most of the other studies including Monnappa and Balasubramanium (2015), Hasanvand et al. (2017), Mendes et al. (2023), and Spazzapan et al. (2022), and Sujitha and Pebbili (2022).

This gender difference in CPPS values could be because of inherent differences in anatomical and physiological characteristics of vocal folds, including vocal fold mass, vocal fold contact quotient, etc. (Buckley et al., 2023). Females tend to show lower value of vocal fold contact quotient (Holmberg et al., 1988; Tsutsumi et al., 2017), higher asymmetry quotient, and decreased maximum area declination rate (Patel et al., 2014) than males. These changes contribute to the variations in the acoustic source spectrum which might be contributing to the reduced CPPS values in females. Alongside, a longer closed phase in males contributes to a lower spectral tilt, that indicates more energy in the higher frequencies that in turn might be leading to increased CPPS values in males (Fraile & Godino-Llorente, 2014).

Effect of Stimulus on CPPS

Amongst all the vowels, /a/ is found to have the highest CPPS value, followed by /i/, and /u/. Similar results have been found in other studies as well. Awan et al. (2012) found out that low vowels [/ɑ/ and /æ/] yielded greater CPPS value vs high vowels [/i/ and /u/]. The rationale given by them is that an open vowel like /a/ is spoken with a wider open mouth in comparison to closed vowels like /u/ and /i/ which involve narrower mouth opening. This wider mouth opening leads to higher intensity being radiated out of the mouth in low vowels like /a/; and CPPS is directly corelated with intensity (Awan et al., 2012), thus yielding higher CPPS values for /a/ in comparison to /i/ and /u/. Similar results were also obtained by Sujitha and Pebbili (2022) and Buckley et al. (2023).

With respect to reading samples, voiced-sentence yielded higher CPPS value than mixed-sentence with both algorithms and across both the gender groups. This finding is supported by the same fact that tells why AVD yields higher CPPS value, i.e., voiceless segments in a mixed-sentence leads to a reduction in the CPPS value which isn't the scenario in the case of voiced-sentence on which there are only minimal voiceless segments in form of few gaps between the words in the sentence.

Should AVD be Used?

Automatic Voicing Detection (AVD) presents with the major advantage of removal of unvoiced segments in an utterance that might otherwise lead to reduction in CPPS values leading to increased false positives. On the other hand efficacy of AVD has been questioned in various studies. Findings of a study done by Awan et al. (2010) suggest ineffective detection and filtering out of voiceless segments by the AVD algorithm of ADSV. Murton et al. (2020) found the CPPS value to be higher for the all-voiced sentence in comparison to other sentences used from CAPE – V with both Praat and ADSV (given the fact that ADSV uses automatic voicing detection).

Alongside AVD also presents a logical risk of increasing false negatives. In cases of voice samples of dysphonic individuals presenting with voice breaks, the AVD might wrongly filter out the intended voiced segments which became voiceless because of the dysphonic voice breaks thus leading to an increased CPPS value.

Thus use of AVD can actually lead to increased false negative rate when used with dysphonic individuals irrespective of the type of stimulus used. Whereas, use of AVD can lead to increased false positive rates when a continuous speech sample is under consideration.

CHAPTER 6

Summary and Conclusion

Acoustic voice analysis is an easy, cost effective, non-invasive, and instrumental voice assessment tool. Amongst all the acoustic parameters of voice quality which have been investigated so far, CPPS comes out to be one of the most reliable parameter of voice quality and has shown to be having highest correlation with voice quality across various studies. CPPS can be computed with or without using the Automatic Voicing Detection (AVD) algorithm and can be computed on various sample types including sustained vowel phonation and continuous speech samples.

Since continuous speech samples are also valid indicator of voice quality, they are also preferred along with sustained vowel phonation samples. Reading task is used as the most frequent way of continuous speech sampling as it prevents inter – text variability across recordings. All-voiced reading texts have been a choice for voice sampling because of reasons such as better correlation with perceptual dysphonia severity (Awan et al., 2010; Roy, 2010), lower inter-text variability (Kitayama et al., 2020), and an alternative to Automatic Voicing Detection (AVD). Use of AVD is less preferred because of efficacy issues, filtering out of voice breaks/ dysphonic segments leading to false negatives, and filtering out of natural pauses and breaks in normal samples leading to increased chances of false positives.

