

Research Article

Normative Values of Cepstral Peak Prominence Measures in Typical Speakers by Sex, Speech Stimuli, and Software Type Across the Life Span

Daniel P. Buckley,^{a,b}  Defne Abur,^{a,c,d}  and Cara E. Stepp^{a,b,e} 

^aDepartment of Speech, Language, and Hearing Sciences, Boston University, MA ^bDepartment of Otolaryngology – Head and Neck Surgery, Boston University School of Medicine, MA ^cDepartment of Computational Linguistics, University of Groningen, the Netherlands ^dResearch School of Behavioral and Cognitive Neurosciences, University of Groningen, the Netherlands ^eDepartment of Biomedical Engineering, Boston University, MA

ARTICLE INFO

Article History:

Received August 22, 2022

Revision received December 15, 2022

Accepted March 16, 2023

Editor-in-Chief: Katherine C. Hustad

Editor: Rita R. Patel

https://doi.org/10.1044/2023_AJSLP-22-00264

ABSTRACT

Purpose: The purpose of this study was to determine normative values for cepstral peak prominence measures across the life span as a function of sex using clinically relevant stimuli (/a/, /i/, and two sentences of The Rainbow Passage) and two commonly used software types: Praat (Version 6.0.50) and Analysis of Dysphonia in Speech and Voice (ADSV).

Method: One hundred fifty speakers (75 males, 75 females; evenly distributed into three age groups) without voice disorders aged 18–91 years were recorded via headset microphone in a sound-treated booth. Cepstral measures were analyzed using common analysis methods in Praat and ADSV by sex, stimuli, and software type. Kruskal–Wallis tests and post hoc Mood’s Median tests for significant factors were performed on cepstral measures to assess the effects of age group, sex, stimuli, and software type.

Results: The results revealed statistically significant effects of sex, stimuli, and software type on cepstral measures, but no statistical effect of age group on cepstral values. Females had lower average cepstral values compared to males. Across stimuli, the highest average cepstral measure was found for sustained /a/, followed by sustained /i/, and then of the two sentences of The Rainbow Passage. Average cepstral measures in Praat were higher than those from ADSV.

Conclusions: The current work did not find a statistical effect of age group on cepstral values; thus, normative cepstral values were reported by sex, stimuli, and software type. Future work should examine the applicability of these normative values for discriminating speakers with and without voice disorders.

It is estimated that approximately 30% of individuals will develop a voice disorder during their lifetime (Roy et al., 2005). The act of voice production includes the contributions from multiple subsystems: respiration, phonation, and resonance. Voice disorders can be broadly

classified into organic or functional disorders. Organic disorders present with physiological changes that result from alternations in voice subsystems, and consist of structural (physical changes to the vocal mechanism) or neurogenic (any of numerous neurological disruptions affecting the vocal mechanism). Contrary to organic voice disorders, functional disorders include inadequate and disrupted phonation in the absence of anatomical or physical changes to the larynx. In response to this wide range of etiologies, voice evaluations conducted by speech-language pathologists (SLPs) consist of multiple components that aim to develop a detailed understanding of an individual’s voice disorder and vocal function. A typical voice evaluation

Correspondence to Daniel P. Buckley: buckleyd@bu.edu. **Disclosure:** Cara E. Stepp has received consulting fees from Altec, Inc./Delsys, Inc., companies focused on developing and commercializing technologies related to human movement. Stepp’s interests were reviewed and are managed by Boston University in accordance with their conflict of interest policies. The other authors have declared that no other competing interests existed at the time of publication.

consists of an interview to assess subjective symptoms, auditory-perceptual assessment of the patient's voice quality, endoscopic imaging to obtain direct visualization of voice production, aerodynamic measures of airflow and pressures related to voice and respiratory function, and an acoustic evaluation (Patel et al., 2018; Roy et al., 2013). Together, these measures provide the SLP with data that may assist with developing treatment targets or monitoring the response to intervention. These multiple components of the voice evaluation consist of objective and subjective components, which possess various strengths and limitations.

Voice Evaluation Methods for SLPs

One important component included in a comprehensive evaluation of voice is the auditory-perceptual evaluation. This involves one of multiple tools and methods that exist for documenting these impressions, including the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V; Kempster et al., 2009) and the grade, roughness, breathiness, asthenia, strain (Hirano, 1981) scale. Auditory-perceptual judgments have demonstrated varying degrees of reliability, with relatively higher reliability for the perceived overall severity of dysphonia, but conflicting and variable reliability for other individual components such as roughness and strain (Zraick et al., 2011). The subjective nature of these judgments allows for individual interpretation of potentially unquantifiable variables (i.e., auditory-perceptual impressions of voice quality that may not be reflected nor captured through objective, acoustic analysis) and relies on a listener's auditory perception, which can be argued as highly important for assessing the perceived dysphonia of unfamiliar listeners in a real communication environment. However, the varying degrees of reliability are a limitation.

Similar to the auditory-perceptual evaluation of voice, endoscopic imaging is a critical component of the voice evaluation that is also limited by subjectivity. An endoscope is used, either trans-orally or trans-nasally, to visualize the larynx and surrounding structures. This allows for visualization of the vocal fold tissue at rest, to assess for masses or vocal fold lesions, and during phonation, to assess the various vibratory and kinematic features of voice production. Laryngeal imaging is necessary for a diagnosis and is typically completed by a referring physician prior to the voice evaluation. However, the SLP may implement voice therapy techniques during endoscopy to assess for stimulability (i.e., improvements to voice function), via the use of various behavioral methods and interventions. This may, in turn, assist the SLP with selecting voice treatment targets. Despite these benefits, endoscopy relies on the subjective impression of the clinician. In contrast to auditory-perceptual and endoscopic assessments,

aerodynamic evaluation of voice relies strictly on objective measures. Aerodynamic measures are gleaned through collection of airflows and pressures via a pneumotachograph, with or without an intraoral pressure tube. Measures such as vital capacity, mean glottal airflow during voicing, and indirect estimates of subglottal pressure and phonation threshold pressure can be collected. These objective data are useful because they can elucidate various aspects of an individual's phonatory physiology, which may be useful for developing treatment targets (e.g., reducing increased subglottal pressure during phonation). However, measures such as subglottal pressure and phonation threshold pressure collected via clinical systems are indirect estimates, as their direct counterparts involve invasive methods such as a tracheal puncture (Isshiki, 1964).

