ACOUSTIC ANALYSIS OF THE VOICE IN NATIVE AND

NON-NATIVE ENGLISH SPEAKERS

A Thesis Presented to the Faculty of the Department

of Communication Sciences and Disorders

University of Houston

In Partial Fulfillment of the

Requirements for the

Degree of Master of Arts

By

Teresa Procter

April 2019

ACOUSTIC ANALYSIS OF THE VOICE IN NATIVE AND

NON-NATIVE ENGLISH SPEAKERS

Teresa B. Procter

Ashwini Joshi, Ph.D.
Committee Chair
Department of Communication Sciences and Disorders

Laura Cizek, M.A. CCC-SLP
Department of Communication Sciences and Disorders

Elizabeth Goodin-Mayeda, Ph.D.
Department of Hispanic Studies

Antonio D. Tillis, Ph.D.
Dean, College of Liberal Arts and Social Sciences
Department of Hispanic Studies

ACOUSTIC ANALYSIS OF THE VOICE IN NATIVE AND NON-NATIVE ENGLISH SPEAKERS

An Abstract Presented to the Faculty of the Department

of Communication Sciences and Disorders

University of Houston

In Partial Fulfillment of the

Requirements for the

Degree of Master of Arts

By

Teresa Procter

April 2019

ABSTRACT

There are significant differences in voice parameters for nonspeech tasks between culturally and linguistically diverse sample populations. Normative data of standard clinical software programs is typically comprised of North American speakers of Standard American English (SAE). There is a need for normative spectral and cepstral data across sociolinguistic groups to ensure clinical objective measurements are accurately classifying the voice quality of all individuals. The purpose of this study was to (1) compare objective measures of voice quality assessment of monolingual speakers of SAE with native speakers (L1) of French and Spanish on acoustic spectral and cepstral analyses; and (2) compare ratings on auditory-perceptual assessment with acoustic data secondary to degree of accentedness. Perceptual analyses and voice quality measures for frequency, cepstral measures, and perturbation measures were analyzed using the Analysis of Dysphonia in Speech and Voice (ADSV) and the MultiDimensional Voice Program (MDVP) in 10 L1 English, 10 L1 Spanish, 10 L1 French speakers. An informal measure of accentedness of SAE was performed. ANOVAs were included for identifying group differences in perceptual ratings and acoustic data, and the relationship between the degree of accentedness and CAPE-V scores. Results indicate relative accuracy of objective measures of voice assessment for multicultural clinical populations. In this study, two variables demonstrated significance for main effect of language, with 17 other variables of voice quality at minimum, approaching significance. Therefore, appropriate consideration of cultural competence is warranted in therapeutic and diagnostic services.

ACKNOWLEDGMENTS

The diversity of Houston, a vibrant symbol to the world, became the impetus for this thesis. I was inspired to glean if all voices from culturally diverse, sociolinguistic backgrounds were considered equal to established normative data for clinical voice evaluations. Sincere gratitude is to be given to my thesis advisor, Dr. Ashwini Joshi, from the Department of Communication Sciences and Disorders at the University of Houston, for her endless assistance and support in this research project. I extend a sincere thank-you as well to my other thesis committee members, Ms. Laura Cizek (Department of Communication Sciences and Disorders) and Dr. Elizabeth Goodin-Mayeda (Department of Hispanic Studies), who encouraged my research questions further. I am appreciative of the faculty, staff, and fellow graduate student body in the Department of Communication Sciences and Disorders at the University of Houston for positively influencing my clinical perspective as a budding speech-language pathologist, and advocating investment in research for the advancement of our field.

TABLE OF CONTENTS

INTRODUCTION

The process of phonation allows for language to be communicated through one's voice. Language may be expressed through verbal and nonverbal communication, but speech production of language occurs through the engagement of an intricate network among the neurological, respiratory, phonatory, resonatory, and articulatory systems of an individual to communicate a signal composed of voiceless and voiced phonemes. While the physiological integration for speech production is a universal process, the reflection of phonological and phonetic characteristics unique to each spoken language begets each voice as a cultural construct. As the demographics of the world continue to expand and diversify, the acoustic characteristics of a speaker represent an infusion of the community and the accent of speakers, as well as an internal to external transmission of a message shared via speech. Assessment of voice must therefore be composed of international standardized measures and awareness of language and cultural contributions to voice quality, to ensure high sensitivity in evaluation of voice.

I. Anatomical, Physiological and Functional Differences Between Ethnicities/Races

A. Anatomy and Physiology

With an intact, typical laryngeal structure, the initial phonation for speech occurs through the systematic interaction of respiration, phonation and resonance. Vibration of the vocal folds becomes acoustic energy immediately as it moves into the supraglottis and into the vocal tract. Before the voice signal emerges from the mouth or nose, numerous factors contribute to its acoustical variations: context, stress, and audience of the message (e.g., speaking to a large crowd unamplified or in an intimate conversation), individual differences in the size and shape of the vocal tract and the laryngeal structure, and the interaction of the vocal tract filter (pharyngeal, nasal, and oral cavities) (Awan & Mueller, 1996; Berg et al., 2017). The vocal tract filter has the capacity to manipulate sound for a particular spoken language (i.e., opening of the velopharyngeal port for nasalized vowels in French (Basset, Amelot, Vaissière, & Roubeau, 2001), contributing to the unique acoustic message each speaker shares.

Anatomical differences in vocal tract parameters and nasal cross-sectional areas for culturally diverse populations have also been identified (Corey et al., 1998; Xue & Hao, 2006). Using acoustic reflection technology, Xue and Hao (2006) measured vocal tract dimensions including oral length, oral volume, pharyngeal length, pharyngeal volume, and vocal tract length of 120 Caucasian American, African American, and Chinese female and male speakers. Resultant data revealed significant gender and race main effects for these parameters and established normative data for these culturally diverse populations. Xue and Hao (2006) identified that within their sample population, Caucasian American females had significantly larger pharyngeal volume than that of Chinese and African American female subjects.

