Investigating the Effects of Amateur Singing on Speaking Voice Quality, Function, and Motor Control Across Adulthood

[#,†,‡,¶]Amélie Nadeau, [#,†,¶]Pascale Tremblay, and [#,‡,§,¶]Maude Desjardins, #†‡§¶*Quebec City, Canada*

SUMMARY: Objectives. Singing is an accessible activity that may promote vocal health, including vocal function, quality, and motor control. However, evidence remains limited and heterogeneous, as some studies have shown an association between singing practice and dysphonia-related complaints in certain populations. This study investigated whether amateur singing practice influences adult speaking voice outcomes related to function, quality, and motor control, and whether it may moderate age-related vocal changes.

Methods. Participants were 143 French-speaking adults (20–98 years; 72 singers, 71 non-singers). Amateur singers were defined as individuals singing in a choir for at least two years with a minimal weekly practice of 60 minutes. Vocal function and quality were assessed using the Dysphonia Severity Index, Acoustic Voice Quality Index, speaking habitual intensity, and perceptual judgment of overall severity. Voice Onset Time (VOT) and its variability during a diadochokinetic task were used as proxies for vocal motor control. Linear and mixed-effects regression models tested the main effects of group (singers/non-singers), age, and their interaction, while controlling for relevant variables.

Results. Singers exhibited higher harmonic-to-noise ratios and lower shimmer values than non-singers. Age predicted lower maximal fundamental frequency values, independent of singing status. A moderating effect of singing was found for overall severity ratings, showing a trend toward better perceptual vocal quality in younger and middle-aged singers, but not in older singers, resulting in a steeper association between age and overall severity among singers. Interactions also revealed differential associations between VOT and age in singers compared to non-singers for specific consonants, suggesting potential differences in laryngeal-articulatory coordination between groups.

Conclusion. Findings suggest that amateur singing is associated with specific acoustic and perceptual benefits, although some of these associations may not be found in older adults. Longitudinal studies are needed to clarify the potential benefits of amateur singing.

Key Words: Singing–Aging–Vocal function–Voice quality–Vocal motor control–Voice onset time.

Professional vocal training and practice have significant effects on measures of vocal function and quality. Previous studies have reported enhanced resonance indicators,[1,2] a broader phonation range, and greater vocal intensity in the singing voice of trained singers.[3,4] Benefits extend beyond singing, with trained singers demonstrating longer maximum phonation times (MPT), increased voice stability in sustained vowels, better (ie, higher) scores on the Dysphonia Severity Index (DSI),[5,6] as well as increased vocal clarity and resonance in speech tasks.[1,2] Nevertheless, singers are also known to exhibit a high prevalence of voice-related problems due to important vocal demands, including self-perceived dysphonia, though these results vary according to population (eg, voice students, voice teachers, classically trained singers), training background, and assessment method.[7] Beyond vocal quality, professional singing practice has also been shown to influence vocal motor control, including Voice Onset Time (VOT), a marker of fine articulatory-laryngeal motor control,[8] although evidence remains limited and mixed. For instance, researchers have reported shorter VOTs in a sung voiced consonant in 24 trained female singers compared to 25 untrained female singers.[9] Conversely, other work found longer VOTs in five male singers compared to five male non-singers for voiceless consonants during singing,[10] which the authors interpreted as reflecting more precise and controlled articulatory movements in singers. However, the same authors later reported that the effect of singing on VOT varied based on the nature of the task (sung versus spoken syllables),[11] and found no benefits for female singers.[12]

Recent research has also explored the impact of amateur singing on vocal function, quality, and motor control during speech. In the literature, amateur singing is usually described as participation in recreational singing activities, for example, in settings such as community choirs. Although research in this area is limited with somewhat heterogeneous results, emerging evidence suggests that amateur singing may have beneficial effects on the speaking voice. For example, Ravi et al[13] reported lower perturbation measures and noise-to-harmonics ratios (NHR) in

Accepted for publication September 9, 2026.

From the #School of Rehabilitation Sciences, Faculty of Medicine, Université Laval, Quebec City, Quebec, Canada; †CERVO Brain Research Center, Quebec City, Quebec, Canada; ‡Center for Interdisciplinary Research in Rehabilitation and Social Inclusion (CIRRIS), Quebec City, Quebec, Canada; §Research Center of the CHU de Québec-UL, Université Laval, Quebec City, Quebec, Canada; and the ¶Centre for Research on Brain, Language and Music (CRBLM), Quebec City, Quebec, Canada.