Acoustic Evaluation of Voice

Acoustic evaluation of voice is comprised of various noninvasive, objective measurements that rely solely on an acoustic signal captured via a microphone. These measures evade the subjectivity of auditory-perceptual or visual judgments of endoscopy but are less invasive and require less expensive instrumentation than what is required for aerodynamic evaluation. Fundamental frequency and sound pressure level are acoustic measures that have robust relationships with the auditory-perceptual qualities of pitch and loudness, respectively (Fletcher, 1934; Jenkins, 1961). However, many individuals with voice disorders have decrements in their voice quality, which is less cleanly associated with any one acoustic measure. Historically, acoustic measures focused on voice quality in clinical voice evaluations for individuals with voice disorders consisted of time-based, acoustic perturbation measures, such as jitter (cycle-to-cycle frequency perturbation; Lieberman, 1961), shimmer (cycle-to-cycle amplitude perturbation; Kitajima & Gould, 1976), and harmonics-to-noise ratio (the ratio of periodic energy to nonperiodic energy in a voice signal; Awan & Frenkel, 1994). These measures have multiple drawbacks that limit their utility in the voice evaluation. First, many of these measures have weak, or conflicting, correlations to auditory-perceptual features (Heman-Ackah et al., 2003). Second, the algorithms for these acoustic measures rely on the accurate estimation of fundamental frequency, which can be problematic for dysphonic voices (Watts et al., 2017). Third, these measures are only valid for steady-state signals, which limit their use to sustained vowels (e.g., /a/, /i/; not running speech). This reduces their clinical utility as they may not accurately capture an individual's dysphonia in their day-to-day voice use (i.e., in typical conversation). In consideration of these known issues, the American Speech-Language-Hearing Association revised recommendations for the acoustic evaluation of

voice in 2018 to replace these earlier measures with a more universal measure of dysphonia: cepstral peak prominence (CPP; Patel et al., 2018).

CPP is an acoustic measure of “the overall level of noise in the vocal signal” (Patel et al., 2018). A Fourier transform of the power spectrum of the voice signal is performed to create the power “cepstrum.” Within this cepstrum, a cepstral peak is estimated that represents the periodic harmonic energy present from the original acoustic source spectrum. The CPP (the amplitude of this peak), relative to a regression line created through the overall cepstrum, is measured in decibels (Heman-Ackah et al., 2003, 2014; Oppenheim & Schaffer, 2004). CPP is not time based, nor does it rely on fundamental frequency across multiple cycles of vibration or cycle-to-cycle variations in amplitude. Hence, it is robust in dysphonic voices that are more prone to aperiodic content and inconsistent fundamental frequency. This is also supported by the strong relationship that has been reported between CPP and overall severity of dysphonia as perceived by listeners (Awan et al., 2010; Maryn et al., 2009). Furthermore, CPP can be estimated from running speech and is not limited to sustained productions of vowels, so it can be calculated from more ecologically valid speech samples. Although CPP is very promising, some barriers remain to the full implementation of CPP as a clinical measure for the evaluation of dysphonia.

Influential Factors on Cepstral Measures and Available Literature

Voice clinicians commonly use two different software packages to calculate CPP: Analysis of Dysphonia in Speech and Voice (ADSV; PENTAX Medical), and Praat software (Boersma & Weenink, 2011). When CPP is calculated in Praat, it is referred to as smoothed cepstral peak prominence (CPPS) due to a slight variation in the algorithm. Research by Watts et al. (2017) compared CPP values using ADSV and CPPS values in Praat, and found differences in the measures by software type (i.e., CPPS values in Praat were higher than CPP values in ADSV) using the same samples. However, the cepstral measures from the software types were strongly related. Thus, the utilization of either software (ADSV or Praat) is appropriate for calculating cepstral measures of voice. Most studies of CPP/CPPS use either ADSV or Praat, but typically do not use both software types, which creates an obstacle in determining normative values (the cepstral values that distinguish a typical voice from a dysphonic voice).

Only two studies have completed analysis of cepstral measures in both Praat and in ADSV (Murton et al., 2020; Sauder et al., 2017). To distinguish typical voices from individuals with various voice disorders, Murton

et al. (2020) used receiver operating characteristic (ROC) curves to classify cepstral values from a large sample of English speakers with dysphonia ($n = 295$) and a sample of speakers with typical voices ($n = 50$) consisting largely of recordings from the 1990s. Furthermore, the speakers with typical voices in this study consisted of one heterogeneous group of 50 individuals (30 females, 20 males) with an age range of 22–59 years. Sauder et al. (2017) similarly used ROC curves to classify cepstral values from English native speakers from combined groups of males and female individuals aged 18–85 years without voice disorders ($n = 70$) and with voice disorders ($n = 100$). However, neither Murton et al. nor Sauder et al. reported normative values for cepstral measures by age and sex, which are important considerations for measures of voice (Benjamin, 1997; Klatt & Klatt, 1990; Torre & Barlow, 2009; Whiteside & Irving, 1998).

Two additional studies have proposed cutoff scores for using CPP to distinguish typical voices from dysphonic voices speaking in English using ADSV software only (Heman-Ackah et al., 2003, 2014). Heman-Ackah et al. (2003) did not use any participants with typical voices, and Heman-Ackah et al. (2014) used only 50 such control speakers, without controlling for their age or sex. Alongside Sauder et al. (2017) and Murton et al. (2020), other available CPP data on larger datasets of individuals with voice disorders have been presented but in languages other than English, including Turkish (Esen Aydinli et al., 2019), Korean (Lee et al., 2020; Yu et al., 2018), Brazilian Portuguese (Oliveira Santos et al., 2021), and Spanish (Delgado-Hernández et al., 2019; Núñez-Batalla et al., 2019). However, none of these previously reported datasets of cepstral measures examined both software types (ADSV and Praat) by evenly balanced age groups across the life span and by sex. Thus, the lack of stratification by sex in previous studies may have influenced the reported ranges for typical cepstral measures. Two studies have reported investigations of the impact of age, alone, on cepstral values. Garrett (2013) analyzed the CPP from ADSV in individuals from 20 to 30 and 40 to 50 years old using sustained vowels and a portion of The Rainbow Passage, and found reduced CPP in speakers aged 40–50 years compared to the 20–30 group. However, this study was limited in sample size; there were only 60 total participants, with 15 individuals per age group per sex. Furthermore, there were no individuals aged 31–39 years or greater than 50 years of age included in the study, which limits the generalizability of these findings to several other age groups. A study by Taylor et al. (2020) analyzed CPPS from Praat in 169 individuals aged 17–87 years and found no age-related changes in CPPS. However, the speech stimuli examined only consisted of running speech and the age groups were inconsistently balanced, consisting

of roughly 5-year brackets (i.e., 17–20, 21–25, 26–30) from ages 17 to 90 years for males and females with as many as 19 to as few as zero individuals per group. It is also not clear if the two studies have conflicting findings due to the differing software type (Praat vs. ADSV); thus, the impact of age on cepstral measures needs to be examined in both software types and in evenly distributed age groups across the life span. Other studies have examined the impact of both age and sex on cepstral measures in one of the two software types. Lee et al. (2018) reported that in 144 age-matched adults aged 19–49 years, females had higher average CPP values calculated in ADSV compared to males. Yet, Awan et al. (2012) found that CPPS calculated from Praat was significantly lower in females than in males in a sample of 92 typical speakers of English aged 18–30 years. Although the results from these two studies may still stem from different software types, it likely that it is a result of the younger age range investigated in the work of Awan et al. (2012) compared to Lee et al. (2018). In support of this possibility, Oliveira Santos et al. (2021) reported evidence that differences in cepstral measures by sex may be age dependent. In their study, Oliveira Santos et al. analyzed CPPS from Praat in 265 speakers of Brazilian Portuguese aged 30–79 years without voice disorders (140 females) and found that CPPS was lower for females than males in the third decade of life, but also that females demonstrated increases in CPP in the seventh decade of life compared to females in other decades of life. Thus, there appear to be effects of both age and sex on cepstral measures in speakers with typical voices, highlighting the need to examine cepstral measures in both ADSV and Praat across the life span and by sex.