B. Functional

While acoustic and aerodynamic parameters in male-female differences, and normative measures for some culturally diverse populations have been obtained, there is a paucity of empirical studies on acquired acoustic and aerodynamic characteristics for other cultures, with little implementation of rigorous research controls. In response to this gap in the literature, Andrianopoulous et al. (2001) studied the effects of four homogeneous groups comprised of native Mandarin Chinese speakers, native Hindi-Indian speakers, and Caucasian and African-American speakers of standard American English (SAE) on spectral and fundamental frequency measures in vowel tasks (prolongation of three isolated vowels [a], [i], and [u]). As expected regarding the main effect of gender between the female and male participants, the female participants of all homogeneous groups demonstrated significantly higher fundamental frequencies (F_0) for the vowel tasks, yet within the results of this dependent variable, the research also indicated that both female and male Mandarin Chinese speakers had significantly higher fundamental frequencies than the other cultural groups. After controlling for dialectal, physiological, and linguistic variables, significant differences were found among the varying cultural background participants in acoustic and aerodynamic assessments for this sample population. Generalization of the data is limited by the isolated nature of the speech tasks. Other researchers evidenced significant differences in F_0 and spectra measures between two languages in data on connected speech for proficient bilingual speakers (Ng, Chen, & Chan, 2012; Järvinen, Laukkanen, & Geneid, 2017; Bahmanbiglu, Mojiri, & Abnavi, 2017). Voice quality in linguistically diverse speakers may vary as a result of the manipulation of the vocal tract and articulators to produce the phonetic properties unique to a language (Esling & Wong, 1983; Ng, Chen, & Chan, 2012).

II. The Impact of Sociodemographic Factors on the Perception of the Voice: Voice Quality, Accent, and Personal Characteristics

A. Influence of Accent

Native speakers of a language may perceive a voice as normal if its dialect is commonly found within the community and upholds the expected and anticipated variations of prosodic, rate, fluency and voice quality features of speakers within a culture (Celce-Murcia, Brinton, & Goodwin, 1996). A particular pattern of pronunciation within a standard language that may characterize an individual by his/her sex, social class, age, geographic region, profession, and culture can be defined as a speaker's *accent* (Giles, 1970; Wells, 2012; Andrianopoulous, Darrow, & Chen, 2001). An accent is representative of one idiolect, a speaking style influenced by the aforementioned characteristics of an individual and unique to each speaker. Listeners' perceptions of a nonnative accent are therefore impacted if a nonnative speaker conveys a message that interrupts the typical pattern of speech of a particular community and is marked with enough contrastive elements in temporal, tonal, and dynamic features of the native language (Trofimovich & Baker, 2006). Accents influence segmental and suprasegmental properties of the voice: rate of speech, prosody, stress timing, pause frequency and duration (Trofimovich & Baker, 2006; Chen, 2010), as well as the perception of the speaker, though currently there is little agreement to which parameters (segmental or suprasegmental) are more detrimental to comprehension by the listener. The literature reveals that the degree of accentedness is influenced by a plethora of variables: age of second language learning (L2), length of residence in the L2 country, continued use of native language, brain maturation, and motivation (Pisk, MacKay, & Flege, 2001). Numerous studies have also suggested that the authenticity of L2 pronunciation decreases as time between L2 learning and the neurological critical period (learning L2 before 12 years of age) increases (Lenneberg, 1967; Asher and Garcia, 1969; Patkowski, 1980; Oyama, 1982; Scovel, 1988; Flege, 1988). Research by Flege, Yeni-Komshian, & Liu (1999) revealed foreign accents were greater in native Korean speakers in English with increased with age of acquisition. Factors such as phonological system differences between the L1 and L2 languages, education, and language use may also be contributing factors.

B. Perception of Accent

No matter the age of acquisition of L2, articulatory precision of speech through a nonnative accent has shown to affect the speaker, the listener, and the communicative exchange between speaker and listener. Subjectively, the credibility of an individual and perception of professionalism, comprehensibility, and intelligence have all been associated with the degree of nonnative accent (Derwing & Munro, 1997; Munro & Derwing, 1995; Fuse, Navichkova, & Allogio, 2018; Carlson & McHenry, 2006, Flege, 1988). Research suggests that discrepancies in perception of these personal evaluations may be influenced by personal characteristics of the listener: whether a listener is monolingual or bilingual, a native or nonnative speaker of the L2 language (Fuse et al., 2018), or how comprehensible the accent is to a native speaker (Major, Fitzmaurice, Bunta, & Balasubramanian, 2005). In a study by Carlson and McHenry (2006), perceived intelligibility of nonnative speakers positively influenced listeners' perception, but a maximally perceived accent negatively impacted hypothetical employability if comprehensibility was compromised in the speech. Listeners may therefore associate personal evaluations with nonnative speakers, forming judgments on an individual based upon the degree of deviation from the target language.

III. Assessment of Voice

A. Five Areas of Voice Assessment

The most common outcome measures specifically in clinical evaluation of the voice include five approaches: perceptual ratings of voice quality; assessment of the acoustic, aerodynamic, patient perceptual features of the voice; and laryngeal imaging-based assessment (Roy et al., 2013). Acoustic assessments include the recording of a variety of speech tasks such as sustained vowels and continuous speech through spontaneous conversation and reading of written stimuli. Perceptual evaluation of these speech samples is completed in addition to analyzing the recordings in computer software programs. Under the GRBAS Scale (Hirano, 1981), perception of grade (G), roughness (R), breathiness (B), asthenia, weakness in the voice, (A) and strain (S) may be quantifiably evaluated on a four-point scale. CAPE-V (Kempster et al., 2009) is another prominent standardized tool in the auditory-perceptual assessment of the voice due to its sentence stimuli that are constructed to elicit task-dependent differences based on the context (e.g., sentence with absence of nasalization, hard glottal onsets, etc.).

B. Cepstral Analyses

While spectral analyses (jitter, shimmer, noise to harmonic ratio) or perturbation measures provide objective and noninvasive measures of objective data and were considered a robust measure for voice evaluation, a new type of analysis, cepstral analyses, are advantageous in collecting speech samples of connected speech over traditional time-based acoustic measures. Aperiodic signals in a voice sample become unreliable with perturbation measures if the voice sample contains modulation, intermittency, or strong subharmonics (Dejonckere et al, 2001). Additionally, frequency and amplitude perturbation measures may be inflated in voice samples due to excess noise in consonantal production and variations in loudness and pitch (Awan, 2010). Cepstral analyses measure aperiodicity or additive noise in the signal without need of accurate identification of cyclic behavior for these measurements (Awan et al., 2010).