Address correspondence and reprint requests to Maude Desjardins, École des sciences de la réadaptation, Faculté de Médecine, Université Laval, Quebec City, Quebec, Canada. E-mail: maude.desjardins@fmed.ulaval.ca

Journal of Voice, Vol xx, No xx, pp. xxx–xxx
0892-1997

© 2026 The Author(s). Published by Elsevier Inc. on behalf of The Voice Foundation. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

https://doi.org/10.1016/j.jvoice.2026.09.019

untrained choir singers aged 20–30 years compared to non-singers. Similarly, Awan and Ensslen[14] found better DSI scores in choir singers aged 18–30 years relative to non-singers, along with higher maximum fundamental frequencies (f_0High) and reduced jitter measures. Regarding vocal or speech motor control, the literature on amateur singers is scarce, but a study from our group has shown that amateur singers demonstrate better articulatory accuracy in an oral diadochokinetic (DDK) task compared to non-singers.[15]

Overall, amateur singing may be associated with improvements in specific vocal outcomes, but further research is needed to clarify these effects and the conditions under which they apply. For example, it is unclear whether any singing-related benefits persist into older adulthood despite the physiological changes associated with aging, which impact vocal function and quality. In fact, aging is associated with structural and functional alterations in the larynx (eg, vocal fold atrophy), resonant cavities, and respiratory system (for a review, see Mallick et al[16]) that manifest acoustically in older adults' voices. For example, studies have reported greater voice perturbation demonstrated by increased jitter and shimmer, reduced harmonic-to-noise ratio (HNR)[17–19] and sound pressure level (SPL) (ie, voice intensity),[20] and declines in maximal function parameters such as f_0High and MPT.[5] Perceptually, aging voices have been shown to exhibit signs of breathiness, hoarseness, and instability.[21,22] Beyond vocal function and quality, aging also impacts speech and vocal motor control. Declines in articulation and speech rates have been reported in older speakers,[23–25] as well as increased variability and reduced articulatory accuracy in tasks such as syllable reading[26,27] and non-word repetition.[28] Global VOT duration appears generally stable across age,[29–31] although age-related effects have been observed for specific consonants like voiceless plosives.[32,33] Finally, VOT variability tends to increase with age, suggesting reduced stability in articulatory-laryngeal timing among older speakers.[30,31,33,34]

Emerging evidence suggests that amateur singing may have a moderating effect on age-related vocal decline. Older choir singers' voices have been described as clearer, with less roughness and strain, as assessed using the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V),[35] compared to age-matched non-singers.[36] Acoustically, older choir singers have been shown to exhibit lower jitter than aged non-singers, suggesting improved vocal stability,[37] as well as longer MPT[38] and f_0 stability during sustained vowels (eg, smaller f_0 SD).[39] A pilot study by Stager et al[40] reported that choral singing may contribute to the preservation of speaking voice quality, as indicated by reduced noise energy between 2 and 3 kHz in the voice of choir singers aged 65+ compared with age-matched controls. A subsequent longitudinal study extended these findings by examining the effects of a structured choral singing program in older participants. The authors reported improvements in aerodynamic (eg, subglottal pressure, flow) and acoustic measures (eg, long-term average spectrum [LTAS])—with evidence of maintained gains over time—although the parameters that improved differed according to prior singing experience.[41] Furthermore, a recent systematic review underscored the potential of singing-based approaches to improve respiratory function and respiratory-phonatory control (ie, airflow control measured through MPT) in aging adults with a respiratory disease or cognitive impairment diagnosis, although the effect on phonatory function remains unclear.[42] Specifically, the review highlighted several limitations in the literature, reducing the ability to form a comprehensive understanding of how singing impacts voice production, including small sample sizes and substantial variability in intervention protocols and results. Importantly, further research is needed to clarify the specific effects of amateur singing on the aging speaking voice in healthy adults.

The general objective of the present study was to determine whether regular amateur singing practice influences speaking voice outcomes in adults spanning a wide age range, without formal vocal training. The first specific objective was to assess how amateur singing practice influences: A) objective measures of vocal function and quality, operationalized through two multiparametric measures—the Dysphonia Severity Index (DSI) and the Acoustic Voice Quality Index (AVQI)—and their individual components, as well as through a measure of habitual intensity; B) subjective assessment of vocal quality, using perceptual ratings based on the overall severity dimension of the CAPE-V; and C) vocal motor control, operationalized through VOT and its variability in voiceless plosives. The second specific objective was to assess whether amateur singing moderates age-related changes across the same voice-related measures.

Based on previous evidence suggesting beneficial effects of amateur singing practice on the speaking voice, our first hypothesis was that adults who regularly engage in amateur singing would demonstrate enhanced vocal function, vocal quality, and articulatory-laryngeal coordination (ie, longer VOTs in voiceless plosives) compared to non-singers. Our second hypothesis was that amateur singers would show reduced age-related declines on these measures, consistent with the proposition that amateur singing practice may provide beneficial effects on the speaking voice across a wide age range.