Purpose and Research Aims

The purpose of this study was to calculate measures of CPP from ADSV and CPPS from Praat in individuals with typical voices across the life span, with age brackets matched for sex. In order to emulate a clinical voice evaluation, we used the standard clinical voice protocol stimuli of sustained vowels and the rainbow passage. Our sample consisted of 75 males and 75 females, evenly distributed into three age brackets: ages 18–39, 40–64, and 65–91 years. Considering the recent data reported by Murton et al. (2020) and the need for more clinically translatable data (i.e., using voice stimuli that are elicited in clinical evaluations) regarding CPP and CPPS values in typical speakers, the current work had three aims. The first aim was to establish a set of values for CPP and CPPS in ADSV and Praat, respectively, that could be used as normative reference data for clinical evaluations for voices across the life span. The second aim was to test the hypothesis that CPP via ADSV would yield lower values, regardless of age, than the CPPS via Praat. The final aim

was to test the hypothesis that, controlling for age range, males would have higher CPP and CPPS values compared to females.

Method

Participants

A total of 150 individuals aged 18–91 years participated in the study. Participants were recruited from 2013 until 2020 through flyers and recruiting at various local academic and research events. To investigate the study aims, individuals were divided equally into three age groups (18–39, 40–64, 65–91 years old); each age group had an equal number of male and female speakers,¹ with 25 males and 25 females per age group. All participants denied a history of speech, voice, hearing, language, or neurological disorder. An SLP specializing in voice disorders screened all participants through an auditory-perceptual evaluation to determine that their voices were within normal limits. The SLP listened to each recording (including sustained vowels and running speech) and noted whether the individual presented with a voice quality within normal limits using the framework of the CAPE-V. Some individuals presented with trace to low–mild increased overall severity of dysphonia, breathiness, roughness, and/or strain, but only to a degree the SLP attributed to the typical variance present among individuals without voice disorders. Via verbal interview before the study visit, no participants reported taking hormones or medication that can impact speech function. All participants reported no current or prior history of hearing disorders or wearing of assistive hearing devices (i.e., hearing aids). All study participants completed informed consent in compliance with the Boston University Institutional Review Board. All participants were native English speakers.

Data Collection

Speech data were collected in a sound-attenuated booth (either IAC Acoustics or Eckel Noise Control Technologies) at Boston Medical Center or Boston University. Sound-attenuated booths had comparable ambient noise levels at 31.0, 32.1, or 34.1 dBA as measured via external sound pressure level meter. All participants were given the instructions to use their typical voice at a comfortable pitch and loudness and were recorded while producing the following stimuli: The Rainbow Passage (Fairbanks, 1960), three sustained vocalizations of the vowel /a/, and three sustained vocalizations of the vowel /i/. Acoustic

¹Gender information was not collected for all participants.

recordings at both sites were collected using an omnidirectional microphone (Model MX153) or a dynamic headset microphone (Model WH20XLR or SM35XLR). For all recordings at both locations, the headset microphone was positioned at approximately 45° from the midline and 7 cm from the corner of the mouth. The microphone signal was amplified via a microphone preamplifier (Model Quadmic II, RME Audio) and digitized via a soundcard (Model Ultralite-mk3 Hybrid, MOTU) for data collected at the Boston University. The microphone signal at Boston Medical Center was captured via the headset microphone connected to a handheld audio recording device (Model H4nPro, ZOOM Corporation). All data were collected at a sampling rate of 44100 Hz and 16 bits resolution.

Data Analysis

All speech samples were first amplitude-normalized via peak normalization using Praat by selecting “Modify” and then “Scale peak” with the new absolute peak set to 0.99. Files were then cropped through manual selection in Praat (.wav format) to include the following, per participant: (a) the beginning of the second sentence to the end of the third sentences of The Rainbow Passage; (b) the middle 1 s of each of three productions of /a/; and (c) the middle 1 s of three productions of /i/. The middle 1-s portions of the vowels were selected to capture the steadiest portion of the vowel and to eliminate capturing the onset or offset of phonation, which similarly follows the methods of Watts et al. (2017).

ADSV. ADSV (PENTAX Medical) settings were all set to default except for two components. Under “Advanced setup,” the “CPP Threshold (dB)” was set to 1. The “Apply Vocalic Event Detection” option was also activated. These settings were used to capture more clinically meaningful productions of voice (i.e., those over 1 dB) and to reduce the likelihood of background noise from being detected as voicing. Further algorithmic specifications can be observed in Table 1. Audio files were

imported individually into ADSV. For The Rainbow Passage, the auto select feature (the “A” icon in the taskbar) was used to select the speech signal, and for any instances in which the auto-select feature missed a portion of the speech signal, the entire signal was manually selected. The investigator confirmed accurate selection of the voicing signal and proceeded to analyze the files through the program’s analysis function by pressing the icon “Compute/Display New ADSV Results.” These processes were completed for all participants and stimuli. For vowels, a single rater manually highlighted the middle second of the vowel (shift + click at start and finish timepoints in the signal). The bottom of the task bar displays the length of the highlighted signal in seconds, and this was manually adjusted until each signal was between 1 and 1.05 s. Once 1 s was highlighted, the auto select feature was clicked to set the analysis portion to the highlighted middle second. Next, the “Computer/Display New ADSV Results” was pressed to calculate CPP of the highlighted section. Figure 1 shows a screenshot where these mentioned icons are circled for reference. Individual vowel stimuli were all completed individually and later averaged per participant. To calculate reliability of the portions of ADSV that required manual selection (/a/, /i/, and The Rainbow Passage), 10% of samples were repeated by the initial rater, which resulted in the intrarater reliability of $r > .99$.

Praat. For each of the extracted speech stimuli, Praat software (Boersma & Weenink, 2016) Version 6.0.50 was used to calculate the CPPS values using the methods described in the work of Watts et al. (2017). Each sound file was loaded into Praat and selected in the program window. Under the “Analyze Periodicity” drop down menu, the “To PowerCepstrogram” option was selected. For the resulting settings window, the pitch floor was set to 60 Hz, the time step was set to 0.002 s, and the maximum frequency was set to 5000 Hz. Once the power cepstrogram was generated, the resulting file was selected

Table 1. Algorithmic specification differences for calculating CPP/CPPS between the acoustic analysis platforms Analysis of Dysphonia in Speech and Voice (ADSV) and Praat, as discussed in the work of Watts et al. (2017).