In the present study, 92 participants (46 females and 46 males) within the age range of (19 – 30 years) were recruited. Samples of all participants were subjected to perceptual and acoustic screening to ascertain normal quality of voice. Final screened sample consisted of 37 females and 36 males. All the screened samples were subjected to

CPPS analysis using Praat software. Mean CPPS values for /a/, /i/, /u/, mixed-sentence, and voiced-sentence were found out to be 15.13 dB, 13.28 dB, 11.63 dB, 11.23 dB, and 12.50 dB respectively in females and 17.74 dB, 16.40 dB, 14.53 dB, 12.21 dB, and 13.09 dB respectively in males. Use of AVD yielded higher CPPS values in comparison to when AVD wasn't used. Overall, male samples yielded higher CPPS values than female samples. Amongst sustained phonation samples, /a/ had highest CPPS values followed by /i/ and /u/ in both female and male subgroups. Similarly voiced-sentence yielded higher CPPS values in comparison to mixed-sentence in both female and male subgroups.

This study reveals CPPS normative for the young adult (19 – 30 years), Hindi speakers for various stimulus types and both the algorithm types, i.e., with AVD and without AVD. The established values can be used as a screening tool for the clinical purposes.

Given the advantages of using all voiced text and disadvantages of using AVD as mentioned before, use of the developed "voiced-sentence" is recommended for clinical voice analysis in Hindi speakers taking the established normative value as the reference value.

Limitations and Future Scope

- This study was conducted only on a limited age range, i.e., young adults (19 – 30 years). Thus similar future studies should be planned for the younger and older age groups.
- The sensitivity and specificity of the established CPPS normative should further be established by using a similar protocol on dysphonic individuals.
- Normal participants were not screened over endoscopic visualization.

CHAPTER 7

References

Awan, S. N., Giovinco, A., & Owens, J. (2012). Effects of vocal intensity and vowel type on cepstral analysis of voice. *Journal of Voice*, *26*(5), 670.e15-670.e20. https://doi.org/10.1016/j.jvoice.2011.12.001

Awan, S. N., & Roy, N. (2006). Toward the development of an objective index of dysphonia severity: A four-factor acoustic model. *Clinical Linguistics & Phonetics*, *20*(1), 35–49. https://doi.org/10.1080/02699200400008353

Awan, S. N., Roy, N., & Dromey, C. (2009). Estimating dysphonia severity in continuous speech: Application of a multi-parameter spectral-cepstral model estimating dysphonia severity in continuous speech. *Clinical Linguistics and Phonetics*, *23*(11), 825–841. https://doi.org/10.3109/02699200903242988

Awan, S. N., Roy, N., Jetté, M. E., Meltzner, G. S., & Hillman, R. E. (2010). Quantifying dysphonia severity using a spectral/cepstral-based acoustic index: Comparisons with auditory-perceptual judgements from the CAPE-V. *Clinical Linguistics and Phonetics*, *24*(9), 742–758. https://doi.org/10.3109/02699206.2010.492446

Balasubramanium, R. K., Shastry, A., Singh, M., & Bhat, J. S. (2015). Cepstral Characteristics of Voice in Indian Female Classical Carnatic Singers. *Journal of Voice*, *29*(6), 693–695. https://doi.org/10.1016/j.jvoice.2015.01.002

Barsties v. Latoszek, B., Mayer, J., Watts, C. R., & Lehnert, B. (2023). Advances in Clinical Voice Quality Analysis with VOXplot. *Journal of Clinical Medicine*, *12*(14). https://doi.org/10.3390/jcm12144644

Bhattacharyya, N. (2014). The prevalence of voice problems among adults in the United States. *The Laryngoscope*, *124*(10), 2359–2362. https://doi.org/https://doi.org/10.1002/lary.24740

Bhattacharyya, N. (2015). The prevalence of pediatric voice and swallowing problems in the United States. *The Laryngoscope*, *125*(3), 746–750. https://doi.org/https://doi.org/10.1002/lary.24931

Brinca, L. F., Batista, A. P. F., Tavares, A. I., Gonçalves, I. C., & Moreno, M. L. (2014). Use of cepstral analyses for differentiating normal from dysphonic voices: A comparative study of connected speech versus sustained vowel in European portuguese female speakers. In *Journal of Voice* (Vol. 28, Issue 3, pp. 282–286). Mosby Inc. https://doi.org/10.1016/j.jvoice.2013.10.001

Brockmann-Bauser, M., Van Stan, J. H., Carvalho Sampaio, M., Bohlender, J. E., Hillman, R. E., & Mehta, D. D. (2021). Effects of Vocal Intensity and Fundamental Frequency on Cepstral Peak Prominence in Patients with Voice Disorders and Vocally Healthy Controls. *Journal of Voice*, *35*(3), 411–417. https://doi.org/10.1016/j.jvoice.2019.11.015