METHOD

Participants

Recruitment

A total of 151 participants (mean age: 53 years, SD: 20 years, range: 20–98 years, 85 females/66 males) were initially recruited in the context of a broader project through emails, posters, and social media posts aimed at the general community and Université Laval's community, as well as targeted emails and social media posts directed at choirs in

the Quebec City area. The study was approved by the *Comité d'éthique de la recherche sectoriel en neurosciences et santé mentale, Institut Universitaire en Santé Mentale de Québec* (#192–2017; The Quebec City Singing dataset).[43–46] Informed consent was obtained for each participant.

Eligibility criteria

To be included, participants had to be right-handed according to the Edinburgh Handedness Inventory[47] (due to the magnetic resonance imaging (MRI) portion of the broader project, not reported here), non-smokers, and native speakers of Canadian French. Participants were excluded if they reported a history of language, psychiatric, neurological, or neurodegenerative disorders, speech, voice, or swallowing disorders, or uncontrolled gastroesophageal reflux. They also had to report normal or corrected-to-normal vision. People with professional vocal training as well as regular instrument players (> 25 hours/week) were excluded to minimize potential confounding effects of vocal training and instrumental practice. For the purposes of this specific study, wind instrument practice was also considered an exclusion criterion due to its potential influence on voice-related parameters such as MPT. Three participants were excluded on this basis. Participants were also excluded if they showed evidence of impaired vocal quality based on an auditory-perceptual evaluation using the overall severity dimension of the CAPE-V scale[35] (see Data Extraction and Voice Measurements subsections). This screening was implemented to minimize the impact of vocal impairment that may reflect underlying pathological processes on the relationships of interest. Five participants, aged 21, 22, 30, 86, and 87 years, were excluded following this evaluation.

Participants were divided into two groups according to their singing background. The present study adopts a definition of amateur singing based on the absence of formal professional vocal training and the recreational nature of the activity (eg, choral singing). The control group consisted of individuals not involved in any form of singing—solo or in group, professional or amateur. The singing group consisted of individuals with at least two years of regular amateur choral singing experience, with a frequency of at least once a week for a minimum of 60 consecutive minutes.

Questionnaires and covariates

All participants completed five questionnaires: *(1)* A sociodemographic questionnaire with questions about current/former occupation, perceived physical condition (rated on a scale from 1 (poor) to 7 (excellent)), and items on social participation (questionnaire accessible on Borealis, see Data Availability section); *(2)* a questionnaire assessing past and present musical and singing experiences; *(3)* the long version of the Geriatric Depression Scale (GDS)[48]; *(4)* the Montreal Cognitive Assessment (MoCA)[49] and *(5)* the Quebec French version of the Voice Handicap Index (VHI-QF)[50,51] assessing the functional impact of one's voice on daily life. Refer to Appendix A for more detailed information on the questionnaires.

Two scores were derived from combining items from the questionnaires mentioned above and were used as covariates. *(1)* Social Participation Score: Given the associations between mood disorders, social isolation, and vocal behavior and function,[52–54] a composite score was constructed from items in the sociodemographic questionnaire to quantify social participation, inspired by the social isolation score developed by Saito et al.[55] The developed score ranges from 0 to 17, with higher scores indicating greater social engagement (see Appendix A for details); *(2)* Professional Vocal Use Score: Derived from occupational data obtained from the sociodemographic questionnaire, participants were classified using a binary system as either professional voice users (eg, teachers, tour guides, call center agents) or nonprofessional voice users. Retired participants with a history of professional voice use were classified as professional voice users.

Procedures

The experiment took place in a double-walled sound-attenuated room at the Speech and Hearing Neuroscience Laboratory in Quebec City, Canada. The visit lasted three hours, during which participants underwent an audiometric evaluation, completed a series of voice tasks (described below), as well as a series of speech, language, and cognitive tasks—the results of which have been published elsewhere.[43,45,56] Additionally, a subsample of 85 participants underwent an MRI session (not used in the present study), the results of which have also been published elsewhere.[44,46,57] Assessments relevant to the present study are detailed next.

Audiometric evaluation

Pure-tone thresholds in Decibels Hearing Level (dB HL) were measured with a calibrated clinical audiometer (AC40, Interacoustic, Denmark). Thresholds were obtained separately for each ear at three frequencies: 0.5, 1, and 2 kHz. For each frequency, the lower threshold between the two ears was retained. These three best thresholds were then averaged to obtain an overall best pure-tone average (PTA) score.

Speech and voice production tasks

All speech and voice recordings were performed in the sound-attenuated room under identical conditions for all participants throughout the entire data collection. Recordings were made using a head-worn condenser Shure Microflex Beta 53 microphone connected to a Shure Microflex RK100PK preamplifier and a Quartet USB audio interface (Apogee Electronics, Santa Monica, CA, USA), which was connected to an iMac computer (27-inch, 3.5 GHz Intel Core i7, macOS Mojave 10.14.3). The

microphone was equipped with the RPM304 windscreen. The equipment and recording chain were the same throughout data collection. The head-worn configuration ensured a stable placement with an approximate mouth-to-microphone distance of 5 cm and a consistent angle across all recordings, which was visually inspected before each session. Recordings were acquired via the *Sound Studio 4* software (Felt Tip Inc., New York, NY, USA) with a sampling rate of 48 kHz and 24-bit quantization.