Variable	ADSV	Praat
Voicing activity detection (VAD)	On	Not available
Power cepstrum construction	Logarithm of the amplitude spectrum → power cepstrum → forward Fourier transform of the log spectrum → cepstrum	Logarithm of the amplitude spectrum → power cepstrum → inverse Fourier transform of the log spectrum → cepstrum
Line of best fit	Simple least squares linear regression	Theil robust fitting method
Regression line	Quefreny value = 0.0001 s (10 kHz)	Quefreny value = 0.001 s (1 kHz)
Window function	Hamming window	Gaussian window
Interpolation	No interpolation	Parabolic
Sampling frequency	Dependent pre-emphasis	Independent pre-emphasis

Note. CPP/CPPS = cepstral peak prominence/smoothed cepstral peak prominence.

Figure 1. Screenshot of the window in Analysis of Dysphonia in Speech and Voice (ADSV) during the analysis of The Rainbow Passage’s CPP. The “A” icon is highlighted in an orange circle, which represents the function “Apply automatic data selection,” as well as the compute results button, which represents that function “Computer/Display new ADSV results.” CPP = cepstral peak prominence.



again in the program window. From the “Query” menu, the “Get CPPS” option was selected. In the resulting settings window, the “Subtract tilt before smoothing” option was unchecked, the time-averaging window was set to

0.01 s, the quefrency averaging window was set to 0.001 s, the peak search pitch range was set to 60–330 Hz, the tolerance was set to 0.05, the interpolation was set to parabolic, the tilt line quefrency range was set to 0.001–0.0 s,

the line type was set to straight, and the fit method was set to “robust.” The resulting value was used as the CPPS measure for each of the stimuli. Further algorithmic specifications can be observed in Table 1. This process was completed for all participants. Individual vowel stimuli were all completed individually and later averaged per participant.

Statistical Analysis

All statistical analyses were performed using Minitab 19 software. A significance level of $p < .05$ was set a priori. Since the data failed to meet conditions for data normality, nonparametric statistics were conducted. Kruskal–Wallis tests were conducted on the cepstral measures using factors of age group (18–39 vs. 40–64 vs. 65–91), sex (female vs. male), stimuli (/a/ vs. /i/ vs. The Rainbow Passage), and software (Praat vs. ADSV). Eta squared based on the H-statistic (η^2 [H]) was used to quantify the effect size of statistically significant effects from the Kruskal–Wallis tests (small effect = 0.01, medium effect = 0.06, and large effect = 0.14; Kennedy, 1970). Post hoc Mood’s Median tests were used to evaluate directional effects from significant factors in the Kruskal–Wallis tests with more than two factors (age group and stimuli).

The cepstral measure normative values were determined by groups based on statistically significant factors in the Kruskal–Wallis tests. The normative values for cepstral measures were set as 2 *SDs* below the group mean. This method is consistent with prior methods of determining normative scores in other areas in the field of speech-language pathology, such as quality of life measures (Arffa et al., 2012; Zraick et al., 2011). The resulting norms are designed to encapsulate roughly 95% of the data.

Results

The Kruskal–Wallis tests revealed that sex ($df = 1$, $H = 14.18$, $p < .001$, η^2 [H] = 0.01, small effect size), software ($df = 1$, $H = 198.25$, $p < .001$, η^2 [H] = 0.22, large effect size), and stimuli ($df = 2$, $H = 431.43$, $p < .001$, η^2 [H] = 0.48, large effect size) were statistically significant factors for cepstral measures. Age group did not show a statistically significant effect on cepstral measures ($p = .93$). Table 2 catalogs the mean cepstral values as a function of the statistically significant factors.

Post hoc Mood’s Median tests were used to determine directional effects. Males had statistically higher cepstral values ($M = 12.62$ dB, $SD = 4.34$ dB) compared to females ($M = 11.47$ dB, $SD = 3.87$ dB; $df = 1$, chi-square = 10.24, $p < .001$). Praat software yielded statistically higher cepstral values ($M = 14.05$ dB, $SD = 4.27$ dB) than those of ADSV software ($M = 10.04$ dB, $SD = 2.87$ dB; $df = 1$, chi-square = 94.74, $p < .001$). All stimuli yielded statistically different cepstral measures from each other. The /a/ stimulus had higher average cepstral values ($M = 14.76$ dB, $SD = 3.27$ dB) compared to /i/ ($M = 13.05$ dB, $SD = 4.09$ dB) and The Rainbow Passage ($M = 8.32$ dB; $SD = 1.44$ dB; $df = 2$, chi-square = 443.86, $p < .001$).

Normative values for males were /a/, ADSV (CPP) = 8.86 dB; /a/, Praat (CPPS) = 11.72 dB; /i/, ADSV (CPP) = 6.68 dB; /i/, Praat (CPPS) = 12.01 dB; The Rainbow Passage, ADSV (CPP) = 5.40 dB; The Rainbow Passage, Praat (CPPS) = 6.40 dB. For females, the values were /a/, ADSV (CPP) = 8.09 dB; /a/, Praat (CPPS) = 11.05 dB; /i/, ADSV (CPP) = 4.65 dB; /i/, Praat (CPPS) = 10.37 dB; The Rainbow Passage, ADSV (CPP) = 5.04 dB; The Rainbow Passage, Praat (CPPS) = 6.49 dB. This information is represented in Table 3.

Normative values for males were /a/, ADSV (CPP) = 8.86 dB; /a/, Praat (CPPS) = 11.72 dB; /i/, ADSV (CPP) = 6.68 dB; /i/, Praat (CPPS) = 12.01 dB; The Rainbow Passage, ADSV (CPP) = 5.40 dB; The Rainbow Passage, Praat (CPPS) = 6.40 dB. For females, the values were /a/, ADSV (CPP) = 8.09 dB; /a/, Praat (CPPS) = 11.05 dB; /i/, ADSV (CPP) = 4.65 dB; /i/, Praat (CPPS) = 10.37 dB; The Rainbow Passage, ADSV (CPP) = 5.04 dB; The Rainbow Passage, Praat (CPPS) = 6.49 dB. This information is represented in Table 3.

Discussion

Normative Values for CPP and CPPS

The first aim of the current work was to establish a set of normative values for CPP and CPPS. This was accomplished through the calculation of cepstral measures from the vowels /a/ and /i/, as well as the second and third sentences of The Rainbow Passage in males and females across the life span (aged 18–91 years) and using two

Table 2. Mean and standard deviation values for the significant factors of sex, stimulus, and software, presented in decibels (dB).

