AGE (2015) 37: 117
DOI 10.1007/s11357-015-9854-1

Effects of age on the amplitude, frequency and perceived quality of voice

Catherine L. Lortie · Mélanie Thibeault · Matthieu J. Guitton · Pascale Tremblay

Received: 12 May 2015 / Accepted: 3 November 2015 / Published online: 14 November 2015

Abstract The manner and extent to which voice amplitude and frequency control mechanisms change with age is not well understood. The related question of whether the assessment of one's own voice evolves with age, concomitant with the acoustical changes that the voice undergoes, also remains unanswered. In the present study, we characterized the aging of voice production mechanisms (amplitude, frequency), compared the aging voice in different experimental contexts (vowel utterance, connected speech) and examined the relationship between voice self-assessment and age-related voice acoustical changes. Eighty healthy adults (20 to 75 years old) participated in the study, which involved computation of several acoustical measures of voice (including measures of fundamental frequency, voice amplitude, and stability) as well as self-assessments of voice. Because depression is frequent in older adults, depression and anxiety scores were also measured. As was expected, analyses revealed age effects on most acoustical measures. However, there was no interaction between age and the ability to produce high/low voice amplitude/frequency, suggesting that voice amplitude and frequency control mechanisms are preserved in aging. Multiple mediation analyses demonstrated that the relationship between age and voice self-assessment was moderated by depression and anxiety scores. Taken together, these results reveal that while voice production undergoes important changes throughout aging, the ability to increase/decrease the amplitude and frequency of voice are preserved, at least within the age range studied, and that depression and anxiety scores have a stronger impact on perceived voice quality than acoustical changes themselves.

Keywords Voice · Aging · Amplitude control · Frequency control · Connected speech · Self-assessment

C. L. Lortie · P. Tremblay
Département de Réadaptation, Université Laval, Quebec City, Quebec, Canada

C. L. Lortie · M. J. Guitton
Département d'ophtalmologie et ORL - chirurgie cervico-faciale, Université Laval, Quebec City, Quebec, Canada

C. L. Lortie · M. J. Guitton · P. Tremblay (✉)
Centre de Recherche de l'Institut Universitaire en Santé Mentale de Québec (CRIUSMQ), 2601 chemin de la Canardière, Quebec City, Quebec G1J 2G3, Canada
e-mail: pascale.tremblay@fmed.ulaval.ca

M. Thibeault
Nuance Communications, Inc., Montréal, Quebec, Canada

Introduction

The ability to communicate our thoughts, opinions, and feelings verbally is a key component of social relationships, and integral to full participation in society at all ages. Communication relies on a healthy voice production system to express both complex ideas and emotions. However, the human voice undergoes significant perceptual and acoustic transformations with age. Some of these changes may negatively affect the communication process (Stathopoulos et al. 2011), and, in turn, the quality of life. Specifically, voice aging can have a negative impact on independence, integration, and effective communication (Kendall 2007; Plank et al. 2011). Indeed, voice-related effort and discomfort, combined with increased anxiety and frustration, can cause seniors to avoid social situations and withdraw from certain kinds of activities such as telephone conversations, or large parties (Verdonck-de Leeuw and Mahieu 2004; Roy et al. 2007; Etter et al. 2013).

Voice aging is caused by normal anatomical and physiological changes associated with this phase of life (Goy et al. 2013; Forero Mendoza et al. 2014). In particular, physiological changes occur in the larynx (Honjo and Isshiki 1980; Bloch and Behrman 2001; Ximenes Filho et al. 2003; Kersing and Jennekens 2004; Pontes et al. 2005, 2006; Sato et al. 2010, 2011), the vocal tract (Pontes et al. 2006), and the respiratory system (Linville 1996; Teles-Magalhães et al. 2000; Ramig et al. 2001; Sauder et al. 2010). These changes include calcification and ossification of cartilages, muscles and vocal fold atrophy, vocal fold bowing (i.e., inward curve) and reduced mucosal wave, and reduced pulmonary lung pressures, volumes, elasticity, and recoil. These changes have an impact on voice production and quality (Mazzetto de Menezes et al. 2014). Indeed, aging negatively affects voice stability and amplitude in a sex-dependent manner (Linville and Fisher 1985; Ma and Love 2010; Dehqan et al. 2013; Goy et al. 2013). For instance, men generally show a gradual increase in fundamental frequency (f0) with age (Honjo and Isshiki 1980; Harnsberger et al. 2008; Torre and Barlow 2009; Ma and Love 2010; Dehqan et al. 2013), while women exhibit a decrease in f0 with age (Honjo and Isshiki 1980; Torre and Barlow 2009; Ma and Love 2010; Da Silva et al. 2011; Dehqan et al. 2013; Goy et al. 2013). However, age-related changes in amplitude are less consistent across studies. Some studies report an amplitude decrease for both women and men (Baker et al. 2001; Da Silva et al. 2011), while other show it only for men (Morris and Brown 1994; Goy et al. 2013). The effect of age on voice perturbation measures is less clear: while some studies have shown that voice stability defined as jitter and shimmer declines with age (Wilcox and Horii 1980; Baken 2005; Dehqan et al. 2013), other studies did not find significant changes with age (Baker et al. 2001; Goy et al. 2013). Finally, while some studies have reported reduced maximum phonation time (MPT) in aging (Ptacek et al. 1966; Kreul 1972), recent reports observed no significant decrease in MPT values in older participants (Maslan et al. 2011). There is even one report of longer MPT values in older females compared with younger females (Goy et al. 2013).

From a perceptual point of view, the aged voice has been associated with increased hoarseness and breathiness, vocal fatigue, instability, and crackling (Kendall 2007; Gregory et al. 2012). However, most of the previous studies have focused so far on steady vowel utterances, a task that only bears limited resemblance with natural language production, which requires dynamic adjustments to voice frequency and amplitude. Thus, very little is known about the effect of age on the ability to control the frequency and amplitude of voice. Only two studies have examined voice amplitude control mechanisms in aging. The first study did not observe differences in the amplitude of a sustained vowel /a/ at normal and high amplitude in 30 young adults compared to 30 elderly women (young, 20–35 years old; older, 60–82 years old) (Mazzetto de Menezes et al. 2014). The second study did not observe difference in voice amplitude of a sustained /a/ at low, normal and high amplitude in 15 young compared to 14 old men (young, 20–26 years old; older, 56–71 years old) (Bier et al. 2014). To the best of our knowledge, no study has examined the aging of voice frequency control mechanisms. Therefore, the effect of age on voice amplitude and frequency control mechanisms is still largely unknown, and so is the effect of age on the production of voice in different contexts (sustained vowel, connected speech).

An important question that follows from the observation of changes in voice physical and acoustical properties is whether voice perception is directly affected by these changes. Studies that have examined the relationship between voice perception and voice acoustics have reported significant correlations between actual voice instability (i.e., jitter, shimmer and noise to harmonic ratio) and perceived breathiness and hoarseness on Hirano's grade index for perceptual voice assessment (GRBAS) (Eskenazi et al. 1990; Dejonckere et al. 1996; Wolfe and Martin 1997). Perceived roughness and tension was also significantly correlated to levels of harmonic energy and harmonic to noise ratios (HNRs) (Yanagihara 1967; Whitehead and Whitehead 1985). However, the listeners included in these studies were either experienced speech pathologists (Dejonckere et al. 1996; Wolfe and Martin 1997) or faculty members in Speech departments (Eskenazi et al. 1990). It is possible that the perception of one's own voice relies on different, and perhaps more subjective, criteria than those used by experts in the field judging other people's voice. In addition, other factors could also modulate the relationship between age and voice self-assessment. In particular, because elderly adults frequently experience depressive and anxious states, it is possible that mood affects the perception of one's own voice. Indeed, while depression ranges in prevalence from 10 to 20 % in the elderly (Steffens et al. 2009; Solhaug et al. 2012; Wu et al. 2012; Williams et al. 2015), the prevalence of anxiety ranges from 3 to 15 % (Kessler et al. 2005; Miloyan et al. 2014; Wassertheil-Smoller et al. 2014; Reynolds et al. 2015). Importantly, depression scores can change the way people perceive themselves on multiple levels, including their self-esteem (Fox 2000; Watson et al. 2002; Furegato et al. 2008; Sowislo and Orth 2012; Orth and Robins 2013; Legrand 2014; Wegener et al. 2015). Moreover, a few cross-sectional studies have reported negative moderate to strong correlations between self-esteem and anxiety (Watson et al. 2002; Riketta 2004; Lee and Hankin 2009; Sowislo and Orth 2012). Though it has never been examined, it is possible that anxiety and depression scores negatively affect voice self-assessment.

The objective of this study was thus to examine the effect of age on voice production, focusing on amplitude and frequency control mechanisms. We also compared conversational voice in different contexts (steady vowel utterances and connected speech). Finally, we also examined, for the first time, the relationships between age, voice acoustics, voice self-assessment, and depression and anxiety scores.

Methods

Participants

Eighty-one nonsmoking healthy adults (35 women) with normal or corrected-to-normal vision and no self-reported history of speech, voice, language, swallowing, hearing, severe respiratory restrain, neurological or neurodegenerative disorder, ranging in age from 20 to 75 years old (mean±standard deviation [SD] 54.63± 17.57) were recruited from the general community in Quebec City (QC, Canada) through e-mails, flyers, journal ads, and posters over the course of a 1-year period. One participant was excluded because he did not meet the inclusion criteria. The remaining 80 participants were included in the analysis. Participants were native speakers of Canadian French (17.76±3.5 years of education, range 12–29 years). A French version of the Mini Mental State Examination (MMSE) (Folstein et al. 1975; Hudon et al. 2009) confirmed that their cognitive functioning was within normal limits given their age (29.39±.88, range 25–30 points). Participants' characteristics are reported in Table 1. The procedures were approved by the Institutional Ethical Committee of the "Institut Universitaire en Santé Mentale de Québec" (protocol no. 353-2014) and the "CHU de Québec" (protocol no. C14-01-1908). Informed written consent was obtained from all participants, and they were compensated for their participation.

Voice and communication quality assessment

Voice-related quality of life was assessed for three categories of difficulties (physical, emotional and functional) using a French version of the Voice Handicap Index (VHI-30) (Jacobson et al. 1997; La voix 2006). A high score on the VHI indicates important voice difficulties. We translated and modified the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) (Zraick et al. 2011) in order to perform perceptual self-assessment of dysphonia in 6 points (i.e., voice strain, roughness, breathiness, etc.) on a 100 mm visual analog scale (0= not at all, 100=extremely). In this modified version of the CAPE-V, participants completed the assessment of their voice quality and filled the evaluation tool themselves, according to the instructions given by the examiner (C.L.L.). A high score again indicates voice dissatisfaction. Finally, the Bordeaux' Verbal Communication scale (ECVB) (Darrigrand and Mazaux 2000) was used to evaluate communication abilities (e.g., to express intentions, to hold a conversation, and to shop alone). A low score on the ECVB indicates communication difficulties.

Depression and anxiety scores

Participants were screened for depression and anxiety scores using the Hospital Anxiety and Depression Scale (HAD) (Zigmond and Snaith 1983).

Voice recordings

All recordings were performed by the same examiner (C.L.L.) under identical conditions in a calm room at the

Table 1 Participants' characteristics, for each age group and overall

N	Age		Education (in years)		HAD		MMSE		ECVB		VHI		Modified CAPE-V	
	Mean±SD	Range	Mean±SD	Range	Mean±SD	Range	Mean±SD	Range	Mean±SD	Range	Mean±SD	Range	Mean±SD	Range
20	28.6±6.26	20–39	18.75±4.83	12–29	7.05±2.98	1–12	29.65±0.5	29–30	60.15±4.73	51–66	4.05±5.36	0–23	8.75±4.52	2.3–18.9
28	55.39±8.74	40–65	17.79±3.48	12–25	6.89±6.18	1–33	29.32±1.22	25–30	60.86±4.9	42–66	5.79±9.26	0–47	11.2±8.24	0.2–32.4
32	70.22±2.77	66–75	17.13±2.32	13–22	4.66±2.22	1–12	29.28±0.68	28–30	59.91±4.25	50–66	2.03±2.78	0–12	7.52±5.6	0–20.1
80	54.63±17.57	20–75	17.76±3.5	12–29	6.04±4.29	1–33	29.39±0.88	25–30	60.3±4.56	42–66	3.85±6.48	0–47	9.11±6.56	0–32.4

HAD scale ranges from 0 to 42. For this test, scores between 0 and 16 are considered normal. The MMSE score ranges from 0 to 30 and a cutoff score of 23 optimizes sensitivity and specificity of detection of impairment (Murden et al. 1991; Uhlmann and Larson 1991)

HAD hospital anxiety and depression scale, *MMSE* mini-mental state examination, *ECVB* Bordeaux' verbal communication scale, *VHI* voice handicap index, *Modified CAPE-V* modified and self-administered version of the consensus auditory-perceptual evaluation of voice

CRIUSMQ, using a Shure headset microphone (Microflex Beta 53) placed at ±5 cm at a 45° angle to the subject's mouth to decrease aerodynamic noise from the mouth. A headset microphone was used to ensure that the distance between the microphone and the mouth was kept constant throughout the procedure and across participants. The microphone was connected to an Edirol U-25EX analog interface, which was in turn connected to a Toshiba PC through a USB port. The recordings were performed using the Audacity software at a sampling signal of 44.1 kHz and 32 bits of quantization. Throughout the experiment, water and short breaks were given to the participants as needed, and participants generally completed the session within 10 min.