C. Analysis of Dysphonia in Speech and Voice (ADSV)

The Analysis of Dysphonia in Speech and Voice (ADSV), employing cepstral analyses, is emerging as a prominent clinical assessment software for objective measurement of voice disorders. ADSV uses frequency-based acoustic analyses to track the cepstral peak within a cepstrum, computed through a forward Fourier transformation of the logarithmic power spectrum of a recorded sound wave (Awan, 2011). Within ADSV, the Cepstral Spectral Index of Dysphonia (CSID) is a multifactorial estimate of dysphonia severity that is correlated to a visual analog scale to analyze continuous speech samples. CSID is constructed as a function of results from ADSV's analysis of CAPE-V sentences and has established high correlation to these sentences in the literature (Watts & Awan, 2011; Lowell, Kelley, Awan, Colton, & Chan, 2011). Cepstral Peak Prominence (CPP) is an additional measure that correlates to the harmonic structure and well-defined fundamental frequency of the voice signal. A more periodic signal (i.e., higher harmonic energy) corresponds to a more distinct, high amplitude and therefore, higher CPP (Awan, 2011). Due to these aforementioned benefits of cepstral analysis, ADSV is actively used as an objective measurement for baseline measures in voice evaluation, as well as a tool in treatment change and the trajectory toward disorder specific assessment (Gillespie, Dastolfo, Magid, & Gartner-Schmidt, 2014).

IV. Clinical Impact of Sociolinguistic Diversity in Voice Evaluation

Linguistic diversity is emulated in the voice, and globalization and migration throughout the world continue to impact the cultural demographic of the United States. According to the 2011-2015 American Community Survey, there are 62.43 million people (21% of the United States population) five years and over that speak a language other than English at home (U.S. Census Bureau, 2016a). Spanish speakers make up 13% and other Indo-European speakers make up 3.7% of this population. In 2011, French speakers of the United States population were the most dominant Indo-European language other than Spanish with over 1.3 million speakers. In Texas alone, Spanish speakers (5-64 years of age) who speak English as an L2 compose 28% of the state's population (U.S. Census Bureau, 2016b).

A. Perceptual Evaluation of Voice Quality in Cross-Linguistic Studies

Speech-language pathologists (SLPs) encounter patients of varying sociolinguistic diverse backgrounds in daily clinical practice due to the expanding diversity within the United States population as seen above from the U.S. Census Bureau. Many of these clients are L2 learners of English and may continue using their L1. Regarding the prevalence of voice disorders at a national level, at least one-third of the United States population may encounter vocal impairment at some point in their lives (Roy, Merrill, Gray & Smith, 2005). Though there is not solidified consensus within the literature, the language background of an SLP may influence the auditory perceptual evaluation of voice quality and features of the voice (Järvinen, Laukkanen, & Geneid, 2017). Using the GRBAS scale (Hirano, 1981), listeners of different languages evaluated the perceptual parameters of this scale differently based upon their language background. Native American English and Japanese professionals evaluated strain (S) and asthenia (A) differently in English and Japanese voice samples (Yamaguchi, Shrivastav, Andrews, & Niimi, 2003), and native French and Italian professionals evaluated French and Italian voice samples in the perceptual parameter of roughness (R) differently in Ghio et al. (2011), yet Ribeiro Chaves, Campbell, & Côrtes Gama (2017) found no significant differences in the influence of native language on perceptual evaluation of the voice when native Brazilian and Canadian evaluated Brazilian and Canadian voice samples. Using the same voice samples from Yamaguchi et al, Brazilian evaluators in Behlau et al. (2001) judged asthenia (A) and strain (S) differently than Japanese evaluators. While the GRBAS scale used by the investigators was constant in these studies, the evaluators and speakers did not have commensurate language backgrounds.

Breathiness and hypernasality are two parameters of auditory-perceptual evaluation that are included in numerous standardized voice assessments, yet linguistically, aspiration and nasality may contribute to contrastive elements within language or be characteristic of an accent. In addition to French, over 30% of languages in the WALS database use vowel nasality as a phonemic contrast, while nasality in the English language is solely present in coarticulation, in which a nasal consonant preceding and/or following a vowel influences nasal properties of the vowel. Styler (2017) found that in comparing nasality in French and English, the degree of nasality, the enhancement of nasality through breathy voicing (also identified in a cross-linguistic study by Garellek, Ritchart, & Kuang, 2016), and timing differences created acoustically different vowel nasality in French. While more similarities than differences were found in acoustic features of nasality between French and English participants, the acoustical nature of nasality in the voice had marked linguistic differences. Additionally, an L2 speaker of a language with nasalization could be judged as nasal if the L2 has less or no nasalization, due to application of the L1 rule to their L2.

B. **Objective Measure Standardizations in Cross-Linguistic Studies**

Currently, the normative data for clinical acoustic measures are bereft of cultural and linguistically diverse demographic representation of healthy, nonnative English speakers. With new research indicating significant differences in acoustic and aerodynamic parameters for nonspeech tasks between culturally and linguistically diverse sample populations, there is a need to ensure that objective measures are accurately classifying the voice quality of all individuals to avoid misidentification of individuals having a voice disorder. In a cross-sectional study by Malki, Al-Habib, Hagr, and Farahat (2009) to compare acoustic differences in normal adult Saudi male and female voices to the normative acoustic sample database of KayPentax, 15 of 33 Multidimensional Voice Program (MDVP) variables were significantly different for the Saudi male participants in comparison to the KayPentax database, normative data acquired from only North American participants, and 10 of the 33 MDVP variables were significantly different in the Saudi female participants compared to the norms.

Commensurate with MDVP, preliminary normative measures for ADSV were solely obtained from North American speakers, and while diverse populations have been studied using ADSV, with one study including native Flemish speakers, (Awan, Roy, & Cohen, 2014; Watts, Awan, & Maryn, 2017) normative data was not established nor was there disclosure of nonnative speakers of English as study participants using CAPE-V as a connected speech sample. Given that there are physiological differences between races and ethnicities, as supported by research findings, and that accents and languages have an effect on voice quality and its perception, it is important to establish normative data for persons from different populations and backgrounds, as acoustical differences exist between cultures.

V. Objective of Research

The purpose of this study was to (1) compare objective measures of voice quality assessment of monolingual speakers of SAE with native speakers (L1) of French and Spanish on acoustic spectral and cepstral analyses; and (2) compare ratings on auditory-perceptual assessment with acoustic data secondary to degree of accentedness.