Factor	Males		Females	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
/a/, ADSV (CPP)	13.28 dB	2.21 dB	12.07 dB	1.99 dB
/a/, Praat (CPPS)	17.52 dB	2.90 dB	16.17 dB	2.56 dB
/i/, ADSV (CPP)	10.96 dB	2.14 dB	8.75 dB	2.05 dB
/i/, Praat (CPPS)	17.23 dB	2.61 dB	15.27 dB	2.45 dB
The Rainbow Passage, ADSV (CPP)	7.78 dB	1.18 dB	7.40 dB	1.18 dB
The Rainbow Passage, Praat (CPPS)	8.92 dB	1.26 dB	9.17 dB	1.34 dB

Note. ADSV = Analysis of Dysphonia in Speech and Voice; CPP/CPPS = cepstral peak prominence/smoothed cepstral peak prominence.

Table 3. Proposed normative values from this study are calculated as the value that is 2 SDs below the group mean.

Stimulus	Normative value (males)	Normative value (females)	Murton et al. (2020) cutoff value	Sauder et al. (2017) cutoff value
/a/, ADSV (CPP)	8.86 dB	8.09 dB	11.46 dB	DNT
/a/, Praat (CPPS)	11.72 dB	11.05 dB	14.45 dB	DNT
/i/, ADSV (CPP)	6.68 dB	4.65 dB	DNT	DNT
/i/, Praat (CPPS)	12.01 dB	10.37 dB	DNT	DNT
The Rainbow Passage, ADSV (CPP)	5.40 dB	5.04 dB	6.11 dB	5.53 dB
The Rainbow Passage, Praat (CPPS)	6.40 dB	6.49 dB	9.33 dB	19.10 dB

Note. Murton et al. (2020) suggested cutoff values are listed as a comparison. DNT represents “did not test,” where no specific value was presented. ADSV = Analysis of Dysphonia in Speech and Voice; CPP/CPPS = cepstral peak prominence/smoothed cepstral peak prominence.

software types (ADSV and Praat). Statistical analyses revealed that age group did not statistically impact CPP/CPPS values. Thus, normative values were reported by averaging all age groups by sex (male and female), stimuli type (/a/, /i/, and The Rainbow Passage), and software (Praat and ADSV). Of note, Mood’s Median tests were applied post hoc in order to evaluate directional effects from statistically significant factors in the Kruskal–Wallis tests. Only two of the post hoc tests were related to study aims (i.e., comparing ADSV to Praat and males to females) with the other tests presented in order to provide directionality and associated effect sizes. Considering this, no *p*-value corrections were applied to this analysis. Table 3 presents the proposed normative values, in comparison to the two studies of cepstral values using both ADSV and Praat (Murton et al., 2020; Sauder et al., 2017).

Compared to the two prior investigations that used both ADSV and Praat, the current work found distinctly lower average values in both software types using similar speech stimuli, which may be a result of differing methodology as well as the larger age range examined in the current work. Prior to analysis in either software program, all files were normalized via peak-normalization using the “Scale peak” function in Praat. This was conducted in order to facilitate signal visualization during cropping of the audio files for analysis and does not yield any changes to CPP/CPPS values. A sample of the raw data (prenormalization) was reanalyzed and demonstrated that no changes in cepstral values occurred after peak normalization. Thus, it remains a user preference whether or not to complete this step. Regarding other settings, the Praat platform settings for CPPS were adjusted from the program’s default settings in order to match the published settings by Watts et al. (2017), which were also used by Murton et al. (2020). However, Sauder et al. (2017) calculated CPPS in Praat by using the software’s default settings, which are different than those used in the current work or in the work of Murton et al. This could explain the substantial difference in the average CPPS value presented in the work of Sauder et al., 19.10 dB, in

comparison to those of the current work (6.4 dB for males and 6.49 dB for females). In addition, both Sauder et al. and Murton et al. used slightly different running speech stimuli than the current work. Here, we examined CPPS on the second and third sentences of The Rainbow Passage. Instead, Sauder et al. only used the second sentence of The Rainbow Passage and Murton et al. used the first 12 s of the passage, which may have included a variable number of sentences depending on the speaker’s reading rate. Hence, this may have resulted in the CPPS values for The Rainbow Passage reported by Murton et al. (9.33 dB) being higher than those in the current work. Although no statistical impact of age group on cepstral measures was found in the current work, it is possible that the larger, and evenly distributed, age range examined here (150 speakers, aged 18–91 years) resulted in overall lower cepstral values compared to the prior studies (50 speakers aged 22–59 years in Murton et al., 2020; 70 speakers aged 18–85 years in Sauder et al., 2017) given previous reports of lower cepstral values in older compared to younger adults (Garrett, 2013).

Our methods for calculating CPP in ADSV also differed from both of the prior investigations. Murton et al. (2020) and Sauder et al. (2017) reported that CPP was obtained using the default settings in ADSV, which includes the CPP threshold set to 0 dB (this was set to 1 dB in the current work). Murton et al. furthermore reported that the option to “apply vocalic event detection” was turned off (this was set to “on” in the current work). The rationale for our methodology was to increase the likelihood of ADSV excluding background noise or other extraneous sounds unrelated to the voice signal from being included in the CPP calculation, which is more likely to occur in a clinical setting. Inherently, the exclusion of these extraneous sounds (and thus low CPP values likely unrepresentative of the true voice signal, e.g., values below 1 dB) in this methodology would lead to higher cepstral values. Furthermore, our use of vocalic event detection would aim to reduce background noise (and thus lower cepstral values) and potentially inflate our values in

comparison to when it is left off. However, in contrast, we observed lower mean cepstral values in ADSV across stimuli compared to Murton et al. and Sauder et al. It is therefore unlikely that these methodological differences in the application of ADSV explain the lower cepstral values observed here.

The use of sex-specific normative values in the current work differs from previous studies (Murton et al., 2020; Sauder et al., 2017). We found a statistically significant effect of sex on CPP in those without voice disorders, which suggests that not controlling for this factor would inherently disrupt the control group mean. This finding was also observed in the work of Awan et al. (2012) where sex was found to be a statistically significant factor for CPP values. For these reasons, the methodology in the current work and the proposed normative values for CPP may be useful in representing normative CPP values in individuals without voice disorders, as atypical voices (which may present with a wide range of degrees of perceived severity of dysphonia) would be compared to a large dataset of individuals with typical voices and without voice disorders.

Finally, when comparing methodologies between Murton et al. (2020) and Sauder et al. (2017) to the current work, their use of ROC analysis to develop cepstral cutoff values is inherently different than our approach. ROC analysis is primarily used to separate two groups; however, some individuals with voice disorders do not present with dysphonia. Therefore, this methodology may incorporate inherent bias into the group with voice disorders and, thus, impact the specific value used to separate the groups. The methods of the current work were designed to specifically present normative values for those without voice disorders and directly provide that information without the comparative relationship to individuals

with voice disorders. This approach should capture the majority of individuals without voice disorders and may eliminate outliers in the typical population (i.e., those without a voice disorder but with an atypical auditory-perceptual impression of voice quality).