Sustained vowel

The ability to control the amplitude and the frequency of voice can be assessed in different ways, for example during connected speech where they are analyzed as normal, online speech modulation (e.g., the lowest and highest frequency values during normal conversation). Another way to examine amplitude and frequency control mechanisms is to evaluate maximal capacity, either using an alternating crescendo/decrescendo method (Awan 2011; Maruthy and Ravibabu 2015) or an intersperse sustained vowel approach (Goy et al. 2013; Bier et al. 2014; Mazzetto de Menezes et al. 2014) (e.g., how low can my pitch go in absolute terms). Here, amplitude and frequency control mechanisms were assessed using the intersperse sustained vowel methods.

Participants were given two trials for each task. First, participants were asked to produce a sustained vowel /a/ at comfortable frequency and amplitude levels, i.e., under "normal talking voice" condition. Maximum phonation time (MPT) was calculated from the normal talking voice condition (Fig. 1). Next, participants were asked to produce the vowel /a/ for approximately 3 s in four other conditions: lowest amplitude (without whispering), highest amplitude (without yelling), lowest frequency, and highest frequency. Frequency and amplitude levels were self-determined by the participants in a manner similar to that implemented in previous studies (Bier et al. 2014; Mazzetto de Menezes et al. 2014). Participants were allowed to determine their own levels but they were encouraged to make a substantial difference between productions. A demonstration of expected amplitude and frequency levels was given to each

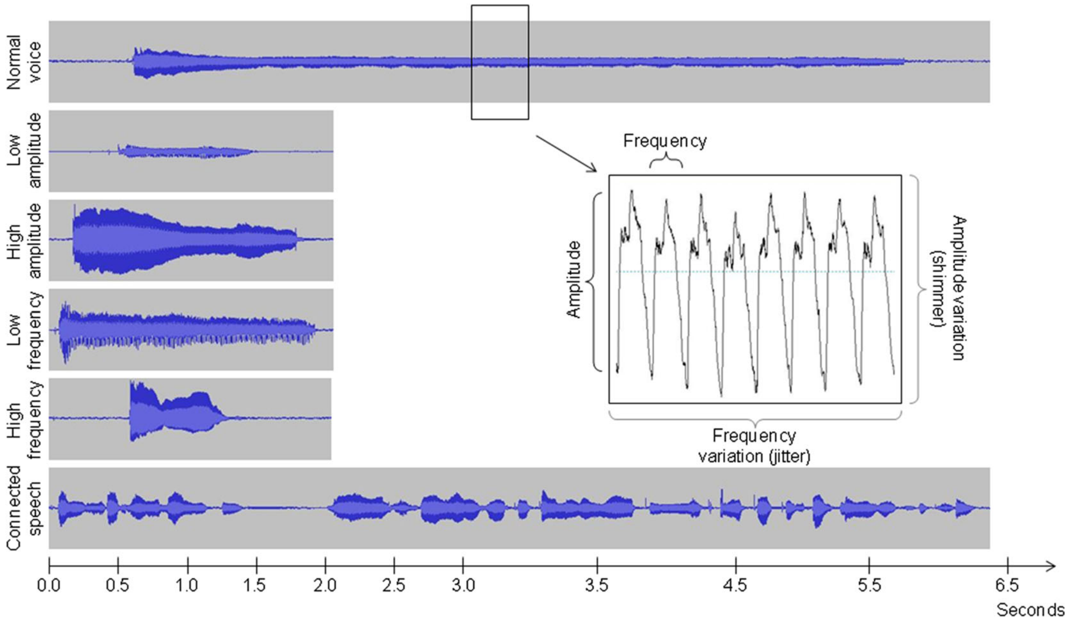

Fig. 1 Voice recordings. Examples of sustained vowels produced under normal voice, lowest amplitude (without whispering), highest amplitude (without yelling), lowest frequency and highest frequency, and connected speech. A representation of some of the acoustical measures extracted from the voice samples is also provided

participant by the same examiner (C.L.L.) under identical conditions by way of example. All vowels were produced as steadily as possible, with no amplitude or frequency variation. The task order was identical across all participants in order to avoid contamination effects across voice conditions (in particular from the high amplitude voice and high frequency voice conditions). Because of the very short duration of the procedure (less than 10 min), no fatigue or habituation effects were expected.

Connected speech

Participants were asked to narrate, using their own words, two popular story tales (i.e., "Red riding hood" and "Three little pigs") at comfortable frequency, amplitude, and rate (Fig. 1). The participants were given a representative illustration of the story tale to help recall. If the story tale was unknown to the participant, he/she was asked to describe the given illustration instead.

Acoustic analysis

Vocal signals were analyzed using the Praat software, version 5.3.39 (Boersma and Weenink 2012). The acoustical parameters used in this study are detailed in Table 2.

Sustained vowel

Original voice samples were visually inspected to identify passages with artifacts such as extraneous noise, laughter or coughing. These passages were excluded from the analysis. The analysis was then performed in two steps. First, the longest and most stable central segments of each vowel were manually selected. A Praat script was applied on that central section to automatically extract all acoustical measures, i.e., minimum f0 (Hz), maximum f0 (Hz), mean f0 (Hz) and f0 SD (semitones), mean amplitude and SD (dB), relative jitter (%), shimmer (dB), and HNR (dB), with the exception of duration. Minimum and maximum f0 target values were adjusted according to the sex of the speaker (men 65–300 Hz; women 80–550 Hz). These segments were visually inspected to correct f0 disruptions manually when necessary. Next, the whole voice sample was manually selected. A different Praat script was applied on the whole sample to automatically extract duration (MPT, seconds). For each participant, data from the two trials were averaged together for each acoustical measure.

Table 2 Acoustic measures extracted with Praat and their definition, along with the windowing and thresholds used to set internal Praat parameters in the scripts

Measure	Definition	Specific internal Praat parameters	
Minimum f0 (Hz)	Minimum fundamental frequency (i.e., number of glottic cycles per second)	To pitch: Time step, 0.0001 s Pitch floor, 65 Hz (men), 80 Hz (women) Pitch ceiling, 300 Hz (men), 550 Hz (women)	Time range, 0 to 0 (= all); unit, Hertz; interpolation, parabolic
Maximum f0 (Hz)	Maximum fundamental frequency		Time range, 0 to 0 (= all); unit, Hertz; interpolation, parabolic
Mean f0 (Hz)	Mean fundamental frequency		Time range, 0 to 0 (= all); unit, Hertz
F0 SD (semitones)	Fundamental frequency standard deviation		Time range, 0 to 0 (= all); unit, semitone
Mean amplitude (dB)	Mean sound pressure level	To intensity: Minimum pitch, min Time step, 0 (= auto) Substract mean, yes	Time range, 0 to 0 (= all); averaging method, dB
Amplitude SD (dB)	Sound pressure level standard deviation		Time range, 0 to 0 (= all)
Duration (MPT; s)	Duration of the voiced utterance	Start, get time of point 1; end, get time of point 2; duration, end–start	
Jitter local (%)	Absolute mean difference between consecutive periods, divided by the average period	Time range, 0 to 0 (= all) Shortest period, 0.0001 s Longest period, 0.02 s Maximum period factor, 1.3	
Shimmer local (dB)	Average absolute base-10 logarithm of the difference between the amplitudes of consecutive periods, multiplied by 20	Time range, 0 to 0 (= all) Shortest period, 0.0001 s Longest period, 0.02 s Maximum period factor, 1.3 Maximum amplitude factor, 1.6	
Harmonic to noise ratio (HNR, dB)	Degree of acoustic periodicity, i.e., the ratio between periodic (vocal fold vibration) and aperiodic (glottal noise) voice components (harmonicity of the voiced parts only)	Time step (s), 0.01 Minimum pitch, 75 Hz (men), 150 Hz (women) Silence threshold, 0.1 Periods per window, 1.0 Mean harmonicity: time range, 0 to 0 (= all)	

Connected speech

Original voice samples were first inspected to identify passages with artifacts such as extraneous noise, laughter or coughing, and caricatured voices. A 10-s central section of each sample was manually selected, avoiding such passages. All acoustical measures (i.e., minimum f0 (Hz), maximum f0 (Hz), mean f0 (Hz) and f0 SD (semitones), mean amplitude and SD (dB), and HNR (dB)) were extracted using a Praat script applied on these 10-s sections.

Statistical analyses

All data were analyzed using SPSS 22 (IBM SPSS Statistics, Armonk, NY). Acoustical measures (f0 minimum, maximum, mean and SD, mean and SD amplitude, MPT, jitter, shimmer, and HNR) were used as dependent measures for statistical analyses. For all statistical procedures, α=0.05 was used to establish significance. A false discovery rate (FDR) correction was applied on all post hoc analyses (Benjamini and Hochberg 1995). In the statistical analyses described below, age was used both as a continuous and a categorical independent variable. It was used as a categorical variable in the ANOVAs, in which participants were divided into three age groups (i.e., young, 20–39; middle-aged, 40–65; and older, 66–75 years old). In the moderated mediation analyses, age was used as a continuous variable.

Amplitude control

To assess age differences on the ability to produce high and low amplitude voice, a series of FDR-corrected (FDR per dependent variable [acoustical measures]: i=9, q=0.05) mixed model 3×3×2 ANOVAs on each of the acoustical parameter as dependent variables (f0min, f0max, f0mean, f0SD, mean amplitude, amplitude SD, jitter, shimmer, HNR) were first performed. For these analyses, sustained vowel amplitude (low, normal, high) was used as a within-subject factor, while age group (three levels, 20–39, 40–65, and 66–75 years) and sex were used as categorical between-subject factors. Post hoc *t* tests were conducted when appropriate.

Frequency control

To assess age differences on the ability to produce high- and low-frequency voice, a series of FDR-corrected (FDR per dependent variable [acoustical measures]: i=9, q=0.05) mixed model 3×3×2 ANOVAs on each of the acoustical parameter as dependent variables (f0min, f0max, f0mean, f0SD, mean amplitude, amplitude SD, jitter, shimmer, HNR) were performed. For these analyses, sustained vowel frequency (low, normal, high) was used as a within-subject factor, while age group (three levels, 20–39, 40–65, and 66–75 years) and sex were used as categorical between-subject factors. Post hoc *t* tests were conducted when appropriate.

Although our original intention was to study the interaction between amplitude and frequency control mechanisms, certain voice conditions (high amplitude at low frequency, low amplitude at high frequency, etc.) were very difficult to execute for nonprofessional speakers during pilot testing; therefore, interactions were not tested.

Connected speech

First, the acoustical measures computed from the two different story tales were compared using FDR-corrected (i=7, q=0.05) *t* tests for dependent samples to ensure that they did not differ significantly. Results demonstrated no significant difference between the acoustical measures across the speech samples; therefore, a mean was calculated for each acoustic parameter and used in further analyses. Next, to assess the effect of context on the voice, a series of FDR-corrected (FDR per dependent variable [acoustical measures]: i=7, q=0.05) mixed model 2×3×2 ANOVAs on acoustical measures (f0min, f0max, f0mean, f0SD, mean amplitude and SD, HNR) with context (sustained vowel, connected speech) as a within-subject variable, and age group (three levels, 20–39, 40–65, and 66–75 years) and sex as categorical between-subject factors were performed. Post hoc *t* tests were conducted when appropriate.

Moderated mediation analyses

In order to identify determinants of voice self-assessment, a conceptual model was developed in which the aging process affects voice acoustical measures, which in turn predicts voice self-assessment after controlling for sex. The indirect effect of age on voice self-assessment was hypothesized to be moderated by anxiety and depression scores. This conceptual model was tested in an operative framework, i.e., a mediated moderation, described in the following sections.

Moderation and mediation analyses allow researchers to examine the mechanisms by which variables affects each other (Baron and Kenny 1986; Shrout and Bolger 2002; MacKinnon et al. 2007; Preacher and Hayes 2008). Moderated mediation analyses estimate the path coefficients in a single mediator and a single moderator model and generate bootstrap confidence intervals for direct and indirect effect of X on Y through a mediator variable (*M*) conditional to a moderator (*W*). The model that was used is illustrated in Fig. 2. In this moderated mediation model, the dependent (*Y*) variable was the self-assessment measures of voice quality, while the independent (*X*) variable was the continuous variable age. One covariate (sex) was included in the model. Voice acoustics were used as the mediator (*M*) and anxiety or depression scores as the moderator (*W*). The moderated mediation analyses were conducted using the PROCESS macro (model no. 59) for SPSS (http://www.afhayes.com/) (Preacher et al. 2007; Preacher and Hayes 2008; Hayes and Preacher 2013).