Standardized reading passage. Participants were given a written version of the standardized passage "La bise et le soleil".[58] While seated, they were first asked to read it silently, then to read it aloud as naturally as possible in a comfortable speaking voice.

Neutral discourse. While seated, participants were asked to reflect on a typical day of their week and to speak about it continuously for two minutes. They were instructed to do so in a natural manner, as if having an informal conversation with a close friend or family member.

Vowel Production. Participants completed four standard voice tasks, each repeated three times with a short pause between each production, while standing. **(1)** Sustained vowel. Participants produced the /ɑ/ vowel at a comfortable pitch and loudness for 5 seconds. **(2)** Maximum phonation time. Participants produced the vowel /ɑ/ for as long as they could at a comfortable pitch and loudness. **(3)** Vocal crescendo. Participants produced the /ɑ/ vowel at the lowest possible loudness without whispering and then increased the loudness to its maximum, in a single breath without interruption and without yelling. **(4)** Vocal glissando. Participants were instructed to produce the /ɑ/ vowel at a natural pitch, to progressively lower it to the lowest pitch they could reach, and then to raise it to the highest pitch they could produce, all in one breath, without interruption and without yelling. For the sustained vowel and MPT tasks, a practice trial was provided following the instructions. For both the crescendo and glissando tasks, a member of the research team demonstrated the task after the instructions. Participants then completed a practice trial, and once performance was judged adequate, three trials were recorded.

DDK. After a short break, participants were seated comfortably in front of a screen and completed a DDK task involving repeated syllables (/pa/, /ta/, /ka/) and syllable sequences (/pata/, /taka/, /pataka/), presented visually on the screen. They were instructed to repeat each item as quickly and accurately as possible. A brief practice period (~10 seconds) was provided, with additional practice allowed if the rhythm was too slow or the task was misunderstood. Items were presented in a fixed order to create the following pattern: /pa/, /ta/, /ka/, /pata/, /taka/, /pataka/. Each item was displayed for 5 seconds. Participants repeated this progression of items three times, resulting in three blocks.

Data extraction and voice measurements

All audio files were cleaned and segmented using Praat (v.6.4.27).[59] Recordings from the sustained vowels, crescendo, and standardized passage tasks were amplitude-calibrated using a white-noise reference played by a portable recorder (YAMAHA W24) placed at the same distance from the microphone as the participant's mouth. A sound level meter (Galaxy Audio Check Mate CM 160), placed at the same distance from the white-noise source as the microphone, measured the absolute sound pressure level of the white noise. A custom normalization Praat script adjusted the intensity (in dB) of each participant's audio file by applying a correction factor derived from the intensity difference between the microphone and the sound level meter. The standardized passage was then segmented to retain only the first sentence, as this segment demonstrated good discriminative capacity for French speakers in AVQI-based acoustic analyses.[60]

Objective measures of vocal function and quality. Multiparametric measures such as the DSI[61] and the AVQI[62] offer a standardized and reproducible assessment of the voice while providing an evaluation of multiple key parameters, as they combine several voice parameters into a single score. The DSI, developed as a composite score of dysphonia severity, is derived from four parameters that capture performance on maximal function tasks (MPT, f_0High, lowest intensity (ILow)) and period perturbation in a sustained vowel (jitter %). The DSI yields a single continuous score ranging from −5 (severely impaired vocal function) to +5 (normal vocal function),[61,63] but scores may exceed these bounds in cases of exceptionally good or poor voice performance.[61] A custom Praat script using the original DSI authors' settings was created to automate the extraction of the DSI components from each participant's vowel production tasks.[64,65] For MPT, f_0High and ILow, the three trials were analyzed, and best performance was retained. Jitter was computed from the third trial of the sustained /ɑ/. The script then applied the DSI formula (DSI script v.02.03)[64,65] to compute a score for each participant.

The AVQI assesses vocal quality by combining six acoustic markers: smoothed cepstral peak prominence (CPPs), HNR, shimmer %, shimmer dB, and the slope and tilt of the LTAS, extracted from a sustained vowel and a segment of continuous speech, with scores ranging from 0 (better vocal quality) to 10 (worse vocal quality).[62] A custom Praat script using the original AVQI authors' default parameters was used to automate the extraction of the six markers from the sustained vowel and reading tasks, and to compute AVQI scores.[62,66] Prior to extracting the measures, a high-pass filter at 34 Hz was applied to the voice samples to eliminate low-frequency noise,[66–68] and continuous speech files were segmented to isolate and then merge voiced portions. From this composite signal, six acoustic markers were extracted and the AVQI formula (v.02.06)[66] was applied.