Cutoff values that are not based specifically on the variability inherent in the voices of those without voice disorders may result in misclassifications. To assess this notion, we applied the available cutoff values to the individuals without voice disorders in the current work. When comparing the cutoff values proposed by Murton et al. (2020) and Sauder et al. (2017) to the proposed normative values of the current work, a much larger portion of individuals (i.e., individuals with typical voices) would not pass their cutoff values. The current work's methodology, assuming that the CPP/CPPS values in typical voices follow a normal distribution, would allow for about 5% of individuals in this study to surpass the proposed normative values. Per these proposed normative values, 1.3%–5.3% of individuals in the current work did not meet our proposed normative values, which is consistent with the percent of individuals below 2 SDs of the mean in a typical bell curve. However, 56% of females and 65.3% of males would not pass the cutoff set by Murton et al. (2020) on the stimulus of The Rainbow Passage using Praat. When comparing to Sauder et al. (2017), 100% of individuals in the current work would not pass the proposed cutoff for the Rainbow Passage using Praat, but only 5.3% of males and females for The Rainbow Passage in ADSV would not pass their proposed cutoff. This demonstrates that cutoff values used in ROC analysis, when applied solely to individuals without voice disorders, do not appear to adequately represent a group of individuals without voice disorders. Table 4 presents this complete information across sex, software, and stimuli.

Table 4. The number of individuals from the current sample of individuals without voice disorders who do not pass the proposed normative values in the current work and those who did not pass the proposed cutoff values in the works of Murton et al. (2020) and Sauder et al. (2017) by software and sex.

Article, sex group	Rainbow – Praat	Rainbow – ADSV	/a/ average – Praat	/a/ average – ADSV	/i/ average – Praat	/i/ average – ADSV
Current work, males	3 (4%)	4 (5.3%)	2 (2.6%)	3 (4%)	1 (1.3%)	3 (4%)
Current work, females	1 (1.3%)	3 (4%)	1 (1.3%)	1 (1.3%)	1 (1.3%)	2 (2.6%)
Murton et al. (2020), males	49 (65.3%)	6 (8%)	9 (12%)	15 (20%)	DNT	DNT
Murton et al. (2020), females	42 (56%)	10 (13%)	20 (26.6%)	30 (40%)	DNT	DNT
Sauder et al. (2017), males	75 (100%)	4 (5.3%)	DNT	DNT	DNT	DNT
Sauder et al. (2017), females	75 (100%)	4 (5.3%)	DNT	DNT	DNT	DNT

Note. The values preceding the parentheses represent the number of participants whose cepstral values were below the proposed threshold value, and the percentage in parentheses represents this as a percent of the total number of participants by sex group. DNT represents “did not test,” so no specific value was presented. ADSV = Analysis of Dysphonia in Speech and Voice.

Cepstral Differences in ADSV and Praat

Our second aim was to test the hypothesis that CPP (via ADSV) would yield lower values, regardless of age, than the CPPs calculated in Praat. Our results showed distinct, statistically significant effects of platform on CPP/CPPS values, indicating that the platforms ADSV and Praat do in fact yield statistically significant differences in CPP and CPPS, respectively, and with a large effect size. These findings were consistent with those of both Watts et al. (2017) and Murton et al. (2020). We found that the values of CPP from ADSV were consistently lower than the CPPS values from Praat across productions. This finding confirms the notion that ADSV values of CPP and Praat values of CPPS are not directly comparable. Therefore, we recommend that individuals should compare CPP values from ADSV to normative values of the same and Praat-derived CPPS values to normative values of the same. Regarding within platform differences, Grillo and Wolfberg (2023) sought to analyze the differences in CPPS values across multiple stimuli calculated in different versions of Praat (Versions 6.0.32 and 6.1.05). They used the procedures described by Maryn and Weenink (2015) and found that different Praat versions only yielded differences in CPPS at the phrase level but not during sustained /a/ (although they did not use the same stimuli that were used in the current work). However, despite these differences between software versions, CPPS at the phrase level still demonstrated moderate reliability across versions. Comparing Praat versions, the current work used Version 6.0.50, Murton et al. (2020) used Version 6.0.40, and Sauder et al. (2017) used Version 6.0.17. None of these studies used the same versions tested by Grillo and Wolfberg, but if their findings apply to other versions of Praat, then moderate reliability would be expected between these compared versions and those of the current work. The current work did find distinctly lower cepstral values from these works that used different Praat software versions, but considering the overarching, aforementioned differences in methodology, it is not clear whether these findings can be attributed to the methodological differences or any potential differences across Praat versions. To account for any potential future differences in Praat versions that may yield reduced reliability of CPPS across platform versions, future work should report which Praat version is used for CPPS analysis in order to best contextualize CPPS values.

Effects of Age and Sex on Cepstral Values

Our third and final aim was to test the hypothesis that males would have higher CPP and CPPS values compared to females, when controlling for age. In order to provide a set of more clinically useful normative values and to reflect broader changes across the life span, we

chose to analyze age groups (ages 18–39, 40–64, 65–91 years), rather than treating age as a continuous variable in our model. We found statistically significant differences in CPP/CPPS between males and females. We hypothesized that males would have higher CPP values than females due to inherent differences in vocal fold anatomy and physiology, considering vocal fold mass, glottal contact, and vibratory characteristics such as vocal fold contact quotient. Contact quotient alone may explain this hypothesis as it has been demonstrated that females tend to have greater open quotient and thus less total glottal contact as a function of time during phonation than males (Holmberg et al., 1988; Tsutsumi et al., 2017). Furthermore, sex-specific vibratory characteristics have been described, such as lower asymmetry quotient and increased maximum area declination rate in males compared to females (Patel et al., 2014). These differences, subsequently, may contribute to variations in the acoustic source spectrum derived from the vocal folds themselves and contribute to reduced cepstral values in females. Acoustically, a longer closed phase contributes to a lower spectral slope (i.e., reduced spectral tilt), in which the amplitude of higher harmonics is increased and spread across higher frequencies in the spectrum. This general increase in harmonic amplitudes may at least partially explain increased CPP values in males than in females, where harmonic peaks are impacted by harmonic amplitudes (Fraile & Godino-Llorente, 2014). The lack of an observed effect of age on CPP/CPPS values is consistent with Taylor et al. (2020), but differs from the findings of Oliveira Santos et al. (2021) and Garrett (2013). However, the current work contained wider age brackets than those used in the work of Garrett, and those participating in the work of Oliveira Santos et al. were not speakers of English and reported a single sustained vowel. Therefore, it is possible that small changes in CPP/CPPS may be present between smaller stratified groups and that our larger age ranges may have masked these small potential differences. Alternatively, the limited sample sizes and different stimuli across the previous and current work may have also contributed to the observed differences.