For each self-assessment measure, a linear regression was used to test for a direct effect of age on voice self-assessment (the c' path in the model). Linear regressions were also conducted to test for an effect of age on voice acoustical measures (the a paths) and an effect of acoustical measures on voice self-assessment (the b paths). Next, a series of regressions were conducted, each including either the mediator, the moderator or both, to examine (1) whether there was an effect of age on voice self-assessment mediated by voice acoustics (the ab paths), (2) whether there was an indirect effect of depression and anxiety scores on voice self-assessment

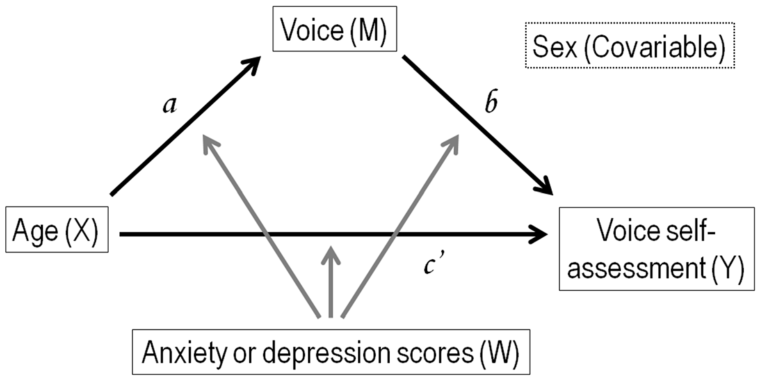

Fig. 2 Moderated mediation model. Conceptual moderated mediation model used to uncover the effect of age (*X*) on voice self-assessment (*Y*; the *c*' path), and whether this relationship was mediated by acoustic measures (*M*; the *ab* paths). This conceptual model was also used to test if depression and anxiety scores (*W*) moderated the effect of age on voice acoustics (the *a* path), the effect of voice acoustics on voice self-assessment (the *b* path), the effect of age on voice self-assessment mediated by voice acoustics (the *ab* paths), and the effect of age on voice self-assessment (*c*' path)

through age (c' paths), or through the effect of age on voice self-assessment mediated by voice acoustics (the ab paths), and (3) whether depression and anxiety scores moderated the effect of age on voice acoustics (the a paths) and the effect of voice acoustics on voice self-assessment (the b paths). A bootstrapping approach was used to test for the significance of the indirect effects (Shrout and Bolger 2002) (p=0.05, using bias-corrected bootstrapping with 10,000 samples). Bootstrapping involves the repeated extraction of samples, with replacement, from a dataset and the estimation of the indirect effect in each resampled dataset. The moderated mediation analyses were performed separately for each acoustic parameter (M; n=10 [9 acoustic measures and MPT]), self-assessment measures of voice (Y; n=2) and anxiety and depression scores (W; n=2), for a total of 40 moderated mediation analyses.

Finally, to assess age differences on psychological states (anxiety and depression), a series of independent t tests between age groups were performed with anxiety and depression scores as the dependent variables.

Results

Amplitude control

Significant main effects of age were found on all acoustic measures with the exception of maximum f0 ($F_{(2,74)}$=1.97, p=0.19), mean amplitude ($F_{(2,74)}$=0.02, p=0.98), and amplitude SD ($F_{(2,74)}$=0.13, p=0.99). The results also revealed significant main effects of amplitude on all voice measures ($F_{(2,74)}$=23.12 to 596.74, $p=1.83^{-9}$ to 1.3^{-70}). Significant main effects of sex were also observed on all acoustic measures with the exception of f0 SD ($F_{(1,74)}$=1.07, p=0.3), mean amplitude ($F_{(1,74)}$=3.19, p=0.09), and jitter ($F_{(1,74)}$=4.18, p=0.06). An interaction was observed between sex and amplitude on amplitude SD ($F_{(1,74)}$=7.01, p=0.01). There was no interaction between age and sex, age, and amplitude, or between amplitude, age, and sex.

First, we explored the main effects of age using post hoc analyses, which showed that f0 SD was higher for the older group [young vs. middle-aged: $t_{(46)}$=−2.68, p=0.01; young vs. older: $t_{(50)}$=−5.02, $p=1.1^{-5}$; middle-aged vs. older: $t_{(58)}$=−3.17, p=0.003]. The voice of the older group had higher jitter values than middle-aged ($t_{(58)}$=−2.1, p=0.04) and young adults ($t_{(50)}$=−2.98, p=0.005) (Fig. 3a). The older adults also had higher shimmer compared to the young ($t_{(50)}$=−2.86, p=0.006) (Fig. 3b). Finally, for HNR the young adults displayed higher values than the middle-aged ($t_{(46)}$=2.71, p=0.009) and the older adults ($t_{(50)}$=3.53, p=0.001) (Fig. 3c).

Next, we explored the main effects of amplitude using post hoc analyses, which showed, as expected, that mean amplitude significantly differed across the experimental conditions ($t_{(79)}$=10.23 to 30.57, $p=3.98^{-16}$ to 1.57^{-45}), with amplitude in the low amplitude condition being the softest, followed by normal and high amplitude voice (Fig. 4). Results revealed that the high amplitude voice had the expected highest frequency compared to normal ($t_{(79)}$=−8.52, $p=8.61^{-13}$) and low amplitude voices ($t_{(79)}$=−6.59, $p=4.58^{-9}$), similar to previous results by Gramming and colleagues

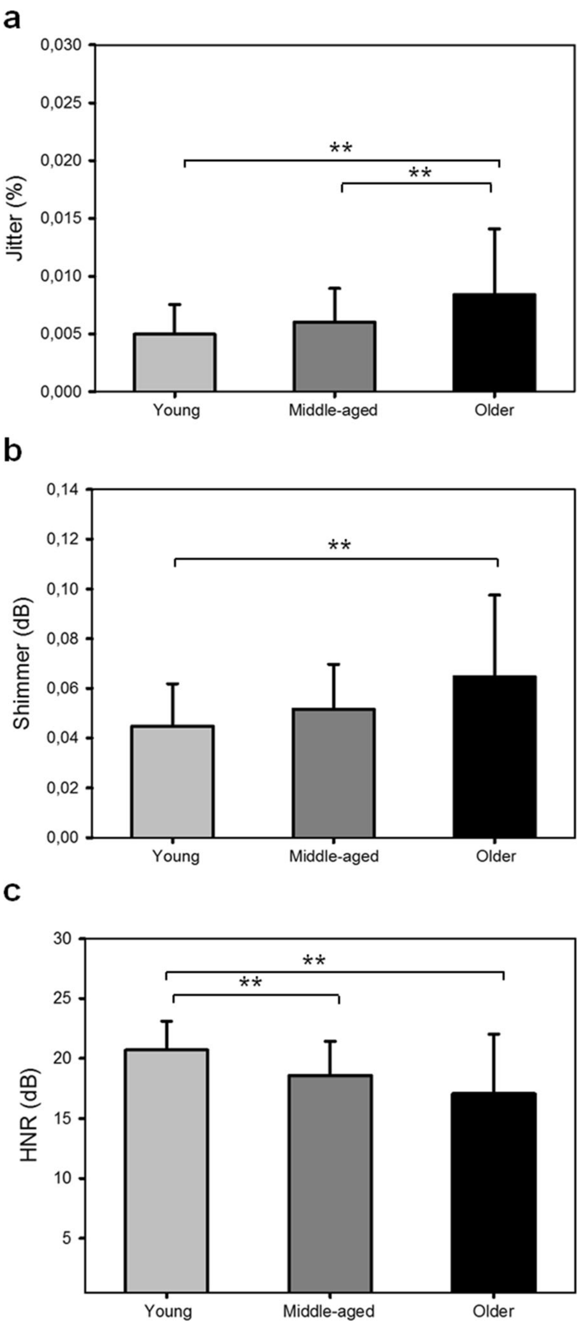

Fig. 3 Main effect of age on voice stability. Voice jitter (**a**), shimmer (**b**), and HNR (**c**) are displayed by age groups. *Single asterisks* indicate significance at $p<0.05$, while *double asterisks* indicate significance at $p<0.01$. *Error bars* represent the standard error of the mean

(Gramming et al. 1988). The high amplitude voice had lower jitter and shimmer values and a greater HNR value than normal voice ($t_{(79)}=6.67$ to 8.51, $p=3.14^{-9}$ to 8.66^{-13}), which in turn had lower jitter and shimmer and higher HRN values compared with the low amplitude voice ($t_{(79)}=4.24$ to 8.04, $p=6.1^{-5}$ to 7.39^{-12}).

Then, we explored the main effects of sex using post hoc analyses, which confirmed that women had significantly higher minimum f0 ($t_{(78)}=-14.88$, $p=1.81^{-19}$), maximum f0 ($t_{(78)}=-15.81$, $p=4.64^{-21}$), and mean f0 ($t_{(78)}=-15.46$, $p=2.75^{-20}$). Women had higher amplitude SD ($t_{(78)}=-2.35$, $p=0.02$) and higher HNR values ($t_{(78)}=-2.70$, $p=0.008$), but lower shimmer values than men ($t_{(78)}=3.31$, $p=0.001$).

To decompose the amplitude by sex interaction, post hoc analyses were performed on the amplitude SD of men and women separately. Results showed that men had the highest amplitude SD values in the normal voice condition [significantly higher than low amplitude ($t_{(44)}=-7.76$, $p=8.95^{-10}$) and high amplitude ($t_{(44)}=6.06$, $p=2.8^{-7}$)], meaning that normal voice was the least stable in amplitude. The normal voice of women was also the least stable in amplitude [amplitude SD in normal voice significantly higher than low amplitude ($t_{(34)}=-7.76$, $p=5^{-9}$) and high amplitude ($t_{(34)}=5.37$, $p=6^{-6}$)], but the amplitude effects were generally larger in men compared to women.

Frequency control

The repeated measures ANOVAs revealed significant main effects of age on three voice measures, i.e., f0 SD ($F_{(2,74)}=13.16$, $p=5.85^{-5}$), jitter ($F_{(2,74)}=5.25$, $p=0.02$), and HNR ($F_{(2,74)}=15.09$, $p=2.7^{-5}$). The results also revealed significant main effects of frequency on all acoustical measures ($F_{(2,74)}=11.98$ to 672.52, $p=1.5^{-5}$ to 3.17^{-74}). Significant main effects of sex were also observed on most acoustical measures with the exception of f0 SD ($F_{(1,74)}=0.08$, $p=0.78$), mean amplitude ($F_{(1,74)}=1.11$, $p=0.4$), jitter ($F_{(1,74)}=3.38$, $p=0.11$), and HNR ($F_{(1,74)}=0.79$, $p=0.42$). An interaction was observed between age and frequency but only for HNR ($F_{(4,74)}=6.26$, $p=9.9^{-4}$). Interactions were also observed between sex and frequency control on f0 minimum ($F_{(2,74)}=71.59$, $p=5.37^{-22}$), maximum ($F_{(2,74)}=73.89$, $p=5.03^{-22}$) and mean ($F_{(2,74)}=73.14$, $p=3.66^{-22}$), and on amplitude SD ($F_{(2,74)}=3.8$, $p=0.04$), jitter ($F_{(2,74)}=3.86$, $p=0.04$), and shimmer ($F_{(2,74)}=5.18$, $p=0.02$). Ultimately, there was a three-way interaction between age, sex, and frequency on HNR ($F_{(4,74)}=7.94$, $p=7.2^{-5}$). There were no interactions between age and sex.

For the main effects of frequency, post hoc analyses showed that frequency (f0 mean, minimum, and

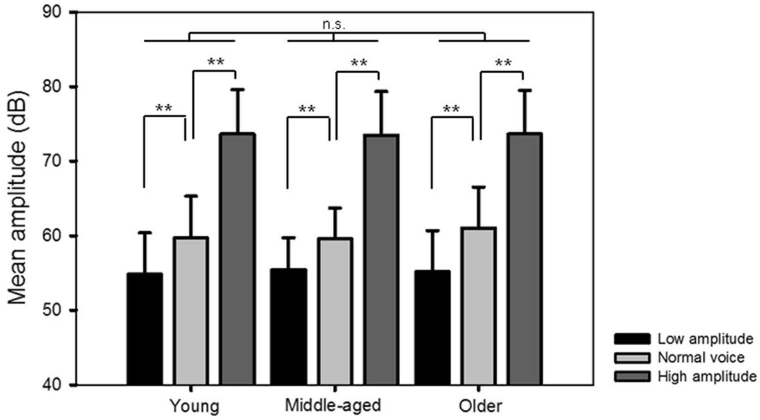

Fig. 4 Main effect of amplitude across conditions. Mean amplitude (dB) is displayed by age group and amplitude condition (low, normal, and high amplitude). *Single asterisks* indicate significance at $p<0.05$, while *double asterisks* indicate significance at $p<0.01$. *Error bars* represent the standard error of the mean. *n.s.* non-significant effect

maximum) in all three conditions were significantly different ($t_{(79)}=4.63$ to 18.71, $p=1.4^{-5}$ to 9.16^{-31}), with the expected linear increase from low to high frequency voice (Fig. 5). Moreover, the results revealed that the high frequency voice had the highest amplitude compared to the normal ($t_{(79)}=-13.74$, $p=1.22^{-22}$) and the low frequency voice ($t_{(79)}=-6.25$, $p=1.95^{-8}$). However, the low-frequency voice had higher amplitude compared to normal voice ($t_{(79)}=9.33$, $p=2.19^{-14}$). The high-frequency voice had lower jitter and shimmer values and a higher HNR value than normal voice ($t_{(79)}=6.16$ to 12.66, $p=2.93^{-8}$ to 1.05^{-20}), which in turn had lower shimmer and higher HNR values compared with the low frequency voice ($t_{(79)}=2.91$ and −2.14, $p=0.005$ and 0.04, respectively).