CHAPTER II: METHODOLOGY

I. Participants

As a preliminary study, thirty participants were recruited via written and verbal recruitment. The participants were matched according to native language, gender, age, and proficiency in English. Subjects were required to meet the following criteria: (1) healthy adults as reported by the participant, (2) 18-70 years of age, (3) normal voice quality with no current or history of a diagnosed voice disorder, neurological or respiratory disorder, (4) non-smoking for at least 5 years, (5) monolingual, native speakers of English or native speakers of French or Spanish who have learned English as a second language after 16 years of age, and who were born, raised, and educated for the first 16 years of his or her life in the same country (6) elementary and secondary education was provided in the subject's native language and/or regional dialect. Participants were included in the study based on their self report and perceptual assessment of voice quality by the study personnel. The University of Houston graduate students in the Communication Sciences and Disorders Department that fit the inclusion criteria were recruited in this study, as well as other participants from the

University of Houston community and the Houston community at large. Two participants did not fit the inclusionary criteria and were exempted from participation in the study.

II. Procedures

As a prospective cohort design, all participants completed a one-time assessment of voice acoustics. All measures within this study are routinely conducted during a voice assessment by a speech pathologist and are clinically applicable in the evaluation of normal and disordered voice.

Participants completed the informed consent form in the voice lab prior to initiating the study procedures. There was no waiting period between informing the subject and obtaining consent. Prior to data collection, participants completed a demographic questionnaire (Appendix A). Inclusion in the study was based upon the alignment of answers from Appendix A to the inclusionary criteria of participants.

Voice samples were recorded in the voice lab at the University of Houston (119 Clinical Research Center). Each subject was seated in a doubled walled sound-proof booth and recorded using the KayPentax computer system.

A. Acoustic Analysis

Voice samples were uploaded to a password protected computer in the voice lab and analyzed using ADSV and MDVP. Acoustic measures included fundamental frequency, jitter, shimmer, noise to harmonic ratio (NHR), and cepstral measures.

B. Auditory Perceptual Analysis

Participants completed the tasks listed on the CAPE-V form, which include sustained prolongation of the vowel /a/, repeating (6) sentences and reading a standard passage (The Rainbow Passage). Each of the six sentences target specific laryngeal behaviors and clinical signs. Participants responded to the prompt, “Tell me about your favorite place to visit,” for a one minute conversation sample. Two study personnel completed an auditory perceptual evaluation using the CAPE-V form to evaluate vocal attributes of the voices. Randomization of the audio recordings occurred to avoid any recognition or familiarity to sociolinguistic characteristic. Interrater reliability was established by comparing 20% of auditory-perceptual analyses by one graduate clinician and one licensed speech-language pathologist with specialization in voice and voice disorders.

C. Accentedness

Perceptual measures of accentedness for each study participant were obtained from a standardized stimulus (The Rainbow Passage) previously recorded for acoustic and auditory-perceptual analyses. Two graduate speech language pathologists, blinded to the purpose of the study and identifying/demographic characteristics of the participants, with strong experience in working with sociolinguistic populations, completed an informal assessment of degree of accentedness on a 100 word sample (Appendix B). An additional training session was established to standardize rating parameters and the scale unique to this study. Indication of accent was analyzed by marking words within the sample that deviated from production of Standard American English. A total percentage of accentedness was calculated.

D. Interrater Reliability:

Interrater reliability was measured for CAPE-V and accentedness tasks via independent speech language pathologist blinded to the previously obtained results. 20% of the samples were randomly selected to determine consistency in perception among raters. Reliability rate was 95%.

III. ANALYSES

Data was analyzed using the ADSV and MDVP software for raw analysis. All data was entered into an Excel spreadsheet and SPSS was used to perform statistical analysis. ANOVA was used to compare the three sociolinguistic groups for differences in perceptual ratings and ADSV.

IV. RESULTS

Comparison Within Groups

Data was analyzed using the *SPSS version 24.0* (IBM, Armonk, NY). Descriptive statistics (mean, SD) are reported in Table 3 for all 19 parameters of voice quality. A univariate ANOVA was performed and group results for the auditory-perceptual, acoustic, and accentedness parameters are reported in Table 4. Significance was established at $p \leq 0.05$. With no significant differences evidenced between groups for age and demographic characteristics (Table 1 and Table 2), native SAE speakers and native French and Spanish participants were commensurate at baseline.

Table 1: Age Characteristics of Subjects

	Native Language					
	English		French		Spanish	
	Female	Male	Female	Male	Female	Male
	n=5	n=5	n=5	n=5	n=5	n=5
Age (Mean)	28.6	31	37.6	30.2	35	35.2
Range	22-41	24-37	26-69	28-34	29-39	29-46

Table 2: Linguistic Background of Non-Native SAE Subjects

	Native Language			
	French		Spanish	
	Female	Male	Female	Male
	n=5	n=5	n=5	n=5
Formal Education in L1	20.5	21.2	17	18.6
Range	14-34	17-27	15-20	16-24
Residency in L1 Country (Yr; Mo)	27.6	23.75	27.9	26.6
Range	22;0- 35;0	20;0- 27;4	16;4 -27;0	23;0 -28;0
Residency in L2 Country (Yr; Mo)	7.6	7.48	11.68	8.52
Range	4;6- 25;0	3;4- 12;1	5;1-19;0	3;4- 23;10
Age of First Exposure to L2	8.4	10.5	6.8	5.2
Range	0-10	09- 11	4 -12	4 - 6
Age of Social Proficiency in L2	21.4	22.2	20	19.8
Range	18-29	18-29	16-25	18-23
Percentage of L2 spoken in daily life	58.2	73.4	66.8	58.8
Range	20-93	63-86	39-77	24-89

Table 3: Descriptive statistics for CAPE- V, Time-Based Perturbation and Noise Measures, and ADSV results

	Native Language – Mean (SD)					
	English		French		Spanish	
	Male	Female	Male	Female	Male	Female
	n=5	n=5	n=5	n=5	n=5	n=5
Overall	1.60 (1.95)	0.40 (0.55)	0.60 (0.55)	1.80 (4.03)	1.40 (1.65)	0.80 (1.80)
Roughness	1.80 (2.17)	0.60 (0.89)	1.20 (0.84)	1.60 (3.58)	2.20 (0.84)	1.00 (2.24)
Breathiness	0.60 (1.34)	0.00 (0.00)	0.00 (0.00)	0.20 (0.45)	1.00 (1.23)	0.20 (0.45)
Strain	0.20 (0.45)	0.00 (0.00)	0.00 (0.00)	1.40 (3.13)	0.00 (0.00)	0.40 (0.55)
(F_0)	108.94 (10.03)	235.07 (28.67)	119.41 (11.13)	188.10 (42.92)	113.37 (12.42)	204.85 (24.22)
Speaking F_0	109.60 (8.00)	201.66 (21.48)	120.20 (12.85)	191.15 (27.58)	119.12 (16.48)	193.88 (10.01)
Jitter	0.59 (0.21)	1.09 (0.67)	0.55 (0.28)	1.02 (0.59)	1.62 (1.19)	1.36 (0.75)
Shimmer	0.30 (0.11)	0.26 (0.06)	0.29 (0.15)	0.22 (0.04)	0.44 (0.37)	0.72 (0.94)