Buckley, D. P., Abur, D., & Stepp, C. E. (2023). Normative Values of Cepstral Peak Prominence Measures in Typical Speakers by Sex, Speech Stimuli, and Software Type Across the Life Span. *American Journal of Speech-Language Pathology*, 1–13. https://doi.org/10.1044/2023_ajslp-22-00264

Castellana, A., Carullo, A., Corbellini, S., & Astolfi, A. (2018). Discriminating Pathological Voice from Healthy Voice Using Cepstral Peak Prominence Smoothed Distribution in Sustained Vowel. *IEEE Transactions on Instrumentation and Measurement*, *67*(3), 646–654. https://doi.org/10.1109/TIM.2017.2781958

Fraile, R., & Godino-Llorente, J. I. (2014). Cepstral peak prominence: A comprehensive analysis. *Biomedical Signal Processing and Control*, *14*, 42–54. https://doi.org/https://doi.org/10.1016/j.bspc.2014.07.001

Hasanvand, A., Salehi, A., & Ebrahimipour, M. (2017a). A Cepstral Analysis of Normal and Pathologic Voice Qualities in Iranian Adults: A Comparative Study. *Journal of Voice*, *31*(4), 508.e17-508.e23. https://doi.org/10.1016/j.jvoice.2016.10.017

Hassan, E. M., Abdel Hady, A. F., Shohdi, S. S., Eldessouky, H. M., & Din, M. H. B. (2021). Assessment of dysphonia: cepstral analysis versus conventional acoustic analysis. *Logopedics Phoniatrics Vocology*, *46*(3), 99–109. https://doi.org/10.1080/14015439.2020.1767202

Heller Murray, E. S., Chao, A., & Colletti, L. (2022). A Practical Guide to Calculating Cepstral Peak Prominence in Praat. In *Journal of Voice*. Elsevier Inc. https://doi.org/10.1016/j.jvoice.2022.09.002

Heman-Ackah, Y. D., Michael, D. D., Baroody, M. M., Ostrowski, R., Hillenbrand, J., Heuer, R. J., Horman, M., & Sataloff, R. T. (2003). Cepstral Peak Prominence: A More Reliable Measure of Dysphonia. *Annals of Otology, Rhinology & Laryngology*, *112*(4), 324–333. https://doi.org/10.1177/000348940311200406

Heman-Ackah, Y. D., Sataloff, R. T., Laureyns, G., Lurie, D., Michael, D. D., Heuer, R., Rubin, A., Eller, R., Chandran, S., Abaza, M., Lyons, K., Divi, V., Lott, J., Johnson, J., & Hillenbrand, J. (2014). Quantifying the cepstral peak prominence, a measure of dysphonia. *Journal of Voice*, *28*(6), 783–788. https://doi.org/10.1016/j.jvoice.2014.05.005

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

Hippargekar, P., Bhise, S., Kothule, S., & Shelke, S. (2022). Acoustic Voice Analysis of Normal and Pathological Voices in Indian Population Using Praat Software. *Indian Journal of Otolaryngology and Head and Neck Surgery*, *74*, 5069–5074. https://doi.org/10.1007/s12070-021-02757-9

Holmberg, E. B., Hillman, R. E., & Perkell, J. S. (1988). Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. *The Journal of the Acoustical Society of America*, *84*(2), 511–529. https://doi.org/10.1121/1.396829

Kim, G. H., Bae, I. H., Park, H. J., & Lee, Y. W. (2021). Comparison of Cepstral Analysis Based on Voiced-Segment Extraction and Voice Tasks for Discriminating Dysphonic and Normophonic Korean Speakers. *Journal of Voice*, *35*(2), 328.e11-328.e22. https://doi.org/10.1016/j.jvoice.2019.09.009

Kim, G. H., Lim, D. W., Kim, J. W., Park, H. J., & Lee, Y. W. (2022). A Cepstral Analysis of Pathological Voice Quality in the Korean Population using Praat. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2022.10.011

Kitayama, I., Hosokawa, K., Iwahashi, T., Iwahashi, M., Iwaki, S., Kato, C., Yoshida, M., Umatani, M., Matsushiro, N., Ogawa, M., & Inohara, H. (2020). Intertext Variability of Smoothed Cepstral Peak Prominence, Methods to Control It, and Its Diagnostic Properties. *Journal of Voice*, *34*(3), 305–319. https://doi.org/10.1016/j.jvoice.2018.09.021