To decompose the frequency by age interaction on the HNR, post hoc analyses were performed of the three age groups separately. Results showed that for the youngest group (20–39 years) and the middle-aged group (40–65 years), HNR of the normal voice and the low-frequency voice did not differ ($t_{(19)}=.17$, $p=0.87$ and $t_{(27)}=-0.14$, $p=0.89$, respectively). However, for the oldest group (66–75 years), HNR values in all three conditions were significantly different, following a linear increase from low to normal voice ($t_{(31)}=-2.73$, $p=0.01$), and from normal to high-frequency voice ($t_{(31)}=-7.66$, $p=1.22^{-8}$).

For the frequency by sex interaction, post hoc analyses demonstrated that for f0 minimum, maximum, and mean, there was an effect of frequency for both men and women, but the frequency effects were generally larger in women compared to men, regardless of age. On the contrary, there was also an effect of frequency for both men and women on jitter values, but the frequency effects were generally larger in men compared to women. For one measure, amplitude SD, a different pattern was found. For men, the amplitude SD values of the normal voice were higher than both low-frequency ($t_{(44)}=-6.45$, $p=7.32^{-8}$) and high-frequency voice ($t_{(44)}=5.22$, $p=5^{-6}$). In contrast, amplitude SD values for women were similarly high for both normal and high-frequency voice ($t_{(34)}=0.21$, $p=0.84$), meaning

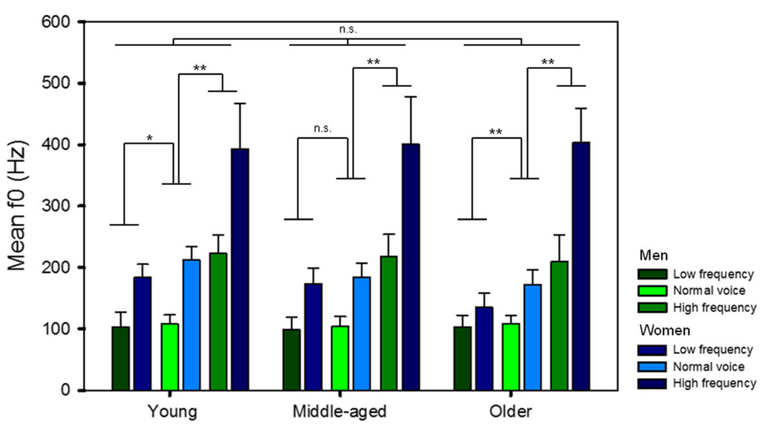

Fig. 5 Main effect of voice frequency for each condition. Mean frequency (Hz) is illustrated as a function of age group, frequency condition, and sex. *Single asterisks* indicate significance at $p<0.05$, while *double asterisks* indicate significance at $p<0.01$. *Error bars* represent the standard error of the mean. *n.s.* non-significant effect

that the normal voice was the least stable in amplitude for men but not women. Post hoc analyses also demonstrated that shimmer values for men decreased from low-frequency to normal voice ($t_{(44)}$=2.79, p=0.008), and from normal to high-frequency voice ($t_{(44)}$=6.51, p=6.08^{-8}). However, shimmer values for women were similarly high for both low-frequency and normal voice ($t_{(34)}$=1.1, p=0.28).

Finally, for the three-way interaction between age, sex, and frequency on HNR, post hoc analyses demonstrated that for the youngest group (20–39 years), the high-frequency voice had higher HNR values compared with the low-frequency voice for men ($t_{(10)}$=−4.58, p=0.001) but not for women ($t_{(8)}$=−0.76, p=0.47). For the middle-aged group, the high-frequency voice had higher HNR values compared with the low-frequency voice for both men and women ($t_{(13)}$=−5.67, p=7.7^{-5} and $t_{(13)}$=−3.72, p=0.003, respectively), and also higher HNR values compared with the normal voice for both men and women ($t_{(13)}$=−4.25, p=0.001 and $t_{(13)}$= −10.03, p=1.75^{-7}, respectively). However, for the oldest group (66–75 years), the HNR values for women in all three conditions were significantly different and followed a linear increase from low to normal voice ($t_{(11)}$=−3.69, p=0.004), and from normal to high-frequency voice ($t_{(11)}$=−3.99, p=0.002), whereas for men, the HNR values were similarly low for both low-frequency and normal voice ($t_{(19)}$=0.01, p=0.99).

Connected speech vs. sustained vowel

The repeated measures ANOVAs revealed significant main effects of context on all acoustical measures ($F_{(1,74)}$=38.93 to 744.45, p=2.48^{-8} to 1.63^{-39}). Significant main effects of sex were also observed on all acoustical measures with the exception of mean amplitude ($F_{(1,74)}$=0.003, p=0.96) and amplitude SD ($F_{(1,74)}$=1.83, p=0.21). Interactions were observed between age and context for minimum f0 ($F_{(2,74)}$=9.73, p=0.001), maximum f0 ($F_{(2,74)}$=3.97, p=0.04) and mean f0 ($F_{(2,74)}$=5.85, p=0.01), and amplitude SD ($F_{(2,74)}$= 4.58, p=0.03). Interactions were also observed between sex and context on f0 minimum ($F_{(1,74)}$=151.98, p= 9.03^{-19}), maximum ($F_{(1,74)}$=95.47, p=2.06^{-14}), and SD ($F_{(1,74)}$=15.95, p=0.0003). An interaction was observed between age and sex on mean f0 ($F_{(2,74)}$=4.61, p=0.045). Finally, a three-way interaction was observed between age, sex, and context on minimum f0 ($F_{(2,74)}$=6.52, p=0.01). There were no main effects of age.

Post hoc analyses were conducted to decompose the main effect of context, which showed that connected speech had a lower minimum f0 ($t_{(79)}$=12.81, p=5.57^{-21}) and a higher maximum f0 ($t_{(79)}$=−17.77, p=2.51^{-29}) and mean f0 ($t_{(79)}$=−6.72, p=2.58^{-9}) than sustained vowels. F0 SD was also higher in connected speech compared with sustained vowel ($t_{(79)}$=−24.94, p=3.38^{-39}), while HNR was lower in connected speech ($t_{(79)}$=13.73, p=1.22^{-22}). Mean amplitude was significantly lower in connected speech ($t_{(79)}$=11.66, p=7.55^{-19}), while amplitude SD was higher in connected speech compared with sustained vowel ($t_{(79)}$= −23.32, p=3.67^{-37}).

Another series of post hoc analyses were also conducted to decompose the context by age interaction on the f0 and the amplitude SD of the three age groups separately. Results showed that minimum f0 was lower in connected speech compared with sustained vowel for the young ($t_{(19)}$=−6.15, p=7^{-6}), middle-aged ($t_{(27)}$= −9.32, p=6.28^{-10}), and older group ($t_{(31)}$=−8.07, p=4.09^{-9}), but the difference between the two contexts was larger for the middle-aged group (Fig. 6a). In contrast, maximum f0 was higher in connected speech compared with sustained vowel for the young ($t_{(19)}$= 10.78, p=1.54^{-9}), middle-aged ($t_{(27)}$=11.77, p=3.88^{-12}), and older group ($t_{(31)}$=9.9, p=4.11^{-1}), but the difference between the two contexts was again larger for the middle-aged group (Fig. 6b). Mean f0 values followed a different pattern: mean f0 was higher in connected speech compared with sustained vowel for the middle-aged ($t_{(27)}$=3.87, p=0.001) and older group ($t_{(31)}$=6.45, p=3.4^{-7}), but similar between the two contexts for the young ($t_{(19)}$=1.15, p=0.26; Fig. 6c). Amplitude SD values were significantly higher in connected speech compared with sustained vowel for the young ($t_{(19)}$=14.85, p=6.59^{-12}), middle-aged ($t_{(27)}$= 16.28, p=1.75^{-15}), and older group ($t_{(31)}$=14.26, p=3.71^{-15}), but the difference between the two contexts was again larger for the middle-aged group (Fig. 6d). In particular, older adults had lower f0 and amplitude SD values compared with younger adults for sustained vowels, but f0 and amplitude SD values were similar between age groups for connected speech.

To decompose the context by sex interaction on f0, post hoc analyses were performed, which revealed that minimum f0 was lower in connected speech compared with sustained vowel for both men ($t_{(44)}$=−13.18, p=7.1^{-17}) and women ($t_{(34)}$=−15.79, p=3.23^{-17}), with a stronger effect of context for women. In addition, the

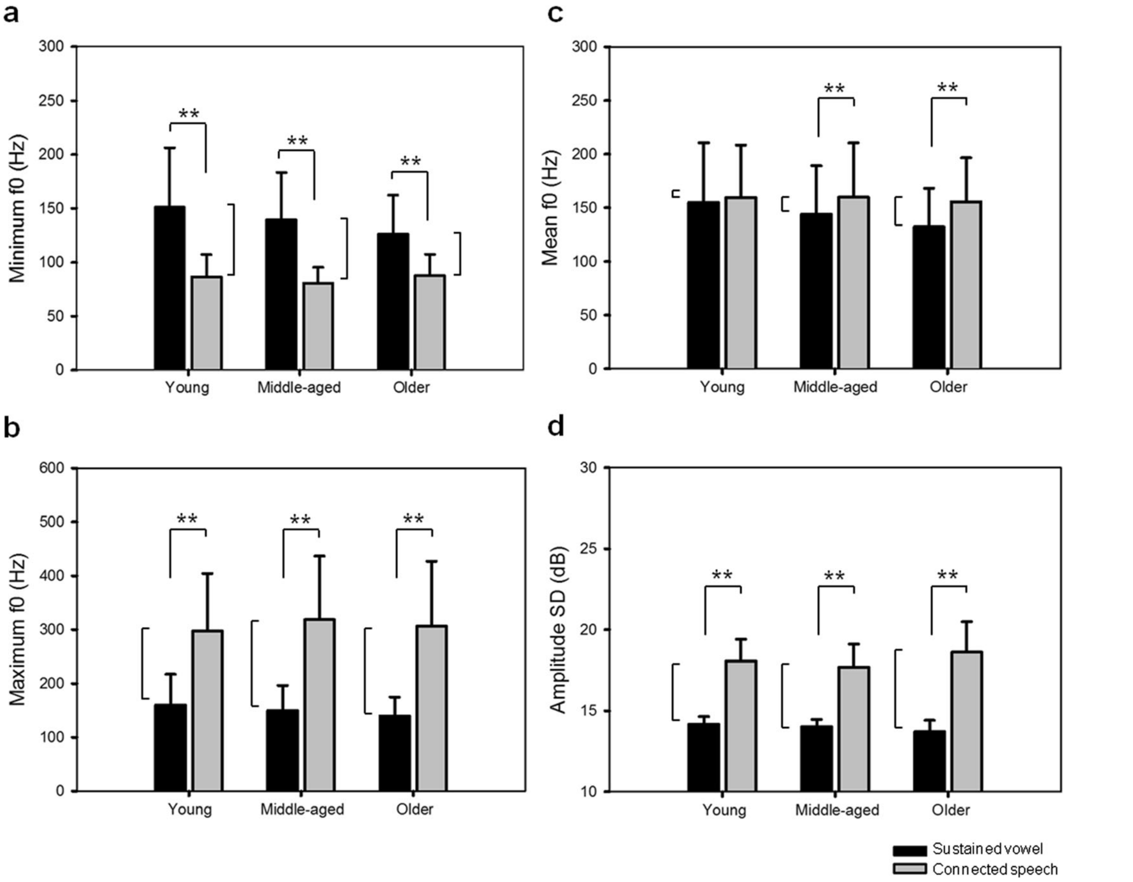

Fig. 6 Interaction between age and context on voice acoustic measures. Minimum f0 (**a**), maximum f0 (**b**), mean f0 (**c**), and amplitude SD (**d**) are displayed for each age group and each context. *Single asterisks* indicate significance at $p<0.05$, while *double asterisks* indicate significance at $p<0.01$. *Error bars* represent the standard error of the mean

maximum f0 was higher in connected speech compared with sustained vowel for both men ($t_{(44)}=16.31$, $p=2.88^{-20}$) and women ($t_{(34)}=20.72$, $p=7.27^{-21}$), with again a stronger effect of context for women. Similar results were observed for f0 SD, which was higher in connected speech compared with sustained vowel for both men ($t_{(44)}=19.16$, $p=5.67^{-23}$) and women ($t_{(34)}=19.2$, $p=8.01^{-20}$), with a slightly stronger effect of context for women.

For the age by sex interaction on mean f0, post hoc analyses demonstrated that there was no significant difference between the mean f0 of the three age groups neither for men nor women. However, the lowest mean f0 value for men was observed in the middle-aged group, whereas for women the lowest mean f0 value was observed in the oldest group.

Finally, to decompose the three-way context by age by sex interaction on minimum f0, post hoc analyses were performed. Results showed that, for both sex, minimum f0 was lower in connected speech compared with sustained vowel for the young (men: $t_{(10)}=-6.93$, $p=4^{-5}$; women: $t_{(8)}=-10.39$, $p=6^{-6}$), middle-aged (men: $t_{(13)}=-7.72$, $p=3^{-6}$; women: $t_{(13)}=-15.09$, $p=1.28^{-9}$) and the older groups (men: $t_{(19)}=-8.16$, $p=1.24^{-7}$; women: $t_{(11)}=-7.76$, $p=9^{-6}$). However, the difference between the contexts decreased with age only for women.