Noise to Harmonic Ratio (NHR)	0.14 (0.02)	0.10 (0.02)	0.14 (0.02)	0.12 (0.03)	0.16 (0.06)	0.12 (0.02)
Sustained Vowel CPP	13.67 (1.90)	10.06 (1.30)	13.48 (2.15)	12.32 (2.60)	13.68 (1.61)	12.40 (0.80)
Sustained Vowel CSID	-0.68 (10.46)	-2.08 (4.90)	-9.42 (13.19)	-9.31 (9.90)	-5.47 (11.26)	-7.96 (4.95)
All Voiced Sentence CPP	6.98 (0.93)	6.55 (1.90)	6.37 (1.14)	4.83 (0.83)	5.52 (0.29)	5.11 (0.61)
All Voiced Sentence CSID	-9.53 (12.08)	-3.61 (11.71)	-6.39 (10.47)	7.97 (10.80)	-1.38 (12.03)	4.09 (5.51)
Easy Onset Sentence CPP	8.11 (0.62)	8.04 (0.64)	8.65 (1.74)	7.42 (1.53)	8.09 (1.53)	7.62 (1.26)
Easy Onset Sentence CSID	-7.89 (6.16)	-2.81 (13.20)	-9.14 (6.66)	2.79 (15.34)	-9.94 (12.45)	-8.93 (5.46)
Hard Glottal Sentence CPP	5.46 (0.56)	4.85 (0.72)	5.56 (1.03)	4.64 (0.93)	5.01 (1.14)	4.90 (1.07)
Hard Glottal Sentence CSID	1.91 (12.85)	-1.65 (9.25)	-6.70 (12.28)	6.35 (11.94)	-0.24 (18.08)	3.07 (13.50)
Voiceless Plosive Sentence CPP	5.27 (1.03)	5.02 (0.60)	6.30 (1.50)	4.55 (0.77)	5.48 (1.88)	4.68 (0.95)

Voiceless Plosive Sentence CSID	-0.75 (14.30)	7.24 (8.44)	-9.34 (14.02)	7.52 (14.02)	-1.43 (29.07)	11.82 (8.72)

Comparison Between Groups

A. Auditory-Perceptual Analyses

No significant differences were observed in CAPE-V parameters (overall, breathiness, roughness, and strain) for main effects of gender and language, though these main effects and the interaction of gender and language were nearing significance at $p \leq 0.05$.

B. Acoustic Analyses

In this study, acoustic analyses included frequency, intensity, time-based perturbation and noise measures (i.e., jitter, shimmer, & NHR) and cepstral/spectral measures (i.e., CPP and CSID for sustained vowel and CAPE-V sentences). A univariate ANOVA revealed a significant main effect of gender for F_0, Speaking F_0 and NHR. As expected, male participants had a significantly lower F_0 ($F= 112.913$, $p \leq .-000$) and Speaking F_0 ($F =155.049$, $p \leq 0.000$) than female participants. In noise-based measures, a significant main effect of gender for NHR (balance of noise to harmonics in the voice signal) revealed significantly less NHR in female subjects than male subjects ($F=5.431$, $p \leq 0.029$). No main effect of gender or language was identified for shimmer. A main effect of gender in cepstral/spectral analyses was observed for CPP of sustained vowel ($F=9.917$, $p \leq 0.006$) and CPP of the voiceless plosive sentence, “Peter will keep at the peak.” ($F = 4.45$, $p \leq 0.045$). In both of these variables, the group mean of CPP in male subjects was significantly higher than female subjects, indicating increased presence of harmonics in the vocal signal from participating male subjects. There was a significant main effect of gender for CSID of the voiceless plosive sentence of the CAPE-V ($F = 4.565$, $p \leq 0.043$), in which the group mean of male subjects was significantly lower than female subjects with a moderate effect size (.54). Main effect of gender was approximating significance for the CPP of the all-voice sentence of the CAPE-V ($F=4.067$, $p \leq 0.055$) and CSID of the all voiced sentence ($F=4.843$, $p \leq 0.038$).

For main effect of language, significance was observed for two variables, jitter ($F=3.145$, $p \leq 0.061$), and CPP of the all-voiced sentence ($F=5.099$, $p \leq 0.014$). Upon follow up for the CPP of the all-voiced sentence, post hoc for interaction of language revealed significance between dyads English and Spanish ($F= .481$, $p \leq 0.023$) and English and French (Std error .481, $p \leq 0.023$), shown in Table 5. Non-native speakers of SAE had significantly lower CPP results than SAE speakers. Interaction of language for French and Spanish speakers in CPP of all-voiced sentence was approximating significance (.481, $p \leq 0.0562$). In post hoc analyses for interaction of language for jitter (Table 5), native speakers of Spanish had significantly higher group mean results in perturbation of frequency (jitter) than native SAE speakers and French speakers (std err .311, $p \leq 0.034$). Before post hoc analysis, the interaction between SAE and French speakers washed out the significance demonstrated between dyads, English -Spanish speakers and Spanish- French speakers, in the initial ANOVA.