Vocal function was further assessed by extracting measures of habitual intensity (in SPL) derived from the neutral

discourse task. Habitual intensity was calculated using a custom Praat script that computed the mean intensity for each pre-segmented sentence of the participant's discourse and then averaged these values to yield a single habitual intensity measure per participant.

Subjective evaluation of vocal quality. Two master's students in speech-language pathology conducted a perceptual assessment of the whole standardized reading passage using the overall severity scale of the CAPE-V, under the supervision of the last author. The evaluators had prior clinical experience in voice disorders through their clinical internships, and they each completed five hours of individual perceptual training using the All-Voiced[69] platform. This training tool is designed for speech-language pathologists and students. It provides voice samples and related demographic information, as well as a virtual rating interface modeled after the CAPE-V. Both evaluators also completed an auditory anchor selection procedure, as this approach has been shown to stabilize evaluators' internal standards and improve rater reliability in perceptual voice assessment.[70,71] Each evaluator independently reviewed half of the voice sample database and identified potential anchors representing different levels of severity for the following parameters: overall severity, roughness, breathiness, and strain. The final selection was made by the last author. The evaluators listened to the anchor samples at the start of each rating session and again after every 20 samples to maintain a consistent frame of reference. Because the overall severity dimension of the CAPE-V has demonstrated stronger reliability than other dimensions,[72,73] this parameter was selected for analysis of subjective vocal quality. Overall severity reflects a global judgment of voice impairment, rated on a visual analog scale from 0 mm (normal) to 100 mm (severely deviant)[35] (scores were expressed in cm in the present study). To assess interrater reliability, the raters evaluated the same subset of 23 participants (15% of the total sample). To assess intrarater reliability, each rater independently evaluated another subset of 23 participants twice. The remaining samples were randomly distributed between the two raters to balance the age and group distributions across judges. Judge 1 rated participants aged 22–98 years (10 younger [< 40 years], 12 middle-aged [40–59 years], 19 older [≥60 years]; 14 non-singers, 27 singers), and Judge 2 rated participants aged 22–87 years (9 younger, 10 middle-aged, 22 older; 15 non-singers, 26 singers).

Vocal motor control. VOT and its coefficient of variation (VOT-CoV; SD/mean)—calculated from the DDK task (/pataka/ syllable group)—were used as proxies for vocal motor control, providing a standardized and reproducible measure of articulatory-laryngeal coordination. VOT segmentation was performed in Praat by the first author and a research assistant following previously established guidelines,[74] and after receiving the same formal training in VOT segmentation. VOT is defined as the time interval between the release of a stop consonant (eg, /p/) and the onset of vocal fold vibration in a consonant-vowel syllable (eg, /pa/).[75] Initially, the first three repetitions of the /pataka/ sequence in each of the three blocks of the task were extracted, and then each /pataka/ syllable group was segmented into individual syllables (/pa/, /ta/, or /ka/), providing nine VOT measurements per syllable per participant (total of 27 measurements/participant). For each syllable, the beginning of the burst associated with the voiceless consonant (/p/, /t/, or /k/) was identified, along with the onset of the first consistent voiced cycle, with the duration in between corresponding to the VOT measure.[75] If a clear burst was not visible, the spectrogram was used to help identify its location. If the spectrogram could not be used to reliably identify the burst, or if the participant's production did not match the target syllable (eg, incorrect articulation or a substitution), the syllable was deemed unusable. In such cases, the next available /pataka/ sequence within the same block was selected. Additionally, because speech rate is known to influence VOT,[76] it was computed for each participant within each block of the DDK task by dividing the duration of the sequence by the total number of syllables produced.

To ensure consistency, the raters developed a detailed segmentation flowchart adapted from these guidelines, tailored to the specific characteristics of our dataset (accessible on Borealis, see Data Availability section). To do so, they first familiarized themselves with the procedure using recordings from the same 10 participants, after which they conferred to discuss discrepancies and reach consensus. To verify reliability of the process, the two raters then evaluated the same subset of 22 DDK recordings (15% of the total sample) independently. Intrarater reliability was assessed using the first author's blinded reevaluation of a separate subset of 22 participants, as only the VOT segmentations by the first author were used for all analyses.

Statistical analysis

All statistical analyses were conducted in R (v.2025.5.0.496)[77] using custom R scripts. For reliability assessment, intraclass correlation coefficients (ICCs)[78] were calculated using the irr package (v.0.84.1)[79] and interpreted according to the guidelines of Koo and Li.[78] Descriptive statistics were computed for each variable and for each group (singers, non-singers), and visually assessed using scatter or bar plots. Data cleaning involved identifying and removing outliers beyond ±3 SD for each variable separately (both groups combined). Distributional properties were assessed through skewness and kurtosis values, as well as visual inspection of density curves to evaluate normality. Group comparisons were conducted using Welch 2-Sample *t* tests for continuous variables and chi-square tests for categorical variables.