Lechien, J. R., Geneid, A., Bohlender, J. E., Cantarella, G., Avellaneda, J. C., Desuter, G., Sjogren, E. V, Finck, C., Hans, S., Hess, M., Oguz, H., Remacle, M. J., Schneider-Stickler, B., Tedla, M., Schindler, A., Vilaseca, I., Zabrodsky, M., Dikkers, F. G., & Crevier-Buchman, L. (2023). Consensus for voice quality assessment in clinical practice: guidelines of the European Laryngological Society and Union of the European Phoniatricians. *European Archives of Oto-Rhino-Laryngology : Official Journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : Affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery*. https://doi.org/10.1007/s00405-023-08211-6

Liang-Hui, W., Thanh-Nhan, D., Fei-Chun, C., Thi-Lien, T., Wen-Chao, H., & Li-Wei, C. (2023). Prevalence of Voice Disorders in Older Adults: A Systematic Review and Meta-Analysis. *American Journal of Speech-Language Pathology*, *32*(4), 1758–1769. https://doi.org/10.1044/2023_AJSLP-22-00393

Maryn, Y., Roy, N., De Bodt, M., Van Cauwenberge, P., & Corthals, P. (2009). Acoustic measurement of overall voice quality: A meta-analysis. *The Journal of the Acoustical Society of America*, *126*(5), 2619–2634. https://doi.org/10.1121/1.3224706

Mendes, A. P., Nunes, N., Ibrahim, S., Coelho, A. C., & Francisco, M. A. (2023). Cepstral Measures in the Fado Voice: Gender, Age and Phonatory Tasks. *Journal of Voice*, *37*(1), 9–16. https://doi.org/10.1016/j.jvoice.2020.09.010

Menon, U. K., Raj, M., Antony, L., Soman, S., & Bhaskaran, R. (2021). Prevalence of Voice Disorders in School Teachers in a District in South India. *Journal of Voice*, *35*(1), 1–8. https://doi.org/10.1016/j.jvoice.2019.07.005

Monnappa, D., & Balasubramanium, R. (2015). Cepstral analysis of voice in healthy aged individuals. *Journal of Laryngology and Voice*, *5*(2), 34. https://doi.org/10.4103/2230-9748.183963

Murton, O., Hillman, R., & Mehta, D. (2020). Cepstral peak prominence values for clinical voice evaluation. *American Journal of Speech-Language Pathology*, *29*(3), 1596–1607. https://doi.org/10.1044/2020_AJSLP-20-00001

Núñez-Batalla, F., Cartón-Corona, N., Vasile, G., García-Cabo, P., Fernández-Vañes, L., & Luis Llorente-Pendás, J. (2019). Validation of the Measures of Cepstral Peak Prominence as a Measure of Dysphonia Severity in Spanish-Speaking Subjects. In *Acta Otorrinolaringol Esp* (Vol. 70, Issue 4). www.elsevier.es/otorrino

Oliveira, P., Ribeiro, V. V., Constantini, A. C., Cavalcante, M. E. de O. B., Sousa, M. dos S., & da Silva, K. (2023). Prevalence of Work-Related Voice Disorders in Voice Professionals: Systematic Review and Meta-Analysis. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2022.07.030

Oliveira Santos, A., Godoy, J., Silverio, K., & Brasolotto, A. (2021). Vocal Changes of Men and Women from Different Age Decades: An Analysis from 30 Years of Age. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2021.06.003

Patel, R. R., Awan, S. N., Kraemer, J. B., Courey, M., Deliyski, D., Eadie, T., Paul, D., Švec, J. G., & Hillman, R. (2018). Recommended Protocols for Instrumental Assessment of Voice: American Speech-Language-Hearing Association Expert Panel to Develop a Protocol for Instrumental Assessment of Vocal Function. *American Journal of Speech-Language Pathology*, *27*(3), 887–905. https://doi.org/10.1044/2018_AJSLP-17-0009

Patel, R. R., Dubrovskiy, D., & Döllinger, M. (2014). Measurement of glottal cycle characteristics between children and adults: Physiological variations. *Journal of Voice*, *28*(4), 476–486. https://doi.org/10.1016/j.jvoice.2013.12.010

Phadke, K. V., Laukkanen, A. M., Ilomäki, I., Kankare, E., Geneid, A., & Švec, J. G. (2020). Cepstral and Perceptual Investigations in Female Teachers With Functionally Healthy Voice. *Journal of Voice*, *34*(3), 485.e33-485.e43. https://doi.org/10.1016/j.jvoice.2018.09.010

Procter, T., & Joshi, A. (2022). Cultural Competency in Voice Evaluation: Considerations of Normative Standards for Socio-linguistically Diverse Voices. *Journal of Voice*, *36*(6), 793–801. https://doi.org/10.1016/j.jvoice.2020.09.025

Reddy, M. B., Kishore, G. P., & Jose, J. B. (2014). *Smoothened Cepstral Peak Prominence in untrained choral singer and non-singers*.