Table 4: ANOVA of Gender, Language, and Gender*Language

	Gender			Language			Gender * Language		
	df	F	p	df	F	p	df	F	p
Overall CAPEV	1	0.279	0.602	2	0.093	0.911	2	1.116	0.344
Roughness	1	0.810	0.377	2	0.097	0.908	2	0.518	0.602
Breathiness	1	1.946	0.176	2	1.027	0.373	2	1.135	0.338

Strain	1	1.243	0.276	2	0.602	0.556	2	0.951	0.400
F_0	1	112.913	0.000*	2	1.454	0.253	2	3.457	0.048*
Jitter	1	0.904	0.351	2	3.145	0.061+	2	0.969	0.394
Shimmer	1	0.111	0.742	2	1.849	0.179	2	0.516	0.604
NHR	1	5.431	0.029*	2	1.237	0.308	2	0.289	0.752
Speaking F_0	1	155.049	0.000	2	0.008	0.992	2	1.042	0.368
Vowel CPP	1	9.166	0.006*	2	1.243	0.306	2	1.437	0.257
Vowel CSID	1	0.129	0.723	2	1.782	0.190	2	0.046	0.955
Allvoiced CPP	1	4.067	0.055+	2	5.099	0.014*	2	0.893	0.423
Allvoiced CSID	1	4.843	0.038*	2	1.715	0.201	2	0.549	0.584
Easyonset CPP	1	1.569	0.222	2	0.079	0.924	2	0.515	0.604
Easyonset CSID	1	2.401	0.134	2	0.895	0.422	2	0.674	0.519
HardGlottal CPP	1	2.606	0.120	2	0.124	0.884	2	0.487	0.620
HardGlottal CSID	1	0.777	0.387	2	0.040	0.961	2	0.993	0.385
VoicelessPlosive CPP	1	4.454	0.045*	2	0.229	0.797	2	0.985	0.388
VoicelessPlosive CSID	1	4.565	0.043*	2	0.367	0.697	2	0.187	0.830

* Significant ($p \leq 0.05$) + Approaching significance

Table 5: Post Hoc Analyses for Language

Dependent Variable	Language Interactions		p
Allvoiced CPP	English	French	0.023*
	English	Spanish	0.006*
	Spanish	French	0.562

Jitter	English	French	0.866
	English	Spanish	0.048*
	Spanish	French	0.034*

* Significant ($p \leq 0.05$)

C. Accentedness

For main effect of accentedness, significant findings from previous ANOVAs and post hoc analyses were further investigated for influence of accent after averaging percentage of accentedness for each L1 French and Spanish speaker in an informal analysis (Table 6). A univariate ANOVA (Table 7) revealed significance for jitter ($F=4.702$, $p \leq 0.034$) and approaching significance for jitter ($F=3.074$, $p \leq 0.066$) and the all-voiced CPP only for language (Table 8). As previous literature indicates, significant differences exist for perturbation measures (jitter) between diverse sociolinguistic groups (Andrianopoulos et al., 2017, etc.).

Table 6: Means of Accentedness by Participant & Language

French Subjects			Spanish Subjects		
Participant	Gender	Accentedness	Participant	Gender	Accentedness
1	F	21.0%	1	F	12.5%
2	F	16.0%	2	F	28.5%
3	F	29.0%	3	F	16.5%
4	F	23.5%	4	F	15.0%
5	F	25.0%	5	F	27.5%
6	M	7.5%	6	M	5.5%
7	M	6.5%	7	M	22.5%

8	M	46.5%	8	M	21.5%
9	M	42.0%	9	M	16.0%
10	M	24.0%	10	M	20.0%

Table 7: ANOVA of Accentedness and Language

	Accentedness			Language		
	df	F	p	df	F	p
Jitter	1	0.58	0.457	1	4.702	0.045*

* Significant ($p \leq 0.05$)

Table 8: ANOVA of Accentedness, Gender, Language, Gender*Language

	Accentedness			Gender			Language			Gender * Language		
	df	F	p	df	F	p	df	F	p	df	F	p
Jitter	1	1.010	0.325	1	0.885	0.357	2	3.074	0.066+	2	1.129	0.341
Allvoiced CPP	1	1.154	0.294	1	4.047	0.056+	2	1.257	0.303	2	1.059	0.363

* Significant ($p \leq 0.05$) + Approaching significance

V. DISCUSSION

With the rise of diversity in the United States, cultural competence within service delivery becomes necessary to ensure cultural variables and language exposure are taken into consideration when providing care to patients and clients, as outlined by ASHA's views on linguistic and cultural competence (2017). To assess cultural appropriateness of clinically standard voice diagnostic tools, a composite of 30 male and female participants, native speakers of Standard American English, French, and Spanish, were recruited and participated in this study to ensure accuracy in objective measures of voice quality assessments of diverse sociolinguistic groups, and to assess possible relationship of auditory-perceptual assessment and acoustic data secondary to degree of accent. Acoustic analysis was completed with standard clinical software programs, MDVP and ADSV. Auditory-perceptual analysis of the voice quality and degree of accentedness were measured via the CAPE-V and an informal analysis with The Rainbow Passage, respectively. Three homogeneous groups of male and female subjects were selected for participation via relevant exclusionary and inclusionary criteria (age, native language, number of years of formal education in L1 country, age of spoken English proficiency, etc.). It was hypothesized that: 1) there would be no significant difference between monolingual speakers of Standard American English with native speakers of Spanish and French on cepstral and spectral analyses 2) there would be no significant difference between ratings on the auditory perceptual assessment with acoustic data secondary to accentedness to assess if perceptional analysis is different than objective analysis.

The results of the study primarily supported these hypotheses. Of the 19 experimental variables related to voice quality, significant differences for main effect of language were noted in two parameters (all-voiced sentence CPP & jitter). Native Standard American English (SAE) female and male speakers demonstrated significantly higher group mean CPP for the all-voiced sentence than native French and Spanish speakers. The all-voiced phonetic content of this stimulus provides the opportunity to assess the maintenance of vocalization ("linking") in a series of voiced sounds. In reverting to the original data for this sentence stimulus, 19 of the 20 results for native French and Spanish speakers were outside of the normative range (lower CPP) delineated for male and female subjects. Adding SAE speakers to this finding, 87 percent of the CPP all voiced sentence values were under the normative range. With researchers evidencing significant differences in F_0 and connected speech for bilingual proficient speakers (Ng, Chen, & Chan, 2012; Järvinen, Laukkanen, & Geneid, 2017; Bahmanbiglu, Mojiri, & Abnavi, 2017), and significant differences in composition of voice quality and timing of execution between languages, (i.e., breathiness and nasality) (Styler, 2017; Garellek, Ritchart, & Kuang (2016), cross linguistic variabilities among sociolinguistic groups influence the presented acoustic signal and must be taken in account during a comprehensive voice evaluation.

Limitations

Limitations of this study should be noted. First, while moderate effect size was appreciated between linguistic groups, a limited sample size was available per gender-language subset, (Males/Females (15): English (5) French (5) Spanish (5)) that reduced effect size to trivial and small in acoustic variables. Second, even though analysis of demographic data revealed no significant differences between the sociolinguistic groups, due to personal experiences unique to each individual's cultural and language background, heterogeneity in level of nonnative accent, especially in the expressive output of a L2 such as English was variable. Participant inclusion criteria also specified country of origin for native speakers of Spanish, French, and English to be Mexico, France, and the United States of America, respectively, though regional location was not specified, and no other sociolinguistic groups were represented in this study.