Moderated mediation analyses

The direct relationship between age and voice self-assessment, as assessed by the VHI and the modified CAPE-V, was not significant [c' path]. Depression and anxiety scores did not moderate the relationship between age and voice self-assessment [$W \rightarrow$ c' path]. Therefore, the age of a healthy adult did not affect his assessment of his own voice, and his anxiety or depression scores did not influence this relationship. Next, the direct effect of anxiety and depression scores on acoustic measures of the voice was examined [$W \rightarrow M$]. Results revealed that neither anxiety nor depression scores significantly influenced acoustics measures of voice. As expected, most acoustic measures were affected by sex, except mean amplitude [sex $\rightarrow M$]. However, sex had no influence on voice self-assessment [sex $\rightarrow Y$]. Thus, the sex of healthy adult did not affect their assessment of their own voice. The well-accepted notion that the aging process affects voice acoustical measures was again assessed after controlling for anxiety and depression scores. The analyses revealed that the effect of age on acoustic measures [a path] was still significant for some measures when anxiety scores were controlled (Fig. 7, Tables 3 and 5; age is associated with reduced minimum, maximum and mean f0, and increased f0 SD), and for other acoustic measures when depression scores were controlled (Fig. 8, Tables 4 and 6; age is associated with increased f0 SD, and decreased HNR). However, neither depression nor anxiety scores moderated the relationship between age and acoustics [$W \rightarrow$ a path]. Consequently, this analysis confirmed that aging affects voice acoustics and revealed that anxiety or depression scores did not alter the effects of age on voice acoustics.

In our conceptual framework, voice acoustics was hypothesized to predict voice self-assessment. Indeed, voice acoustics had a significant influence on voice self-assessment but only for a few measures [b path]. Specifically, high amplitude SD and MPT values were associated with high voice self-assessment scores, while high jitter and shimmer values were associated with a decline in self-assessment scores as assessed with the VHI when anxiety scores were controlled (Fig. 7, Table 3). High f0 SD and jitter values were associated with a decline in self-assessment as assessed with the VHI when depression scores were controlled (Fig. 8, Table 4). However, anxiety scores had a strong influence on the relationship between acoustics and voice self-assessment [$W \rightarrow$ b path] when assessed with the VHI. Indeed, the relationships between all acoustic measures (with the exception of f0 SD and mean amplitude) and the VHI were moderated by anxiety scores (Fig. 7, Table 3). In particular, the interaction between most voice acoustic measures (i.e., f0 minimum, maximum, and mean, amplitude SD, MPT, and HNR) and high anxiety scores were associated with a decline in voice self-assessment, while the interaction between the remaining voice acoustic measures (i.e., jitter and shimmer) and high anxiety scores were associated with an increase in voice self-assessment. Depression scores also moderated the relationship between half acoustic measures and the VHI (Fig. 8, Table 4; amplitude SD, MPT, jitter, shimmer, and HNR). In particular, the interaction between amplitude SD, MPT, HNR, and high

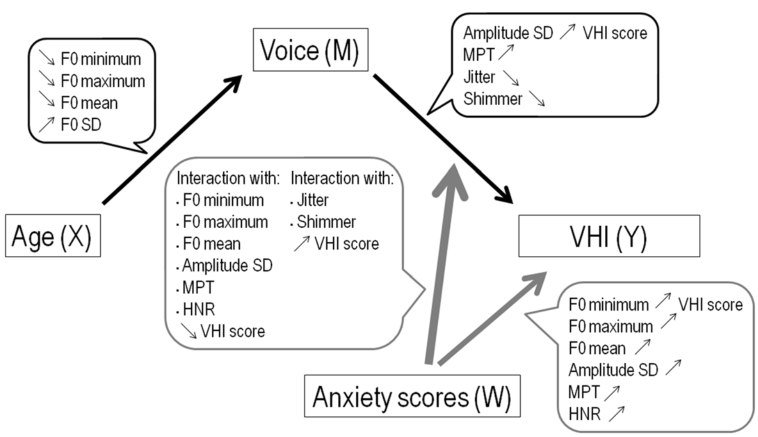

Fig. 7 Results of the moderated mediation analyses with anxiety scores as the moderator (W). The relationship between age and voice self-assessment was mediated by acoustic measures and moderated by anxiety scores. *From the left*: Advancing age was associated with low minimum, maximum and mean f0, and high f0 SD. High amplitude SD and MPT values were associated with high voice self-assessment, while high jitter and shimmer values were associated with low voice self-assessment. The interaction between most voice acoustic measures (i.e., f0 minimum, maximum, and mean, amplitude SD, MPT, and HNR) and high anxiety scores were associated with low voice self-assessment (indicating good voice satisfaction), while the interaction between the remaining voice acoustic measures (i.e., jitter and shimmer) and high anxiety scores were associated with high voice self-assessment (indicating low voice satisfaction). Finally, high anxiety scores were associated with high VHI scores (thus indicating low voice satisfaction) in six models out of ten. The size of the *arrows* corresponds to the quantity of acoustic measures involved in that relationship

Table 3 Results of the moderated mediation analyses with anxiety scores as the moderator (*W*) and VHI as the dependent (*Y*) measure

Paths	a		$W \rightarrow M$		$W \rightarrow$ a		Sex → M		b		c'	
Acoustical measures	β	p	β	p	β	p	β	p	β	p	β	p
F0 minimum	−0.7838	0.0086	−4.6421	0.1604	0.0959	0.0935	75.9296	0.0000	0.0681	0.0673	−0.0322	0.6674
F0 maximum	−0.6836	0.0216	−4.6869	0.1575	0.0930	0.1047	78.7128	0.0000	0.0579	0.1128	−0.0470	0.5248
F0 mean	−0.7139	0.0155	−4.4339	0.1792	0.0908	0.1092	77.3453	0.0000	0.0649	0.0821	−0.0383	0.6068
F0 SD	0.0031	0.0218	0.0017	0.9091	−0.0001	0.6905	−0.0459	0.0397	−23.4412	0.0973	0.0216	0.8009
Mean amplitude	0.0075	0.9174	−0.2881	0.7235	0.0058	0.6782	−0.6039	0.6153	−0.2277	0.3492	−0.0359	0.633
Amplitude SD	−0.0119	0.135	−0.022	0.8045	0.0004	0.7844	0.2694	0.0428	5.6474	0.0057	0.0201	0.7846
MPT	−0.0147	0.8658	−0.0523	0.9575	−0.0052	0.7575	−5.5255	0.0003	0.5611	0.0066	−0.0171	0.8008
Jitter	0.0001	0.3405	−0.0001	0.8326	0.0000	0.7539	−0.0026	0.0092	−588.5953	0.0061	0.0359	0.5982
Shimmer	0.0006	0.0659	0.0025	0.5075	0.0000	0.5599	−0.0161	0.0047	−79.7221	0.0425	−0.0208	0.7668
HNR	−0.0913	0.0731	−0.0155	0.9783	0.0026	0.793	2.0964	0.0142	0.4892	0.1024	0.0045	0.9536

Paths	$W \rightarrow$ b		$W \rightarrow Y$		$W \rightarrow$ c'		Sex → Y		$W \rightarrow$ a+b
Acoustical measures	β	p	β	p	β	p	β	p	β
F0 minimum	−0.0141	0.0048	2.3546	0.0235	0.0162	0.2546	0.2991	0.9030	0
F0 maximum	−0.0135	0.0052	2.2869	0.0272	0.0187	0.1895	0.8470	0.7369	0
F0 mean	−0.0137	0.0059	2.3176	0.0263	0.0172	0.2272	0.3768	0.8808	0
F0 SD	5.0037	0.1548	0.1738	0.8429	0.0011	0.9483	0.7195	0.582	0
Mean amplitude	0.0651	0.1942	−3.1356	0.2919	0.0101	0.4966	0.4534	0.7142	0
Amplitude SD	−1.2454	0.0006	18.5559	0.0005	−0.0002	0.991	1.0584	0.3744	0
MPT	−0.1524	0.0002	3.1097	0.0025	−0.0024	0.861	1.0171	0.4234	0
Jitter	187.6668	0.0000	0.3063	0.6748	−0.0102	0.452	1.5432	0.1736	0
Shimmer	28.0787	0.0002	−0.7597	0.3555	0.004	0.7651	1.5852	0.1825	0
HNR	−0.168	0.0135	4.4862	0.0123	−0.0009	0.9535	1.3839	0.2762	0

For each of the different path, the linear regressions coefficients (β) and significance (p) are reported and listed for each acoustical measure

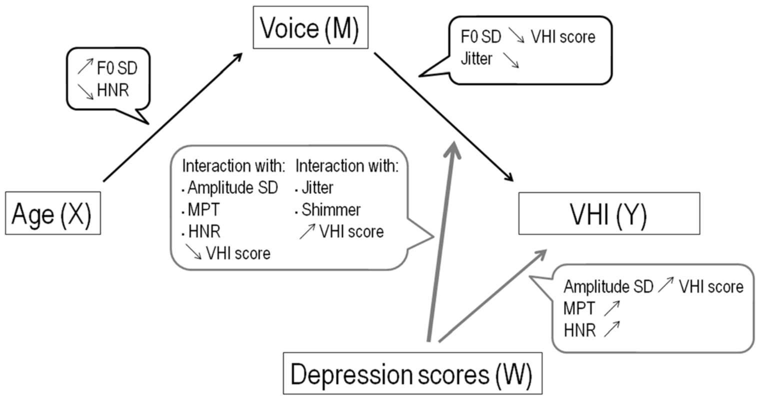

Fig. 8 Results of the moderated mediation analyses with depression scores as the moderator (W). The relationship between age and voice self-assessment was mediated by acoustic measures and moderated by depression scores. *From the left*: Advancing age increased f0 SD and decreased HNR. High f0 SD and jitter values were associated with low voice self-assessment. The interaction between amplitude SD, MPT, HNR, and high depression scores were associated with low voice self-assessment (indicating good voice satisfaction), while the interaction between jitter, shimmer, and high depression scores were associated with high voice self-assessment (indicating low voice satisfaction). Finally, depression scores had a limited direct influence on VHI scores (three models out of ten). The size of the *arrows* corresponds to the quantity of acoustic measures involved in that relationship

depression scores were associated with a decline in voice self-assessment, while the interaction between jitter and shimmer and high depression scores were associated with an increase in voice self-assessment (thus indicating low voice satisfaction). When voice was assessed with the modified CAPE-V, only the relationship between MPT and voice self-assessment was moderated by depression scores (Table 6). Indeed, the interaction between MPT and high depression scores were associated with a decline in voice self-assessment. In summary, these analyses indicate that the actual voice acoustic quality influenced voice self-assessment and that anxiety and depression scores have a strong influence on this relationship.

The direct effect of anxiety and depression scores on voice self-assessment was also examined [$W \rightarrow Y$]. Results revealed that high anxiety scores were associated with an increase in voice self-assessment as assessed with the VHI (thus indicating low voice satisfaction) in six models out of ten (Fig. 7, Table 3). In contrast, high depression scores had a limited influence on voice self-assessment as assessed with the VHI (three models out of ten; Fig. 8, Table 4). Finally, neither anxiety nor depression scores significantly influenced voice self-assessment as assessed with the modified CAPE-V (Tables 5 and 6). Finally, the relationship between age and self-assessment mediated by acoustic measures [ab paths] was moderated by anxiety scores [$W \rightarrow$ ab paths] in two contexts. Specifically, the higher the anxiety scores, the more the interaction between age and shimmer, and between age and HNR, was associated with a high voice self-assessment as assessed with the modified CAPE-V (thus indicating low voice satisfaction). All these results are detailed in Tables 3, 4, 5, and 6.

Anxiety and depression scores

Independent *t* tests on anxiety and depression scores revealed that the older adults had lower anxiety scores compared to the young ($t_{(30)}$=3.129, p=0.004; mean score 3.41±1.72 and 5.45±2.58, respectively) and middle-aged adults ($t_{(58)}$=2.026, p=0.047; mean score 3.41±1.72 and 4.89±3.73, respectively). No difference was observed between the anxiety levels of younger and middle-aged adults ($t_{(46)}$=0.576, p=0.567). For depression scores, older adults had similar scores compared to the young ($t_{(50)}$=0.856, p=0.396; mean score 1.25± 1.14 and 1.6±1.82, respectively) and middle-aged adults ($t_{(58)}$=1.38, p=0.173; mean score 1.25±1.14 and 2.0±2.83, respectively). No difference was observed between the level of anxiety of younger and middle-aged adults ($t_{(46)}$=−0.555, p=0.582). Of note, anxiety and depressive state levels of the three age groups were within normal limits.