Relatedly, the impact of speaker sound pressure level (SPL) on cepstral values has been examined. With increasing SPL values, it has been demonstrated that cepstral values also increase in vowels and in speech (Brockmann-Bauser et al., 2021; de Oliveira Florencio et al., 2021; Sampaio et al., 2020). These works have demonstrated that speakers, when asked to modulate their vocal amplitude, demonstrate significant changes in CPP. One hypothesis for this observed increase of CPP is that higher SPL values may be attained through increasing medial contact of the vocal folds, yielding increased signal periodicity (Awan et al., 2012), alongside higher amplitudes of

harmonics, which may generally raise harmonic peaks and thus raise cepstral values. Furthermore, it has been demonstrated that males have higher average SPL values than females (Awan et al., 2012; Brockmann-Bauser et al., 2018), which may explain the cepstral sex differences found in the current work. Considering this positive relationship with speaking SPL and cepstral values, it has been suggested that SPL be considered when reporting cepstral values (Brockmann-Bauser et al., 2021). Calibrated SPL values were not available for the entire sample and thus not reported in the current work. We asked individuals to use their typical speaking voices, which we expect would encompass various typical speaking intensity levels and thus represent the average of typical SPL variations across speakers without voice disorders.

Overall, the current work establishes normative values for CPP and CPPS by sex, software, and stimuli in the largest sample to date (150 speakers) across the life span (aged 18–91 years). In order to assist with clinical translation, we chose to use all stimuli that are commonly used in typical voice evaluations by SLPs and two commonly used clinically available software platforms. Our findings suggest that CPP and CPPS values are different, and thus not directly comparable, and that normative values for cepstral measures should be further separated by stimuli and by sex. In the current sample, the speaker's age group did not statistically impact cepstral values, which yielded normative values for cepstral measures that were age independent. These conclusions allowed us to present a set of normative values (see Table 3) for determining whether a voice is below what is considered typical for a large pool of speakers with typical voices and without voice disorders, which we believe will increase the utility of both CPP and CPPS measures in clinical environments and in research endeavors.

Limitations and Future Directions

There are some limitations in the current work. The main limitation of our study is that the participants were not evaluated via laryngeal videostroboscopy, which may have allowed for individuals with laryngeal pathology to enter our control data. However, all voices were screened by a voice-specializing SLP to ensure they were perceived via auditory-perception as a typical voice for that individual's age and sex. Another limitation of the current work is that we did not control for the potential influence of speaker SPL on cepstral values. Despite some research demonstrating differences in cepstral values with varying intensity levels, the current work represents the average cepstral values of individuals asked to speak in their normal, typical speaking voice (which we expect to represent varying speaker to speaker SPL variations). However,

future work may consider controlling for SPL in a normative sample to examine its impact on threshold or normative values. Another limitation is that the voicing activity detection (VAD) in ADSV is not specifically detailed in the software's user manual, so, in the current work, we cannot rule out possible differences between the VAD algorithm in ADSV and the cycle recognition feature in Praat. The details of the voicing detection algorithms may be a contributing factor to the two software programs yielding different cepstral values for the same stimuli (as observed in the current work as well as prior literature). Future work may consider reducing the age range of the age brackets by spans of only 10 years, to examine possible differences more closely across age, while maintaining an adequately large sample size in each age bracket. Yet, the current work did not find statistical differences in CPP/PPS values by age group, suggesting that age is unlikely to have a major impact on CPP/PPS values. Future directions may seek to apply the proposed values in a clinical context to assess their discriminative abilities.

Conclusions

Significant main effects for cepstral measures included sex, stimuli, and software, but age was not a significant factor. Thus, the current work yields normative values for cepstral measures by sex (female and male), clinically relevant speech stimuli (*/a/*, */i/*, and The Rainbow Passage), and clinically used software (ADSV and Praat), based on 150 speakers (75 males, 75 females) with typical voices that were evenly distributed into three age brackets throughout the life span (ages 18–39, 40–64, 65–91 years). There were no statistical effects of age group on cepstral measures, so normative values were reported combining speakers across age groups. Future work should evaluate how these normative cepstral values compare to other methods of clinical diagnosis as well as their relation to voice changes before and after successful voice therapy intervention.

Data Availability Statement

The datasets generated during and/or analyzed during this study are not publicly available due to the inability to fully deidentify voice recordings, but are available from the corresponding author on reasonable request.

Acknowledgments

This work was supported by Grants R01 DC015570 (to Cara E. Stepp) and F31 DC019032 (to Defne Abur) from the National Institute on Deafness and Other

Communication Disorders. The authors would like to thank Katherine Brown for assistance with participant selection.