Three series of analyses were conducted, each focusing on a set of different parameters: *(1)* objective measures of vocal function and quality, *(2)* subjective evaluation of vocal quality, and *(3)* vocal motor control. In all analyses, linear regression models (for vocal function and quality data) and linear mixed-effects models (for vocal motor control data) were constructed using the *buildmer* (v.2.11),[80] *lme4* (v.1.1.37)[81] and jtools packages (v.2.3.0).[82] The buildmer package starts with an initial (full) model and employs a backward selection procedure,

retaining only predictors that significantly improve model fit—those that explain the most variance[83]—based on changes in log-likelihood and Akaike Information Criterion (AIC). Consequently, predictors may appear in the results tables without being statistically significant if they contribute to overall fit. Visualizations were generated using the sjPlot (v.2.8.16),[84] ggplot2 (v.4.0.1)[85] and interaction (v.1.2.0)[86] packages. *Post hoc* comparisons and significant interactions were explored using simple slope analysis and estimation of marginal trends (emtrends). Model diagnostics were conducted to ensure that regression assumptions were met, including assessment of multicollinearity and residual normality using Q-Q plots.

Objective measures of vocal function and quality

To assess how amateur singing practice influences objective measures of vocal function and quality (objective 1A), and whether singing moderates age-related changes across these measures (objective 2), dependent variables included the overall DSI score and its four parameters, the global AVQI score and its six markers, as well as habitual intensity, analyzed in separate linear regression models. Independent factors were age, group (singers vs. non-singers) and their interaction, allowing for assessment of both the main effect of singing practice and its potentially moderating effect on age-related influences. All models were adjusted for six covariates: sex, self-reported physical condition, overall PTA, GDS, professional vocal use, and social participation score.

Subjective evaluation of vocal quality

Interrater reliability for CAPE-V overall severity ratings was assessed with a mean-rating (k = 2), absolute agreement, two-way mixed-effects model (ICC [3,2]).[78] Intrarater reliability was computed using a single-rater, absolute agreement, two-way mixed-effects model (ICC [3,1]).[78]

To assess the influence of amateur singing practice on subjective vocal quality (objective 1B), and whether singing moderates age-related changes (objective 2), linear regression models identical to those described above were used, with overall severity assessments on the CAPE-V scale as the dependent variable. Diagnostics revealed non-normal residuals for the CAPE-V overall severity distribution. Therefore, a square-root transformation was applied, and a new regression model was fitted using the transformed data.

Vocal motor control

Interrater reliability for VOT segmentations was assessed using a single-rater, consistency, two-way mixed-effects model (ICC [3,1]).[78] This model was chosen because our goal was to evaluate whether raters applied the segmentation criteria consistently across participants, but only the values segmented by the first author were intended for use in subsequent analyses. ICCs were computed based on VOT variability (VOT-CoV), to focus on consistency rather than absolute agreement. This method was chosen because frequent coarticulation in older participants often blurred VOT boundaries, making inclusion or exclusion of coarticulated segments justifiable as long as raters were consistent. VOT-CoV was calculated separately for each syllable (/pa/, /ta/, /ka/) on all repetitions across the three blocks combined (9 measurements/participant). An additional VOT-CoV value was computed across all syllables (27 measurements/participant). Intrarater reliability was computed using a single-rater, absolute agreement, two-way mixed-effects model (ICC [3,1])[78] following the same procedure as for the interrater reliability (ie, syllable-specific and global VOT-CoV).

To evaluate how amateur singing practice influences vocal motor control (objective 1C), and to assess whether it moderates age-related differences (objective 2), the same approach as described above was applied using linear mixed-effects models with VOT and VOT-CoV as dependent variables. Block and syllable (/pa/, /ta/, /ka/) were also included in the models as independent variables, as well as speech rate as an additional covariate.[76] This design allowed us to examine how vocal motor control varies according to specific articulatory demands (ie, syllables), and how it evolves across repetitions (ie, block). For the VOT-CoV model, residuals showed non-normality; a square-root transformation was therefore applied before refitting the models.

Multiple comparison correction

To control for the risk of type I error, a Benjamini-Hochberg False Discovery Rate (FDR) correction was applied across conceptually grouped families: *(1)* DSI global score and its parameters; *(2)* AVQI global score and its parameters; *(3)* habitual intensity; *(4)* perceptual judgment on the overall severity dimension of the CAPE-V; *(5)* VOT measures; and *(6)* VOT-CoV measurements. Raw P-values were extracted from the final regression models and adjusted using the "stats" package (v.4.4.2).[87] The full analytical script is available on Borealis (see Data Availability section).

RESULTS

For the sake of concision, only hypothesis-relevant predictors, including their main effects and interactions, are described in the text. Full results are detailed in Tables 1–3. Detailed tables related to reliability assessments are presented in Appendix B.