Discussion

The primary goal of this study was to examine the effect of age on voice production, focusing on the ability to

Table 4 Results of the moderated mediation analyses with depression scores as the moderator (*W*) and VHI as the dependent (*Y*) measure

Paths	a		$W \rightarrow M$		$W \rightarrow$ a		Sex $\rightarrow M$		b		c'	
Acoustical measures	β	p	β	p	β	p	β	p	β	p	β	p
F0 minimum	−0.2404	0.182	6.3001	0.1725	−0.0821	0.3117	78.0597	0.0000	0.0348	0.3306	−0.0468	0.3465
F0 maximum	−0.1629	0.3645	5.7418	0.2133	−0.0714	0.3783	8.8135	0.0000	0.0166	0.6351	−0.0534	0.2759
F0 mean	−0.1957	0.2702	6.2396	0.1711	−0.0788	0.3247	79.3861	0.0000	0.0272	0.4485	−0.0506	0.3052
F0 SD	0.0025	0.003	−0.0117	0.5781	0.0002	0.6585	−0.0472	0.0305	−23.3907	0.0388	−0.0112	0.8397
Mean amplitude	0.0857	0.0507	2.1271	0.0584	−0.0352	0.0747	−0.3196	0.7791	−0.0409	0.8146	−0.0766	0.118
Amplitude SD	−0.0083	0.0899	0.0716	0.5637	−0.0012	0.5968	0.2834	0.0283	1.1876	0.3749	−0.0445	0.3684
MPT	0.0223	0.6768	1.9317	0.1607	−0.0337	0.1642	−5.3248	0.0003	0.2071	0.1344	−0.0249	0.5547
Jitter	0.0001	0.0693	0.0001	0.8733	0.0000	0.7404	−0.0025	0.0067	−366.7342	0.0365	−0.0192	0.6733
Shimmer	0.0003	0.201	−0.0057	0.2747	0.0001	0.1798	−0.0177	0.0013	−52.4093	0.0751	−0.0247	0.571
HNR	−0.0826	0.0094	0.1111	0.889	−0.0004	0.9773	2.101	0.0119	0.2404	0.2473	−0.0431	0.3867

Paths	$W \rightarrow$ b		$W \rightarrow Y$		$W \rightarrow$ c'		Sex $\rightarrow Y$		$W \rightarrow$ a+b
Acoustical measures	β	p	β	p	β	p	β	p	β
F0 minimum	−0.0125	0.0747	2.0838	0.248	0.0236	0.2953	−0.7935	0.7677	0
F0 maximum	−0.0117	0.0897	2.1014	0.2498	0.0241	0.2815	0.4942	0.8574	0
F0 mean	−0.0121	0.0828	2.071	0.2529	0.0242	0.2811	−0.2772	0.9197	0
F0 SD	7.3965	0.1252	−0.8822	0.47	0.0169	0.5051	−0.1738	0.8899	0
Mean amplitude	0.0722	0.3906	−4.9624	0.3584	0.0414	0.0608	0.0078	0.995	0
Amplitude SD	−0.9984	0.0304	14.7029	0.0362	0.0136	0.561	0.6636	0.5997	0
MPT	−0.2153	0.0000	5.0561	0.0017	−0.0173	0.4234	0.0107	0.9929	0
Jitter	26.4734	0.0001	−0.6457	0.5541	−0.0031	0.8831	1.1763	0.3261	0
Shimmer	49.1688	0.0000	−0.9566	0.378	−0.0125	0.5612	1.2336	0.2959	0
HNR	−0.2385	0.0124	5.4285	0.0345	0.0117	0.6094	1.1066	0.3938	0

For each of the different path, the linear regressions coefficients (β) and significance (p) are reported and listed for each acoustical measure

Table 5 Results of the moderated mediation analyses with anxiety scores as the moderator (*W*) and modified CAPE-V as the dependent (*Y*) measure

Paths	a		$W \rightarrow M$		$W \rightarrow$ a		Sex $\rightarrow M$		b		c’	
Acoustical measures	β	p	β	p	β	p	β	p	β	p	β	p
F0 minimum	−0.7838	0.0086	−4.6421	0.1604	0.0959	0.0935	75.9296	0.0000	0.0062	0.8919	−0.1268	0.177
F0 maximum	−0.6869	0.0216	−4.6869	0.1575	0.093	0.1047	78.7128	0.0000	0.0079	0.8617	−0.1274	0.1688
F0 mean	−0.7139	0.0155	−4.4339	0.1762	0.0908	0.1092	77.3453	0.0000	0.0049	0.9148	−0.1278	0.1698
F0 SD	0.0031	0.0218	0.0017	0.9091	−0.0001	0.6905	−0.0459	0.0397	−4.7717	0.7773	−0.1024	0.3222
Mean amplitude	0.0075	0.9174	−0.2881	0.7235	0.0058	0.6782	−0.6039	0.6153	0.0691	0.8133	−0.1195	0.19
Amplitude SD	−0.0119	0.135	−0.022	0.8045	0.0004	0.7844	0.2694	0.0428	−0.3166	0.9017	−0.1208	0.2046
MPT	−0.0147	0.8658	−0.0523	0.9575	−0.0052	0.7575	−5.5255	0.0003	0.2638	0.3192	−0.1013	0.2567
Jitter	0.0001	0.3405	−0.0001	0.8326	0.0000	0.7539	−0.0026	0.0092	−71.8739	0.8027	−0.0997	0.2885
Shimmer	0.0006	0.0659	0.0025	0.5075	0.0000	0.5599	−0.0161	0.0047	36.9085	0.4572	−0.1457	0.1074
HNR	−0.0913	0.0731	−0.0155	0.9783	0.0026	0.793	2.0964	0.0142	−0.0353	0.9223	−0.1157	0.2272

Paths	$W \rightarrow$ b		$W \rightarrow Y$		$W \rightarrow$ c’		Sex $\rightarrow Y$		$W \rightarrow$ a+b
Acoustical measures	β	p	β	p	β	p	β	p	β
F0 minimum	−0.0058	0.3367	−0.0544	0.9659	0.0286	0.1073	0.9391	0.7589	0
F0 maximum	−0.0059	0.3179	−0.0262	0.9835	0.0292	0.1005	0.8672	0.7823	0
F0 mean	−0.0057	0.3454	−0.0619	0.9613	0.0289	0.1034	1.012	0.746	0
F0 SD	3.0491	0.4692	−0.9478	0.3707	0.0186	0.3624	−0.0761	0.9614	0
Mean amplitude	−0.0076	0.8991	−0.2497	0.9443	0.026	0.1469	−0.6173	0.6791	0
Amplitude SD	−0.1291	0.7726	1.1513	0.8602	0.0246	0.177	−0.3468	0.8208	0
MPT	−0.0721	0.1571	0.5098	0.6964	0.0178	0.3231	−0.4108	0.8049	0
Jitter	47.3075	0.3757	−0.7488	0.4565	0.0193	0.2998	−0.1886	0.9032	0
Shimmer	7.9005	0.3896	−1.2149	0.2487	0.0251	0.1471	0.6146	0.6844	Yes (+)
HNR	−0.0749	0.3578	1.0502	0.6237	0.0196	0.2956	0.2783	0.857	Yes (+)

For each of the different path, the linear regressions coefficients (β) and significance (p) are reported and listed for each acoustical measure

Table 6 Results of the moderated mediation analyses with depression scores as the moderator (W) and modified CAPE-V as the dependent (Y) measure

Paths	a		$W \rightarrow M$		$W \rightarrow$ a		Sex $\rightarrow M$		b		c'	
Acoustical measures	β	p	β	p	β	p	β	p	β	p	β	p
F0 minimum	−0.2404	0.182	6.3001	0.1725	−0.0821	0.3117	78.0597	0.0000	0.0023	0.9591	−0.0626	0.3061
F0 maximum	−0.1629	0.3645	5.7418	0.2133	−0.0714	0.3783	8.8135	0.0000	−0.001	0.9806	−0.0632	0.2936
F0 mean	−0.1957	0.2702	6.2396	0.1711	−0.0788	0.3247	79.3861	0.0000	−0.0017	0.9694	−0.0639	0.2917
F0 SD	0.0025	0.003	−0.0117	0.5781	0.0002	0.6585	−0.0472	0.0305	0.1312	0.9924	−0.0676	0.3271
Mean amplitude	0.0857	0.0507	2.1271	0.0584	−0.0352	0.0747	−0.3196	0.7791	0.1769	0.4013	−0.079	0.1804
Amplitude SD	−0.0083	0.0899	0.0716	0.5637	−0.0012	0.5968	0.2834	0.0283	−0.6034	0.716	0.0706	0.2519
MPT	0.0223	0.6768	1.9317	0.1607	−0.0337	0.1642	−5.3248	0.0003	0.2921	0.1103	−0.0252	0.6503
Jitter	0.0001	0.0693	0.0001	0.8733	0.0000	0.7404	−0.0025	0.0067	−66.6377	0.7723	−0.0523	0.3906
Shimmer	0.0003	0.201	−0.0057	0.2747	0.0001	0.1798	−0.0177	0.0013	37.043	0.3424	−0.0663	0.2574
HNR	−0.0826	0.0094	0.1111	0.889	−0.0004	0.9773	2.101	0.0119	−0.1575	0.5404	−0.0774	0.2124

Paths	$W \rightarrow$ b		$W \rightarrow Y$		$W \rightarrow$ c'		Sex $\rightarrow Y$		$W \rightarrow$ a+b
Acoustical measures	B	p	β	p	β	p	β	p	β
F0 minimum	−0.0033	0.7018	−0.4851	0.8261	0.0284	0.3062	−0.2625	0.9367	0
F0 maximum	−0.0033	0.6963	−0.4463	0.8416	0.0284	0.3013	0.0069	0.9984	0
F0 mean	−0.003	0.7206	−0.5041	0.82	0.0286	0.3	0.0163	0.9961	0
F0 SD	1.4776	0.8033	−1.213	0.4232	0.0277	0.3787	−0.3938	0.8004	0
Mean amplitude	−0.0688	0.4972	2.9814	0.6468	0.0339	0.1995	−0.4627	0.758	0
Amplitude SD	−0.1356	0.8099	0.9393	0.9129	0.0285	0.3275	−0.2743	0.8612	0
MPT	−0.1945	0.0042	3.4942	0.0915	−0.0134	0.6364	−0.0701	0.9648	0
Jitter	94.5819	0.249	−1.2851	0.3778	0.0175	0.5409	−0.0028	0.9986	0
Shimmer	15.6635	0.2693	−1.0549	0.4663	0.0101	0.7254	0.7471	0.6345	0
HNR	−0.0773	0.5056	0.7323	0.8157	0.0241	0.3971	0.2579	0.8726	0

For each of the different path, the linear regressions coefficients (β) and significance (p) are reported and listed for each acoustical measure

control voice amplitude and frequency. Our results suggest that these abilities are preserved in aging, at least within the age range that we studied (20–75 years), in the context of sustained and steady vowel production (i.e., no online modulation). We also compared conversational voice in different contexts (steady vowels and connected speech), which revealed an age difference in the effect of context on voice production. The second objective of this study was to examine the relationships between age, voice acoustics, voice self-assessment, and depression and anxiety scores. Finally, our results show, for the first time, that anxiety and depression scores moderate the relationships between voice production in aging and voice self-assessment. These results are detailed in the following paragraphs.

Effects of age on voice acoustic measures

The results of the present study confirm that the aging voice is less stable and noisier. Specifically, we showed that jitter values significantly increased with age, consistent with previous findings (Wilcox and Horii 1980; Linville and Fisher 1985; Brown et al. 1989; Orlikoff 1990; Dehqan et al. 2013; Goy et al. 2013).We also observed a decrease in HNR values with age, consistent with previous studies proposing HNR as an indicator of voice aging (Decoster and Debruyne 1997; Ferrand 2002). In addition, voice stability decreased with age as measured by the f0 SD and shimmer, which is consistent with a previous study (Goy et al. 2013). Perturbation of both frequency and amplitude could result from an increasing amount of total vocal fold stiffness, as illustrated by the Ishizaka–Flanagan model described by Baken (2005). However, since no anatomical measures were made in the present study, it is not possible to attribute changes in the acoustics of the voice to precise anatomical alterations.

We also confirmed several effects of age on voice acoustic measures, including a significant decrease in F0 values in women (Honjo and Isshiki 1980; Torre and Barlow 2009; Ma and Love 2010; Da Silva et al. 2011; Stathopoulos et al. 2011; Dehqan et al. 2013; Goy et al. 2013), and no change in f0 values in men (Stathopoulos et al. 2011; Goy et al. 2013). For MPT, consistent with previous reports, we showed that it was unaffected by age (Maslan et al. 2011; Goy et al. 2013), even though MPT should depend upon several factors such as, respiratory volumes, airflow rate, task comprehension, and maximal effort, factors that one would expect to decline with age. However, there is limited evidence suggesting that MPT is not systematically associated with either airflow or vital respiratory capacity (Solomon et al. 2000).

Voice control mechanisms

Although most acoustical measures varied as a function of age, there was no interaction between age and amplitude, suggesting that amplitude control mechanisms are preserved in older adults, at least in the age range that we studied (20 to 75 years). This result is consistent with two previous studies that demonstrated that elderly women (Mazzetto de Menezes et al. 2014) and men (Bier et al. 2014) can produce high and low amplitude voice in aging. Muscular and phonatory compensation might explain the lack of an interaction between age and amplitude. Indeed, elderly adults display a greater expansion of the chest and lungs and more abdominal movement than necessary when asked to increase vocal amplitude (Stathopoulos and Sapienza 1997; Baker et al. 2001; Huber and Spruill 2008). They also initiate phonation at a higher lung volume and use greater percent of their lung volume per syllable (Hoit and Hixon 1987; Hoit et al. 1989; Sperry and Klich 1992), and they have a more pressed pattern of glottal vibration than young adults (Ahmad et al. 2012). The use of these compensatory mechanisms may therefore allow them to produce low and high amplitude voice, especially when produced in isolation, as assessed in the present study.