References

- Arffa, R. E., Krishna, P., Gartner-Schmidt, J., & Rosen, C. A. (2012). Normative values for the Voice Handicap Index-10. *Journal of Voice*, 26(4), 462–465. <https://doi.org/10.1016/j.jvoice.2011.04.006>
- Awan, S. N., & Frenkel, M. L. (1994). Improvements in estimating the harmonics-to-noise ratio of the voice. *Journal of Voice*, 8(3), 255–262. [https://doi.org/10.1016/s0892-1997\(05\)80297-8](https://doi.org/10.1016/s0892-1997(05)80297-8)
- Awan, S. N., Giovanco, 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., Jette, 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 & Phonetics*, 24(9), 742–758. <https://doi.org/10.3109/02699206.2010.492446>
- Benjamin, B. J. (1997). Speech production of normally aging adults. *Seminars in Speech and Language*, 18(02), 135–141. <https://doi.org/10.1055/s-2008-1064068>
- Boersma, P., & Weenink, D. (2011). *Praat: Doing phonetics by computer* [Computer program] (Version 5.2.14). <http://www.praat.org>
- Boersma, P., & Weenink, D. (2016). *Praat: Doing phonetics by computer* [Computer program] (Version 5.3-6.0). <http://www.praat.org>
- Brockmann-Bauser, M., Bohlender, J. E., & Mehta, D. D. (2018). Acoustic perturbation measures improve with increasing vocal intensity in individuals with and without voice disorders. *Journal of Voice*, 32(2), 162–168. <https://doi.org/10.1016/j.jvoice.2017.04.008>
- 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>
- Delgado-Hernández, J., León-Gómez, N., & Jiménez-Álvarez, A. (2019). Diagnostic accuracy of the smoothed cepstral peak prominence (CPPS) in the detection of dysphonia in the Spanish language. *Loquens*, 6(1), Article e058. <https://doi.org/10.3989/loquens.2019.058>
- de Oliveira Florencio, V., Almeida, A. A., Balata, P., Nascimento, S., Brockmann-Bauser, M., & Lopes, L. W. (2021). Differences and reliability of linear and nonlinear acoustic measures as a function of vocal intensity in individuals with voice disorders. *Journal of Voice*. Advance online publication. <https://doi.org/10.1016/j.jvoice.2021.04.011>
- Esen Aydinli, F., Ozebe, E., & Incebay, O. (2019). Use of cepstral analysis for differentiating dysphonic from normal voices in children. *International Journal of Pediatric Otorhinolaryngology*, 116, 107–113. <https://doi.org/10.1016/j.ijporl.2018.10.029>
- Fairbanks, G. (1960). *Voice and articulation drillbook* Harper & Row.
- Fletcher, H. (1934). Loudness, pitch and the timbre of musical tones and their relation to the intensity, the frequency, and the overtone structure. *The Journal of the Acoustical Society of America*, 6(2), 59–69. <https://doi.org/10.1121/1.1915704>
- Fraile, R., & Godino-Llorente, J. I. (2014). Cepstral peak prominence: A comprehensive analysis. *Biomedical Signal Processing and Control*, 14, 42–54. <https://doi.org/10.1016/j.bspc.2014.07.001>
- Garrett, R. K. (2013). *Cepstral- and spectral-based acoustic measures of normal voices* [Doctoral dissertation, University of Wisconsin-Milwaukee].
- Grillo, E. U., & Wolfberg, J. (2023). An assessment of different Praat versions for acoustic measures analyzed automatically by VoiceEvalU8 and manually by two raters. *Journal of Voice*, 37(1), 17–25. <https://doi.org/10.1016/j.jvoice.2020.12.003>
- Heman-Ackah, Y. D., Heuer, R. J., Michael, D. D., Ostrowski, R., Horman, M., Barody, M. M., Hillenbrand, J., & Sataloff, R. T. (2003). Cepstral peak prominence: A more reliable measure of dysphonia. *Annals of Otolaryngology, 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>
- Hirano, M. (1981). *Clinical examination of voice*. Springer-Verlag.
- 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>
- Isshiki, N. (1964). Regulatory mechanism of voice intensity variation. *Journal of Speech, Language, and Hearing Research*, 7(1), 17–29. <https://doi.org/10.1044/jshr.0701.17>
- Jenkins, R. A. (1961). Perception of pitch, timbre, and loudness. *The Journal of the Acoustical Society of America*, 33(11), 1550–1557. <https://doi.org/10.1121/1.1908496>
- Kempster, G. B., Gerratt, B. R., Verdolini Abbott, K., 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\(2008/08-0017\)](https://doi.org/10.1044/1058-0360(2008/08-0017))
- Kennedy, J. J. (1970). The eta coefficient in complex ANOVA designs. *Educational and Psychological Measurement*, 30(4), 885–889. <https://doi.org/10.1177/001316447003000409>
- Kitajima, K., & Gould, W. J. (1976). Vocal shimmer in sustained phonation of normal and pathologic voice. *Annals of Otolaryngology, Rhinology & Laryngology*, 85(3), 377–381. <https://doi.org/10.1177/000348947608500308>
- Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America*, 87(2), 820–857. <https://doi.org/10.1121/1.398894>
- Lee, S. J., Pyo, H. Y., Choi, H.-S. (2018). Normative data of cepstral and spectral measures in Korean adults using vowel phonation and passage reading tasks. *Communication Sciences & Disorders*, 23(1), 208–216. <https://doi.org/10.12963/csd.18474>
- Lee, Y., Kim, G., & Kwon, S. (2020). The usefulness of auditory perceptual assessment and acoustic analysis for classifying the voice severity. *Journal of Voice*, 34(6), 884–893. <https://doi.org/10.1016/j.jvoice.2019.04.013>
- Lieberman, P. (1961). Perturbations in vocal pitch. *The Journal of the Acoustical Society of America*, 33(5), 597–603. <https://doi.org/10.1121/1.1908736>

- 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>
- Maryn, Y., & Weenink, D. (2015). Objective dysphonia measures in the program Praat: Smoothed cepstral peak prominence and acoustic voice quality index. *Journal of Voice*, *29*(1), 35–43. <https://doi.org/10.1016/j.jvoice.2014.06.015>
- 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., & Llorente-Pendás, J. L. (2019). Validation of the measures of cepstral peak prominence as a measure of dysphonia severity in Spanish-speaking subjects. *Acta Otorrinolaringologica* (English Edition), *70*(4), 222–228. <https://doi.org/10.1016/j.otoeng.2018.04.005>
- 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*. Advance online publication. <https://doi.org/10.1016/j.jvoice.2021.06.003>
- Oppenheim, A. V., & Schaffer, R. W. (2004). From frequency to quefrequency: A history of the cepstrum. *IEEE Signal Processing Magazine*, *21*(5), 95–106. <https://doi.org/10.1109/MSP.2004.1328092>
- Patel, R. R., Awan, S. N., Barkmeier-Kraemer, J., Courey, M., Deliyski, D., Eadie, T., Paul, D., Švec J. G., & Hillman, R. E. (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., & Dollinger, 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>
- Roy, N., Barkmeier-Kraemer, J., Eadie, T., Sivasankar, M. P., Mehta, D., Paul, D., & Hillman, R. (2013). Evidence-based clinical voice assessment: A systematic review. *American Journal of Speech-Language Pathology*, *22*(2), 212–226. [https://doi.org/10.1044/1058-0360\(2012/12-0014\)](https://doi.org/10.1044/1058-0360(2012/12-0014))
- Roy, N., Merrill, R. M., Gray, S. D., & Smith, E. M. (2005). Voice disorders in the general population: Prevalence, risk factors, and occupational impact. *Laryngoscope*, *115*(11), 1988–1995. <https://doi.org/10.1097/01.mlg.0000179174.32345.41>
- Sampaio, M., Vaz Masson, M. L., de Paula Soares, M. F., 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
- 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>
- Taylor, S., Dromey, C., Nissen, S. L., Tanner, K., Eggett, D., & Corbin-Lewis, K. (2020). Age-related changes in speech and voice: Spectral and cepstral measures. *Journal of Speech, Language, and Hearing Research*, *63*(3), 647–660. https://doi.org/10.1044/2019_JSLHR-19-00028
- Torre, P., III, & Barlow, J. A. (2009). Age-related changes in acoustic characteristics of adult speech. *Journal of Communication Disorders*, *42*(5), 324–333. <https://doi.org/10.1016/j.jcomdis.2009.03.001>
- 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., & 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., & 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>
- Whiteside, S. P., & Irving, C. J. (1998). Speakers' sex differences in voice onset time: A study of isolated word production. *Perceptual and Motor Skills*, *86*(2), 651–654. <https://doi.org/10.2466/pms.1998.86.2.651>
- Yu, M., Choi, S. H., Choi, C.-H., & Choi, B. (2018). Predicting normal and pathological voice using a cepstral based acoustic index in sustained vowels versus connected speech. *Communication Sciences & Disorders*, *23*(4), 1055–1064. <https://doi.org/10.12963/csd.18550>
- Zraick, R. I., Kempster, G. B., Connor, N. P., Thibeault, S., Klaben, B. K., Bursac, Z., Thrush, C. R., & Glaze, L. E. (2011). Establishing validity of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V). *American Journal of Speech-Language Pathology*, *20*(1), 14–22. [https://doi.org/10.1044/1058-0360\(2010/09-0105\)](https://doi.org/10.1044/1058-0360(2010/09-0105))