As explored previously, voice quality and uniqueness of accent are influenced by geographical, gender, social status, age, and cultural characteristics. Nonnative English participants were also excluded in the study criteria to individuals with a certain level of English language proficiency (ability to independently read aloud paragraphs and speak to a verbal prompt). The study therefore did not represent a complete spectrum of L2 speakers. Third, due to the structure and objective of the study, only one visit for participants was necessary to obtain recorded voice samples for acoustic and auditory-perceptual analysis. As vocal quality is adept to fluctuate throughout a day, let alone alter daily, voice measures collected in this study just offer a snapshot of vocal quality and performance for each participant.

VI. CONCLUSION

The aforementioned limitations notwithstanding, the results of this preliminary study suggest relative accuracy of objective measures of voice assessment in culturally and linguistically diverse populations with spoken English proficiency. While the literature reveals significant differences between diverse sociolinguistic sample populations in acoustic spectra tasks and in vocal tract parameters, acoustic and auditory perceptual measures of speech tasks in this study demonstrated at minimum levels of approaching significance in all voice parameters for main effect of language. Nevertheless, appropriate consideration of personal and patient/client cultural and linguistic backgrounds is warranted in therapeutic and diagnostic services.

Regarding clinical approach with diverse sociolinguistic populations, as part of the clinical voice evaluation, ideally measures of the assessment contribute to the differentiation of normophonic and dysphonic/disordered voice, and normative data is representative of all patients of voice complaints. Baseline measures obtained in subjective and objective voice assessments may therefore serve as a means of tracking absence or presence of progress for an individual with voice complaints throughout his/her plan of care, in addition to forming a comprehensive blueprint of an individual's voice quality in the initial voice evaluation.

APPENDIX A

Participant Questionnaire

Name: ______________________________
Age: _____
Gender: ___________________

Place of birth (city, country): _______________________________, _____________________

Native language and regional dialect spoken in the home, state, and country in which you resided in for elementary and secondary education: ______________________

Number of years of formal instruction in your native language: ______________________

Any other languages and dialects spoken (y/n): __________.
Which one(s)? ____

How much time have you lived in your native country? _______years _______months
How much time have you lived in the United States? _______years _______months

At what age did you begin learning English: ______________________________
If different than the answer above, at what age did you begin using English in spoken conversation: ____________________

Please indicate on the line below the amount of time you speak English versus French/Spanish throughout the week:

____

English French/Spanish

APPENDIX B

University of Houston
Accentedness Speech Sample

Participant Number: ______________________________
Clinician_____________________________
Date: ________________

Accentedness: 100 Word Sample ________% Accented Words

References

American Speech-Language-Hearing Association. (2017). *Issues in ethics: Cultural and linguistic competence.* Available from www.asha.org/Practice/ethics/Cultural-and-Linguistic-Competence/.

Andrianopoulos, M.V., Darrow, K.N., & Chen, J. (2001). Multimodal standardization of voice among four multicultural populations: Fundamental frequency and spectral characteristics. *Journal of Voice, 15*(2):194-219.

Asher, J., & Garcia, R. (1969). The optimal age to learn a foreign language. *Modern Language Journal, 53,* 334-341.

Awan, S.N. (2011). *Analysis of Dysphonia in Speech and Voice (ADSV): An Application Guide*. Montvale, NJ: Pentax Medical

Awan, S.N., & Mueller, P.B. (1996). Speaking fundamental frequency characteristics of white, African American, and Hispanic kindergartners. *Journal of Speech Language Hearing Research*, *39*, 573.

Awan, S., Roy, N., Jetté, M., Meltzner, G., & Hillman, R. (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. doi: 10.3109/02699206.2010.492446

Awan, S., Roy, N., & Cohen, S. (2014). Exploring the relationship between spectral and cepstral measures of voice and the Voice Handicap Index (VHI). *Journal Of Voice*, *28*(4), 430-439. doi: 10.1016/j.jvoice.2013.12.008

Basset, P., Amelot, A., Vaissière, J., & Roubeau. B. 2001). Nasal flow in French spontaneous speech. *Journal of the International Phonetic Association, 31*(1), 87-100.

Behlau, M., Yamaguchi, H., & Andrews, M. L. (2001). Escala GRBAS em três diferentes culturas. Proceedings of the 9th Brazilian Congress of Speech Language

Pathology and Audiology, Guarapari, ES, Brazil, 586

Carlson, H., & McHenry, M. A. (2006). Effect of accent and dialect on employability. *Journal of Employment Counseling, 43*(2), 70–83.

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages.* Cambridge, United Kingdom: Cambridge University Press.

Chaves, C., Campbell, M., & Côrtes Gama, A. (2017). The influence of native language on auditory-perceptual evaluation of vocal samples completed by Brazilian and Canadian SLPs. *Journal Of Voice*, *31*(2), 258.e1-258.e5. doi: 10.1016/j.jvoice.2016.05.021

Chen, H. C. (2010). Second language timing patterns and their effects on native listeners' perceptions. *Concentric: Studies in Linguistics, 36*, 183–212.

Coleman, R. (1971). Male and female voice quality and its relationship to vowel formant frequencies. *Journal Of Speech Language And Hearing Research*, *14*(3), 565. doi: 10.1044/jshr.1403.565

Corey, J., Gungor, A., Liu, X., Nelson, R., & Fredberg, J. (1998). Normative standards for nasal cross-sectional areas by race as measured by acoustic rhinometry. *Otolaryngology-Head And Neck Surgery, 119*(4), 389-393. doi: 10.1016/s0194-5998(98)70085-3

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition, 19,* 1–16.