Our results also indicated that all groups were capable of producing distinct low, normal and high-frequency voice. This suggests that frequency control mechanisms may also be preserved in older adults, at least within the age range that we studied. The one exception was HNR, which was significantly lower in older adults but only for the low-frequency voice compared with the normal voice, a pattern that was not found in young and middle-aged groups. Older adults are able to produce subglottal pressure changes (Stathopoulos et al. 2011; Goy et al. 2013; Maruthy and Ravibabu 2015). Because raised subglottal pressure also yields an increase in fundamental frequency (Gramming et al. 1988), it is possible that older adults rely on subglottal pressure changes to produce different voice frequency levels. These results are particularly important in relation to differential voice diagnostic in an aging population. Indeed, difficulty to produce high or how voice frequency and/or amplitude may be an indicator of abnormal aging. It will however remain important to examine the voice of older adults (80+) to determine whether amplitude and control frequency mechanisms deteriorate later in life, and to investigate potential interactions between voice frequency and amplitude control mechanisms. In addition, it will be important to examine the effect of age on the ability to modulate voice frequency *online*, which could show a different aging trajectory. That is, it is possible that older adults are capable of producing high- and low-frequency voice separately, but they may have difficulty modulating the frequency of their voice online, as they speak, which was not assessed as part of the present study. Future studies should also consider longitudinal protocols in order to track the progression of voice and speech characteristics with age (Hunter et al. 2012), as well as incorporate measures of hearing, which may contribute to the relationship between age and voice production/self-assessment.

Connected speech

One of the most important finding of this study is an age difference in voice acoustics across contexts (sustained vowel versus connected speech). Indeed, our results demonstrated an interaction between age and context on voice acoustics measures: the difference between sustained vowel and connected speech was stronger for middle-aged and older adults compared to young speakers. In particular, older adults had lower minimum, maximum and mean f0 and amplitude SD values compared with younger adults for sustained vowels, but similar f0 and amplitude SD values compared with younger adults for connected speech. Because speech fundamental frequency is due to the physiological characteristics of the vocal folds and control of the larynx musculature (Honjo and Isshiki 1980; Bloch and Behrman 2001; Baken 2005; Sato et al. 2011), the absence of change observed in mean and range speech fundamental frequency here suggests that the state of the tissue and general motor control did not deteriorate with age (at least in the age range studied), or did not deteriorate enough to have a functional impact. These results indicate that, while sustained vowel production was affected negatively by age, connected speech was less vulnerable to age. Previous studies have examined the effect of age on voice acoustics using connected speech, comparing young to older adult performances within each context (Goy et al. 2013; Bier et al. 2014; Watts et al. 2015). However, no other study, to our knowledge, has examined differences between contexts (sustained vowel, connected speech) among age groups.

The type of voice and/or speech task is known to significantly influence voice acoustics, with connected speech associated with more variations and a broader range of frequencies and amplitude (Winkworth et al. 1994; Sapienza and Stathopoulos 1995; Hollien et al. 1997; Bohnenkamp et al. 2002; Zraick et al. 2004, 2005). It has been argued that in order to improve the ecological validity of voice assessments, perceptual and instrumental assessments of voice in sustained vowels and continuous speech contexts should be conducted (Rabinov et al. 1995; Revis et al. 1999; Carding et al. 2000; Maryn et al. 2010; Goy et al. 2013). Because of its large variations in f0 and harmonicity, connected speech may provide a more reliable estimate of a person's intonation in a naturalistic setting (Baken 1987; Yiu et al. 2000; Bhuta et al. 2004). Variability in connected speech may reflect the continuous changing vocal fold settings and conditions of voice onset and offset that occur in real-life situations (Fourcin and Abberton 2008; Fourcin 2009). And yet, connected speech is not the preferred signal for objective measurement of voice, in part because it is more difficult to elicit and because it is intrinsically a less controlled measure. However, in the last years, innovative analyses techniques for connected speech such as *cepstral peak prominence* (an acoustic measure of voice quality correlated with dysphonia severity, which integrates several measures describing the aperiodicity and waveform of the acoustic voice signal) have facilitated and popularized measures on connected speech in voice clinics (Heman-Ackah et al. 2003; Maryn et al. 2010; Fraile and Godino-Llorente 2014; Watts et al. 2015). The present finding adds to the growing literature suggesting that including connected speech measures to the standard clinical assessments could be useful in obtaining otherwise inaccessible data on voice.

Voice self-assessment

Given the changes that the human voice undergoes throughout normal aging, the second objective of this study was to examine the factors that affect the perception of one's voice in aging. Two types of factors were examined: the acoustical properties of the voice, as well as anxiety and depression scores. Our results demonstrate, for the first time, that anxiety and depression scores in healthy, nonclinically anxious or depressed adults moderated the relationships between age, voice acoustics and voice self-assessment, meaning that anxiety and depression scores are associated with an increase in the negative effect that some voice instability measures have on voice self-assessment. These results thus support the hypothesis that both depression and anxiety scores play an important role in voice self-assessment.

Depression and anxiety scores have been widely associated with poor self-esteem levels (Watson et al. 2002; Riketta 2004; Lee and Hankin 2009; Sowislo and Orth 2012; Orth and Robins 2013) and weaker self-rated health and perceived life satisfaction (Rouch et al. 2014). However, the study of the effects of depression and anxiety scores on self-assessment typically does not include an assessment of voice quality, despite the key importance of voice on human interactions. The only studies that examined voice self-assessment in relation with psychological states studied populations with severe voice disabilities and voice pathologies. They observed that negative voice self-assessment and speech in patients who suffered partial or total laryngectomies was associated with higher levels of depression and anxiety and withdrawal from social activities (Devins et al. 1994; Nalbadian et al. 2001; Birkhaug et al. 2002; de Maddalena 2002; Hanna et al. 2004; Boscolo-Rizzo et al. 2008; Danker et al. 2010). Our results demonstrate that the relationship between voice self-assessment and psychological states also holds in the normal population. This suggests that, because of the importance of voice quality in daily communication and social interactions, depression and anxiety scores have a significant influence on the perception of one's voice quality. It might therefore be useful to assess depression and anxiety scores systematically in future studies of voice self-assessment, as well as in people with a voice complaint.

Conclusion

By characterizing maximal voice abilities in terms of frequency and amplitude in young and older healthy adults, the current study increase the knowledge base on the normal aging of the human voice, and holds the potential to inform clinical practice. Since acoustic analyses are widely used in voice evaluation protocols, a thorough understanding of how the voice is affected by internal and external factors such as age, context, the ability to modulate voice frequency and amplitude, and psychological state is critical to improve diagnostic and care for people with voice disorders.

Acknowledgments The authors thank all the participants. This study was supported by grants from the "Fonds de la Recherche du Québec Société-Culture" (FRQ-SC) and the "Fonds de la Recherche du Québec en Santé" (FRQS) to P.T. and from a graduate scholarship from the Natural Sciences and Engineering Research Council of Canada (NSERC) to C.L.L. P.T. and M.J.G. hold Career Awards from the "Fonds de Recherche du Québec – Santé" (FRQS).

Compliance with ethical standards

Conflict of interest All authors report no conflict of interest and no constraints on publishing.

References

Ahmad K, Yan Y, Bless D (2012) Vocal fold vibratory characteristics of healthy geriatric females—analysis of high-speed digital images. J Voice. doi:10.1016/j.jvoice.2011.12.002

Awan SN (2011) The effect of smoking on the dysphonia severity index in females. Folia Phoniatr Logop 63:65–71. doi:10.1159/000316142

Baken RJ (1987) Clinical measurement of speech and voice. College-Hill Press, Boston

Baken RJ (2005) The aged voice: a new hypothesis. J Voice 19:317–325

Baker KK, Ramig LO, Sapir S et al (2001) Control of vocal loudness in young and old adults. J Speech Lang Hear Res 44:297–305

Baron RM, Kenny DA (1986) The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol 51:1173–1182

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B 57:289–300. doi:10.2307/2346101

Bhuta T, Patrick L, Garnett JD (2004) Perceptual evaluation of voice quality and its correlation with acoustic measurement. J Voice 18:299–304. doi:10.1016/j.jvoice.2003.12.004

Bier SD, Watson CI, McCann CM (2014) Using the perturbation of the contact quotient of the EGG waveform to analyze age differences in adult speech. J Voice 28:267–273. doi:10.1016/j.jvoice.2013.10.021

Birkhaug EJ, Aarstad HJ, Aarstad AKH, Olofsson J (2002) Relation between mood, social support and the quality of life in patients with laryngectomies. Eur Arch Otorhinolaryngol 259:197–204

Bloch I, Behrman A (2001) Quantitative analysis of videostroboscopic images in presbylarynges. Laryngoscope 111:2022–2027

Boersma P, Weenink D (2012) Praat: doing phonetics by computer. University of Amsterdam, Amsterdam

Bohnenkamp TA, Andrews M, Shrivastav R, Summers A (2002) Changes in children's voices: the effect of cognitive cues. J Voice 16:530–543

Boscolo-Rizzo P, Maronato F, Marchiori C et al (2008) Long-term quality of life after total laryngectomy and postoperative radiotherapy versus concurrent chemoradiotherapy for laryngeal preservation. Laryngoscope 118:300–306

Brown WS, Morris RJ, Michel JF (1989) Vocal jitter in young adult and aged female voices. J Voice 3:113–119

Carding P, Carlson E, Epstein R et al (2000) Formal perceptual evaluation of voice quality in the United Kingdom. Logoped Phoniatr Vocol 25:133–138

Da Silva PT, Master S, Andreoni S et al (2011) Acoustic and long-term average spectrum measures to detect vocal aging in women. J Voice 25:411–419

Danker H, Wollbrück D, Singer S et al (2010) Social withdrawal after laryngectomy. Eur Arch Otorhinolaryngol 267:593–600. doi:10.1007/s00405-009-1087-4

Darrigrand B, Mazaux JM (2000) L'Échelle de communication verbale de Bordeaux (E.C.V.B.). Ortho Édit, Isbergues

De Maddalena H (2002) The influence of early speech rehabilitation with voice prostheses on the psychological state of laryngectomized patients. Eur Arch Otorhinolaryngol 259: 48–52

de Menezes M, Keyla S, Master S, Guzman M et al (2014) Differences in acoustic and perceptual parameters of the voice between elderly and young women at habitual and high intensity. Acta Otorrinolaringol Engl Ed 65:76–84. doi:10.1016/j.otoeng.2013.11.012

Decoster W, Debruyne F (1997) The ageing voice: changes in fundamental frequency, waveform stability and spectrum. Acta Otorhinolaryngol Belg 51:105–112

Dehqan A, Scherer RC, Dashti G et al (2013) The effects of aging on acoustic parameters of voice. Folia Phoniatr Logop 64: 265–270

Dejonckere PH, Remacle M, Fresnel-Elbaz E et al (1996) Differentiated perceptual evaluation of pathological voice quality: reliability and correlations with acoustic measurements. Rev Laryngol Otol Rhinol 117:219–224

Devins GM, Stam HJ, Koopmans JP (1994) Psychosomal impact of laryngectomy mediated by perceived stigma and illness intrusiveness. Can J Psychiatry 39:608–616

Eskenazi L, Childers DG, Hicks DM (1990) Acoustic correlates of vocal quality. J Speech Hear Res 33:298–306

Etter NM, Stemple JC, Howell DM (2013) Defining the lived experience of older adults with voice disorders. J Voice 27: 61–67

Ferrand CT (2002) Harmonics-to-noise ratio: an index of vocal aging. J Voice 16:480–487

Folstein MF, Folstein SE, McHugh PR (1975) Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 12:189–198

Forero Mendoza LA, Cataldo E, Vellasco MMBR et al (2014) Classification of vocal aging using parameters extracted from the glottal signal. J Voice 28:532–537. doi:10.1016/j.jvoice.2014.02.001

Fourcin A (2009) Aspects of voice irregularity measurement in connected speech. Folia Phoniatr Logop 61:126–136. doi:10.1159/000219948

Fourcin A, Abberton E (2008) Hearing and phonetic criteria in voice measurement: clinical applications. Logoped Phoniatr Vocol 33:35–48. doi:10.1080/14015430701251574

Fox KR (2000) The effects of exercise on self perceptions and self-esteem. Int J Sport Psychol 31:228–240

Fraile R, Godino-Llorente JI (2014) Cepstral peak prominence: a comprehensive analysis. Biomed Signal Process Control 14: 42–54

Furegato ARF, Santos JLF, Da Silva EC (2008) Depression among nursing students associated to their self-esteem, health perception and interest in mental health. Rev Lat Am Enfermagem 16:198–204. doi:10.1590/S0104-11692008000200005

Goy H, Fernandes DN, Pichora-Fuller MK, van Lieshout P (2013) Normative voice data for younger and older adults. J Voice 27:545–555. doi:10.1016/j.jvoice.2013.03.002

Gramming P, Sundberg J, Ternström S et al (1988) Relationship between changes in voice pitch and loudness. J Voice 2:118–126