Final sample

The final sample (N = 143) was composed of 71 non-singers (mean age: 53 ± 19 years; range: 20–98 years, 36 females) and 72 singers (mean age: 54 ± 19 years; range: 21–87 years; 45 females). Groups did not differ in terms of biological sex (male to female ratio) (P = 0.210), age, pure-tone hearing, education level, scores on the VHI, MoCA, GDS, perceived physical condition, professional vocal use,

TABLE 1.
Participants Characteristics Per Group

Characteristics (continuous variables)	Non-Singers (N = 71, F = 36)				Singers (N = 72, F = 45)				Group Difference
	M	*SD*	*Min*	*Max*	*M*	*SD*	*Min*	*Max*	*P*[*]
Age	53.24	19.63	20.00	98.00	54.82	18.76	21.00	87.00	0.623
Education in years[a]	16.69	3.05	9.00	25.00	16.06	3.03	7.00	25.00	0.218
Perceived Physical Condition (/7)	5.29	0.90	3.00	7.00	5.08	1.01	3.00	7.00	0.189
Overall PTA[b]	8.92	8.83	-5.00	33.33	8.11	8.16	-3.33	30.00	0.576
MoCA[c] (/30)	27.31	2.18	21.00	30.00	27.62	1.86	23.00	31.00	0.361
GDS[d] (/30)	2.57	2.48	0	8.00	2.21	2.53	0	10.00	0.405
Social Participation Score (/17)	5.37	5.27	0.25	12.00	5.14	3.22	0	14.50	0.674
VHI-QF[e] (/120)	6.81	6.31	0	33.00	8.00	5.84	0	26.00	0.246
Characteristics (categorical variables)	**N**		**%**		**N**		**%**		***P***[†]
Male	35		49.3		27		37.5		0.210
Female	36		50.7		45		62.5		
Professional Voice Users	16		22.5		12		16.7		0.501
Nonprofessional Voice Users	55		77.5		60		83.3		

Abbreviations: F, number of female participants; M, mean; N, number of participants per group; SD, standard deviation of the mean.
[a] Education = Number of years of education based on the highest degree obtained in Quebec.
[b] Overall PTA = average of the best pure tone average thresholds between 0.5, 1 and 2 kHz, in decibels (dB HL).
[c] MoCA = Montreal Cognitive Assessment, a screening tool scored out of 30, where higher scores reflect better cognitive performance.
[d] GDS = Geriatric Depression Scale, a 30-item yes/no questionnaire where higher scores suggest more severe depression.
[e] VHI-QF = Quebec French Voice Handicap Index, a self-reported 30-item where higher scores indicate a greater perceived voice-related handicap.
[*] Result of the Welch Two Sample *t* test in R.
[†] Result of the Chi-squared test in R.

and social participation (all $P > 0.189$) (Table 1). Singers had an average of 16.2 years of singing experience, ranging from 2 to 62 years, and the majority (63/72) began singing in adulthood. Beyond their weekly choir participation, 30/72 singers practiced singing at home daily, while 28/72 engaged in home practice once a week. The remaining singers reported practicing at home either once a month or less. A detailed table of each participant's characteristics is available on Borealis (see Data Availability section).

Objective measures of vocal function and quality

Results of linear regression models for objective measures of vocal function and quality are shown in Table 2. We analyzed the effects of age, group, and their interaction on vocal function and quality using the overall DSI and AVQI scores and their individual parameters, as well as the habitual intensity measure during spontaneous speech (objectives 1A and 2). Between one and 11 participants were excluded from the final regression models for these measures due to missing data or outlier values.

DSI-related effects

No significant effects were found for the main predictors in the overall DSI score model, as well as in the MPT, jitter, and ILow models. A significant main effect of age was observed for f_0High, indicating a lower maximal f_0 with increasing age (b = −2.77, SE = 1.12, $P = 0.048$). Moreover, significant effects of sex were found for DSI, MPT, jitter, and f_0High (see Appendix C for figures).

AVQI-related effects

No significant effects were found for the main predictors in the overall AVQI score model, as well as in the CPPs and LTAS slope models. Group was a significant predictor for HNR and both shimmer measures, with singers exhibiting higher HNR values than non-singers (b = 1.12, SE = 0.42, $P = 0.030$) (Figure 1A, also showing sex effects), lower shimmer % values (b = −0.62, SE = 0.26, $P = 0.040$) (Figure 1B, also showing sex effects), and lower shimmer dB values (b = −0.04, SE = 0.02, $P = 0.030$). A significant main effect of group was initially found for LTAS tilt but failed to survive FDR correction. See Table 2 for details.

Intensity-related effects

No significant effect was found for the main predictors in the habitual intensity models.