Esling, J., & Wong, R. (1983). Voice quality settings and the teaching of pronunciation. *TESOL Quarterly*, *17*(1), 89. doi: 10.2307/3586426

Flege, J.E. (1988). Factors affecting degree of perceived foreign accent in English sentences. *The Journal Of The Acoustical Society Of America*, *84*(1), 70-79. doi:

10.1121/1.396876

Flege, J., Yeni-Komshian, G., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal Of Memory And Language*, *41*(1), 78-104. doi: 10.1006/jmla.1999.2638

Fuse, A., Navichkova, Y., & Alloggio, K. (2018). Perception of intelligibility and qualities of non-native accented speakers. *Journal Of Communication Disorders, 71*, 37-51. doi: 10.1016/j.jcomdis.2017.12.006

Ghio, A., Cantarella, G., Weisz, F., Robert, D., Woisard, V., Fussi, F., . . . Baracca, G. (2013). Is the perception of dysphonia severity language-dependent? A comparison of French and Italian voice assessments. *Logopedics Phoniatrics Vocology, 40*(1), 36-43. doi:10.3109/14015439.2013.837503

Giles, H. (1970). Evaluative reactions to accents. *Educational Review, 22,* 211-227.

Gillespie, A., Dastolfo, C., Magid, N., & Gartner-Schmidt, J. (2014). Acoustic analysis of four common voice diagnoses: Moving toward disorder-specific assessment. *Journal Of Voice*, *28*(5), 582-588. doi: 10.1016/j.jvoice.2014.02.002

Hirano, M (1981). *Clinical examination of voice.* Springer, New York.

Huber, J., Stathopoulos, E., Curione, G., Ash, T., & Johnson, K. (1999). Formants of children, women, and men: The effects of vocal intensity variation. *The Journal Of The Acoustical Society Of America*, *106*(3), 1532-1542. doi: 10.1121/1.427150

Kempster, G.B., Gerratt, B.R., Verdolini Abbott, K, et al. (2009). Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol. *American Journal of Speech Language Pathology. 18*, 124–132.

Lenneberg, E. (1967). *Biological Foundations of Language* (Wiley, New York).

Lowell, S. Y., Kelley R.T, Awan, S.N., Colton, R.H., & Chan, N. H. (2012) Spectral- and cepstral-based acoustic features of dysphonic, strained voice quality. *Annals of Otology Rhinology & Laryngology,* Vol. 121 (8) 539-548.

Major, R. C., Fitzmaurice, S. M., Bunta, F., & Balasubramanian, C. (2005). Testing the effects of regional, ethnic, and international dialects of English on listening comprehension. *Language Learning, 55,* 37-69.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45*(1), 73–97.

Ng, M., Chen, Y., & Chan, E. (2012). Differences in vocal characteristics between Cantonese and English produced by proficient Cantonese-English bilingual speakers—A Long-Term Average Spectral Analysis. *Journal Of Voice*, *26*(4), e171-e176. doi: 10.1016/j.jvoice.2011.07.013

Oyama, S. (1982). The sensitive period for the acquisition of a nonnative phonological system in child-adult differences in second language acquisition, edited by R. Scarcella and M. Long (Newbury House, Rowley, MA), pp. 20-38.

Patkowski, M. (1980). The sensitive period for the acquisition of syntax in a second language. *Language Learning, 30,* 449-471.

Piske, T., MacKay, I., & Flege, J. (2001). Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics*, *29*(2), 191-215. doi: 10.1006/jpho.2001.0134

Roy, N., Merrill, R. M., Thibeault, S., Gray, S. D., & Smith, E. M. (2004). Voice disorders in teachers and the general population. *Journal of Speech Language and Hearing Research, 47*(3), 542. doi:10.1044/1092-4388(2004/042)

Roy, N., Merril, R.N., Gray, S.D., Smith, E.M. (2005). Voice disorders in the general population: Prevalence, risk factors and occupational impact. *Laryngoscope, 115*, 1988–1995.

Roy, N., Barkmeier-Kraemer, J., Eadie, T., Sivasankar, M., Mehta, D., Paul, D., & Hillman, R. (2013). Evidence-based clinical voice assessment: A systematic review. *American Journal of Speech-Language Pathology*, *22*(2), 212. doi: 10.1044/1058-0360(2012/12-0014)

Titze, I. (1989). Physiologic and acoustic differences between male and female voices. *Journal Acoustic Society of America.* 85,1699-1707.

Trofimovich, P., & Baker, W. (2006). Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition*, 28 (1). doi: 10.1017/s0272263106060013

Tremblay, A., Broersma, M., Coughlin, C., & Choi, J. (2016). Effects of the native language on the learning of fundamental frequency in second-language speech segmentation. *Frontiers in Psychology*, *7*. doi: 10.3389/fpsyg.2016.00985

Scherer R.C. (1991). Physiology of phonation: a review of basic mechanics. In: Ford CN, Bless DM, eds. *Phonosurgery Assessment and Surgical Management of Voice Disorders.* New York, NY: Raven: 77-93.

Stuart-Smith, J. (1999) Glasgow: accent and voice quality. In: Foulkes, P. and Docherty, G.J. (eds.) *Urban voices: Accent Studies in the British Isles.*Arnold: Leeds, UK, pp. 201-222.

Styler, W. (2017). On the acoustical features of vowel nasality in English and French. *The Journal of the Acoustical Society of America, 142*(4), 2469-2482. doi:10.1121/1.5008854

United States Census Bureau (2017a). *Selected social characteristics in the United States: 2011-2015 American Community Survey selected population tables.* [Data file]. Retrieved from: https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_15_SPT_DP02&prodType=table

United States Census Bureau (2017b). *Age by language spoken at home by ability to speak English for the population 5 years and older.* [Data file]. Retrieved from: https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_15_SPT_B16004&prodType=table

Watts, C., Awan, S. (2011) Use of spectral/cepstral analyses for differentiating normal from hypofunctional voices in sustained vowel and continuous speech contexts. *Journal of Speech Language Hearing Research 54*, 1525-1537.

Watts, C., Awan, S., & Maryn, Y. (2017). A comparison of Cepstral Peak Prominence measures from two acoustic analysis programs. *Journal of Voice*, *31*(3), 387.e1-387.e10. doi: 10.1016/j.jvoice.2016.09.012

Wells, J. (2012). *Accents of English*. Cambridge [u.a.]: Cambridge Univ. Press.

Yamaguchi, H., Shrivastav, R., Andrews, M. L., & Niimi, S. (2003). A comparison of voice quality ratings made by Japanese and American listeners using the GRBAS Scale. *Folia Phoniatrica Et Logopaedica, 55*(3), 147-157. doi:10.1159/000070726

Xue, S., & Hao, J. (2006). Normative standards for vocal tract dimensions by race as measured by acoustic pharyngometry. *Journal of Voice*, *20*(3), 391-400. doi: 10.1016/j.jvoice.2005.05.001