Gregory ND, Chandran S, Lurie D, Sataloff RT (2012) Voice disorders in the elderly. J Voice 26:254–258. doi:10.1016/j.jvoice.2010.10.024

Hanna E, Sherman A, Cash D et al (2004) Quality of life for patients following total laryngectomy vs chemoradiation for laryngeal preservation. Arch Otolaryngol Head Neck Surg 130:875–879

Harnsberger JD, Shrivastav R, Brown WS et al (2008) Speaking rate and fundamental frequency as speech cues to perceived age. J Voice 22:58–69

Hayes AF, Preacher KJ (2013) Statistical mediation analysis with a multicategorical independent variable. Br J Math Stat Psychol 67:451–470

Heman-Ackah YD, Heuer RJ, Michael DD et al (2003) Cepstral peak prominence: a more reliable measure of dysphonia. Ann Otol Rhinol Laryngol 112:324–333

Hoit JD, Hixon TJ (1987) Age and speech breathing. J Speech Hear Res 30:351–366

Hoit JD, Hixon TJ, Altman ME, Morgan WJ (1989) Speech breathing in women. J Speech Hear Res 32:353–365

Hollien H, Hollien PA, de Jong G (1997) Effects of three parameters on speaking fundamental frequency. J Acoust Soc Am 102:2984–2992

Honjo I, Isshiki N (1980) Laryngoscopic and voice characteristics of aged persons. Arch Otolaryngol 106:149–150

Huber JE, Spruill J (2008) Age-related changes to speech breathing with increased vocal loudness. J Speech Lang Hear Res 51:651–668

Hudon C, Potvin O, Turcotte M-C et al (2009) Normalisation du mini-mental state examination (MMSE) chez les Québécois francophones âgés de 65 ans et plus et résidant dans la communauté. Can J Aging La Rev Can du Vieil 28:347. doi:10.1017/S0714980809990171

Hunter EJ, Kapsner-Smith M, Pead P et al (2012) Age and speech production: a 50-year longitudinal study. J Am Geriatr Soc 60:1175–1177. doi:10.1111/j.1532-5415.2012.03983.x

Jacobson BH, Johnson A, Grywalski C et al (1997) The voice handicap index (VHI): development and validation. Am J Speech Lang Pathol 6:66–69

Kendall K (2007) Presbyphonia: a review. Curr Opin Otolaryngol Head Neck Surg 15:137–140. doi:10.1097/MOO.0b013e328166794f

Kersing W, Jennekens FG (2004) Age-related changes in human thyroarytenoid muscles: a histological and histochemical study. Eur Arch Otorhinolaryngol 261:386–392. doi:10.1007/s00405-003-0702-z

Kessler RC, Berglund P, Demler O et al (2005) Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. Arch Gen Psychiatry 62:593–602

Kreul EJ (1972) Neuromuscular control examination (NMC) for Parkinsonism: vowel prolongations and diadochokinetic and reading rates. J Speech Hear Res 15:72–83

La voix (2006) La voix - Ses troubles chez les enseignants. Inserm, Paris

Lee A, Hankin BL (2009) Insecure attachment, dysfunctional attitudes, and low self-esteem predicting prospective symptoms of depression and anxiety during adolescence. J Clin Child Adolesc Psychol 38:219–231

Legrand FD (2014) Effects of exercise on physical self-concept, global self-esteem, and depression in women of Low socio-economic status with elevated depressive symptoms. J Sport Exerc Psychol 36:357–365

Linville SE (1996) The sound of senescence. J Voice 10:190–200

Linville SE, Fisher HB (1985) Acoustic characteristics of perceived versus actual vocal age in controlled phonation by adult females. J Acoust Soc Am 78:40–48

Ma EPM, Love AL (2010) Electroglottographic evaluation of Age and gender effects during sustained phonation and connected speech. J Voice 24:146–152

MacKinnon DP, Fairchild AJ, Fritz MS (2007) Mediation analysis. Annu Rev Psychol 58:593–614

Maruthy S, Ravibabu P (2015) Comparison of dysphonia severity index between younger and older carnatic classical singers and nonsingers. J Voice. doi:10.1016/j.jvoice.2014.05.001

Maryn Y, Corthals P, Van Cauwenberge P et al (2010) Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels. J Voice 24:540–555

Maslan J, Leng X, Rees C et al (2011) Maximum phonation time in healthy older adults. J Voice 25:709–713

Miloyan B, Byrne GJ, Pachana NA (2014) Threshold and subthreshold generalized anxiety disorder in later life. Am J Geriatr Psychiatry 1–9. doi: 10.1016/j.jagp.2014.08.010

Morris RJ, Brown WS (1994) Age-related differences in speech intensity among adult females. Folia Phoniatr Logop 46:64–69

Murden RA, McRae TD, Kaner S, Bucknam ME (1991) Mini-mental state exam scores vary with education in blacks and whites. J Am Geriatr Soc 39:149–155

Nalbadian M, Nikolaou A, Nikolaidis V et al (2001) Factors influencing quality of life in laryngectomized patients. Eur Arch Otorhinolaryngol 258:336–340

Orlikoff RF (1990) The relationship of age and cardiovascular health to certain acoustic characteristics of male voices. J Speech Hear Res 33:450–457

Orth U, Robins RW (2013) Understanding the link between Low self-esteem and depression. Curr Dir Psychol Sci 22:455–460

Plank C, Schneider S, Eysholdt U et al (2011) Voice- and health-related quality of life in the elderly. J Voice 25:265–268. doi: 10.1016/j.jvoice.2009.11.002

Pontes P, Brasolotto A, Behlau M (2005) Glottic characteristics and voice complaint in the elderly. J Voice 19:84–94

Pontes P, Yamasaki R, Behlau M (2006) Morphological and functional aspects of the senile larynx. Folia Phoniatr Logop 58:151–158

Preacher KJ, Hayes AF (2008) Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behav Res Methods 40:879–891

Preacher KJ, Rucker DD, Hayes AF (2007) Addressing moderated mediation hypotheses: theory, methods, and prescriptions. Multivar Behav Res 42:185–227

Ptacek PHP, Sander EKE, Maloney WH, Jackson CCR (1966) Phonatory and related changes with advanced age. J Speech, Lang Hear Res 9:353–361

Rabinov CR, Kreiman J, Gerratt BR, Bielamowicz S (1995) Comparing reliability of perceptual ratings of roughness and acoustic measure of jitter. J Speech Hear Res 38:26–32

Ramig LO, Gray S, Baker K et al (2001) The aging voice: a review, treatment data and familial and genetic perspectives. Folia Phoniatr Logop 53:252–265

Revis J, Giovanni A, Wuyts F, Triglia JM (1999) Comparison of different voice samples for perceptual analysis. Folia Phoniatr Logop 51:108–116

Reynolds K, Pietrzak RH, El-Gabalawy R et al (2015) Prevalence of psychiatric disorders in U. S. older adults: findings from a nationally representative survey. World Psychiatry 14:74–81. doi:10.1002/wps.20193

Riketta M (2004) Does social desirability inflate the correlation between self-esteem and anxiety? Psychol Rep 94:1232–1234

Rouch I, Achour-Crawford E, Roche F et al (2014) Seven-year predictors of self-rated health and life satisfaction in the elderly : the proof study. J Nutr Health Aging 18:840–847

Roy N, Stemple J, Merrill RM, Thomas L (2007) Epidemiology of voice disorders in the elderly: preliminary findings. Laryngoscope 117:628–633. doi:10.1097/MLG.0b013e3180306da1

Sapienza CM, Stathopoulos ET (1995) Speech task effects on acoustic and aerodynamic measures of women with vocal nodules. J Voice 9:413–418

Sato K, Umeno H, Nakashima T (2010) Functional histology of the macula flava in the human vocal fold—part 2: its role in the growth and development of the vocal fold. Folia Phoniatr Logop 62:263–270. doi:10.1159/000316962

Sato K, Umeno H, Ono T, Nakashima T (2011) Histopathologic study of human vocal fold mucosa unphonated over a decade. Acta Otolaryngol 131:1319–1325. doi:10.3109/00016489.2011.615067

Sauder C, Roy N, Tanner K et al (2010) Vocal function exercises for presbylaryngis: a multidimensional assessment of treatment outcomes. Ann Otol Rhinol Laryngol 119:460–467

Shrout PE, Bolger N (2002) Mediation in experimental and non-experimental studies: new procedures and recommendations. Psychol Methods 7:422–445

Solhaug HI, Romuld EB, Romild U, Stordal E (2012) Increased prevalence of depression in cohorts of the elderly: an 11-year follow-up in the general population – the HUNT study. Int Psychogeriatr 24:151–158

Solomon NP, Garlitz SJ, Milbrath RL (2000) Respiratory and laryngeal contributions to maximum phonation duration. J Voice 14:331–340

Sowislo JF, Orth U (2012) Does low self-esteem predict depression and anxiety? A meta-analysis of longitudinal studies. Psychol Bull 139:213–240

Sperry EE, Klich RJ (1992) Speech breathing in senescent and younger women during oral reading. J Speech Hear Res 35: 1246–1255

Stathopoulos ET, Sapienza CM (1997) Developmental changes in laryngeal and respiratory function with variations in sound pressure level. J Speech Lang Hear Res 40:595–614

Stathopoulos ET, Huber JE, Sussman JE (2011) Changes in acoustic characteristics of the voice across the life span: measures from individuals 4–93 years of age. J Speech Lang Hear Res 54:1011–1021

Steffens DC, Fisher GG, Langa KM et al (2009) Prevalence of depression among older Americans: the aging, demographics and memory study. Int Psychogeriatr 21:879–888

Teles-Magalhães LC, Pegoraro-Krook MI, Pegoraro R (2000) Study of the elderly females' voice by phonetography. J Voice 14:310–321

Torre P, Barlow JA (2009) Age-related changes in acoustic characteristics of adult speech. J Commun Disord 42:324–333

Uhlmann RF, Larson EB (1991) Effect of education on the mini-mental state examination as a screening test for dementia. J Am Geriatr Soc 39:876–880

Verdonck-de Leeuw IM, Mahieu HF (2004) Vocal aging and the impact on daily life: a longitudinal study. J Voice 18:193–202

Wassertheil-Smoller S, Arredondo EM, Cai J et al (2014) Depression, anxiety, antidepressant use, and cardiovascular disease among Hispanic men and women of different national backgrounds: results from the Hispanic Community Health Study/Study of Latinos. Ann Epidemiol 24:822–830. doi:10.1016/j.annepidem.2014.09.003

Watson D, Suls J, Haig J (2002) Global self-esteem in relation to structural models of personality and affectivity. J Pers Soc Psychol 83:185–197

Watts CR, Ronshaugen R, Saenz D (2015) The effect of age and vocal task on cepstral/spectral measures of vocal function in adult males. Clin Linguist Phon 00:1–9 doi: 10.3109/02699206.2015.1005673

Wegener I, Geiser F, Alfter S et al (2015) Changes of explicitly and implicitly measured self-esteem in the treatment of major depression: evidence for implicit self-esteem compensation. Compr Psychiatry 58:57–67. doi:10.1016/j.comppsych.2014.12.001

Whitehead RL, Whitehead BH (1985) Acoustic characteristics of vocal tension/harshness in the speech of the hearing impaired. J Commun Disord 18:351–361. doi:10.1016/0021-9924(85)90025-5

Wilcox KA, Horii Y (1980) Age and changes in vocal jitter. J Gerontol 35:194–198

Williams ED, Tillin T, Richards M, et al. (2015) Depressive symptoms are doubled in older British South Asian and Black Caribbean people compared with Europeans: associations with excess co-morbidity and socioeconomic disadvantage. Psychol Med 1–11. doi: 10.1017/S0033291714002967

Winkworth AL, Davis PJ, Ellis E, Adams RD (1994) Variability and consistency in speech breathing during reading: lung volumes, speech intensity, and linguistic factors. J Speech Hear Res 37:535–556

Wolfe V, Martin D (1997) Acoustic correlates of dysphonia: type and severity. J Commun Disord 30:403–416

Wu Z, Schimmele CM, Chappell NL (2012) Aging and late-life depression. J Aging Health 24:3–28

Ximenes Filho JA, Tsuji DH, do Nascimento PHS, Sennes LU (2003) Histologic changes in human vocal folds correlated with aging: a histomorphometric study. Ann Otol Rhinol Laryngol 112:894–898

Yanagihara N (1967) Significance of harmonics changes and noise components in hoarseness. J Speech, Lang Hear Res 10:531–541

Yiu E, Worrall L, Longland J, Mitchell C (2000) Analysing vocal quality of connected speech using Kay's computerized speech lab: a preliminary finding. Clin Linguist Phon 14: 295–305

Zigmond AS, Snaith RP (1983) The hospital anxiety and depression scale (HADS). Acta Psychiatr Scand 67:361–370. doi: 10.1016/S0016-5085(01)83173-5

Zraick RI, Marshall W, Smith-Olinde L, Montague JC (2004) The effect of task on determination of habitual loudness. J Voice 18:176–182

Zraick RI, Birdwell KY, Smith-Olinde L (2005) The effect of speaking sample duration on determination of habitual pitch. J Voice 19:197–201

Zraick RI, Kempster GB, Connor NP et al (2011) Establishing validity of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V). Am J Speech Lang Pathol Am Speech Lang Hear Assoc 20:14–22