Roy, N. (2010). Differential diagnosis of muscle tension dysphonia and spasmodic dysphonia. *Current Opinion in Otolaryngology & Head and Neck Surgery*, *18*(3). https://journals.lww.com/co-otolaryngology/fulltext/2010/06000/differential_diagnosis_of_muscle_tension_dysphonia.5.aspx

Roy, N., Merrill, R. M., Gray, S. D., & Smith, E. M. (2005). Voice Disorders in the General Population: Prevalence, Risk Factors, and Occupational Impact. *The Laryngoscope*, *115*(11), 1988–1995. https://doi.org/https://doi.org/10.1097/01.mlg.0000179174.32345.41

Saeedi, S., Aghajanzadeh, M., Khoddami, S. M., Dabirmoghaddam, P., & Jalaie, S. (2022). The Validity of Cepstral Analysis to Distinguish Dysphonia Between Different Levels of Perceptual Dysphonia in the Persian Vocal Tasks. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2022.04.008

Sampaio, M., Masson, M. L. V., Soares, M. F. de P., Bohlender, J. E., & Brockmann-Bauser, M. (2020). Effects of fundamental frequency, vocal intensity, sample duration, and vowel context in cepstral and spectral measures of dysphonic voices. *Journal of Speech, Language, and Hearing Research*, *63*(5), 1326–1339. https://doi.org/10.1044/2020_JSLHR-19-00049

Sataloff, R. T. (2017). *Clinical Assessment of Voice*. Plural Publishing, Incorporated. https://books.google.co.in/books?id=kCZIrgEACAAJ

Sauder, C., Bretl, M., & Eadie, T. (2017). Predicting Voice Disorder Status from Smoothed Measures of Cepstral Peak Prominence Using Praat and Analysis of Dysphonia in Speech and Voice (ADSV). *Journal of Voice*, *31*(5), 557–566. https://doi.org/10.1016/j.jvoice.2017.01.006

Sheyona, V., & Devadas, U. (2022). The Prevalence and Impact of Voice Problems in Nonprofessional Voice Users: Preliminary Findings. *Journal of Voice*, *36*(3), 383–388. https://doi.org/10.1016/j.jvoice.2020.06.010

Soni, A., Jotdar, A., Kapoor, D., & Singh, A. (2023). Acoustic analysis of normal voice patterns in North Indian population. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2023.07.018

Spazzapan, E. A., Marino, V. C. de C., & Fabbron, E. M. G. (2022). Smoothed Cepstral Peak Analysis of Brazilian Children and Adolescents Speakers. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2022.02.002

Sujitha, P. S., & Pebbili, G. K. (2022). Cepstral Analysis of Voice in Young Adults. *Journal of Voice*, *36*(1), 43–49. https://doi.org/10.1016/j.jvoice.2020.03.010

Svec, J. G., & Granqvist, S. (2010). Guidelines for Selecting Microphones for Human Voice Production Research. *American Journal of Speech-Language Pathology*, *19*(4), 356–368. https://doi.org/10.1044/1058-0360(2010/09-0091)

Titze, I. R. (1995). *Workshop on Acoustic Voice Analysis SUMMARY STATEMENT*.

Tsutsumi, M., Isotani, S., Pimenta, R. A., Dajer, M. E., Hachiya, A., Tsuji, D. H., Tayama, N., Yokonishi, H., Imagawa, H., Yamauchi, A., Takano, S., Sakakibara, K. I., & Montagnoli, A. N. (2017). High-speed Videolaryngoscopy: Quantitative Parameters of Glottal Area Waveforms and High-speed Kymography in Healthy Individuals. *Journal of Voice*, *31*(3), 282–290. https://doi.org/10.1016/j.jvoice.2016.09.026

Watts, C. R., & Awan, S. N. (2015). An examination of variations in the cepstral spectral index of dysphonia across a single breath group in connected speech. *Journal of Voice*, *29*(1), 26–34. https://doi.org/10.1016/j.jvoice.2014.04.012

Watts, C. R., Awan, S. N., & Maryn, Y. (2017). A Comparison of Cepstral Peak Prominence Measures from Two Acoustic Analysis Programs. *Journal of Voice*, *31*(3), 387.e1-387.e10. https://doi.org/10.1016/j.jvoice.2016.09.012