Exploratory analyses

Additional exploratory analyses were conducted within the amateur singer group to examine whether variability in singing experience could have concealed potential effects. The normalized singing experience predictor (years of singing/age) indexes the proportion of an individual's life spent singing. This normalized predictor was used because age and years of singing were moderately correlated (r = 0.41), reflecting the tendency for older singers to have accumulated more years of singing experience. The same analytical procedure was applied as in the primary analyses (outlier management, verification of statistical

TABLE 2.
Results Of Linear Regression Models For Vocal Function And Vocal Quality Parameters

Dependent Variable	*Predictors*	*b*	*SE*	*CI*	*P*	*P (FDR)*
DSI	(Intercept)	3.31	0.32	2.67-3.95	**< 0.001**	
Model info: Observations = 133	GDS	−0.01	0.06	−0.13-0.12	0.936	0.936
R^2 / R^2 adjusted = 0.063 / 0.041	PTA	−0.03	0.02	−0.07-0.00	0.084	0.180
AIC = 537.106	Sex [Female]	0.77	0.31	0.15-1.39	**0.015**	**0.048**
MPT	(Intercept)	15.91	3.09	9.79-22.03	**< 0.001**	
Model info: Observations = 131	Sex [Female]	−4.26	1.04	−6.32-−2.21	**< 0.001**	**0.008**
R^2 /R^2 adjusted = 0.159 / 0.125	PTA	−0.04	0.06	−0.16-0.08	0.529	0.610
AIC = 843.471	GDS	−0.25	0.21	−0.65-0.16	0.236	0.393
	Perceived Physical Condition	1.07	0.55	−0.02-2.17	0.054	0.135
	Social Participation	−0.16	0.17	−0.49-0.17	0.345	0.518
Jitter	(Intercept)	0.25	0.01	0.22-0.28	**< 0.001**	
Model info: Observations = 139	Sex [Female]	−0.05	0.02	−0.08-−0.01	**0.016**	**0.048**
R^2 / R^2 adjusted = 0.042 / 0.035						
AIC = −216.208						
ILow	(Intercept)	74.28	3.94	66.49-82.08	**< 0.001**	
Model info: Observations = 132	GDS	−0.05	0.27	−0.59-0.48	0.841	0.841
R^2 / R^2 adjusted = 0.014 / −0.017	PTA	0.05	0.08	−0.11-0.21	0.520	0.732
AIC = 920.034	Social Participation	0.13	0.22	−0.30-0.56	0.549	0.732
	Perceived Physical Condition	−0.83	0.72	−2.26-0.60	0.252	0.732
***f_0*High**	(Intercept)	568.68	107.64	355.65-781.71	**< 0.001**	
Model info: Observations = 132	Sex [Female]	228.12	34.15	160.54-295.71	**< 0.001**	**0.008**
R^2 / R^2 adjusted = 0.343 / 0.311	GDS	9.65	6.69	−3.59-22.88	0.152	0.285
AIC = 1769.941	PTA	−0.87	2.46	−5.72-3.99	0.725	0.777
	Social Participation	3.68	5.45	−7.11-14.47	0.501	0.610
	Perceived Physical Condition	11.97	18.23	−24.11-48.04	0.513	0.610
	Age	−2.77	1.12	−4.99-−0.56	**0.015**	**0.048**
AVQI	(Intercept)	2.38	0.07	2.24-2.51	**< 0.001**	
Model info: Observations = 141						
R^2 / R^2 adjusted = 0.000 / 0.000						
AIC = 341.192						
CPPs	(Intercept)	15.58	0.45	14.69-16.48	**< 0.001**	
Model info: Observations = 134	PTA	0.01	0.02	−0.03-0.05	0.653	0.762
R^2 / R^2 adjusted = 0.054 / 0.025	Sex [Female]	−0.85	0.32	−1.48-−0.23	**0.008**	**0.030**
AIC = 547.469	GDS	−0.02	0.06	−0.15-0.11	0.759	0.839
	Social Participation	−0.01	0.05	−0.11-0.09	0.904	0.916
HNR	(Intercept)	14.74	1.28	12.20-17.28	**< 0.001**	
Model info: Observations = 134	PTA	0.02	0.02	−0.03-0.07	0.440	0.544
R^2 / R^2 adjusted = 0.142 / 0.109	Group [Singer]	1.12	0.42	0.29-1.95	**0.009**	**0.030**
AIC = 621.372	Perceived Physical Condition	0.58	0.22	0.14-1.01	**0.010**	**0.030**
	GDS	−0.01	0.08	−0.17-0.16	0.916	0.916
	Sex [Female]	1.07	0.42	0.24-1.90	**0.012**	**0.032**
Shimmer (%)	(Intercept)	6.14	0.79	4.58-7.70	**< 0.001**	
Observations = 138	Sex [Female]	−1.07	0.26	−1.58-−0.56	**< 0.001**	**0.007**
R^2 / R^2 adjusted = 0.207 / 0.177	PTA	−0.03	0.02	−0.06-0.01	0.170	0.223
AIC = 504,958	Group [Singer]	−0.62	0.26	−1.13-−0.11	**0.017**	**0.040**
	Age	0.02	0.01	0.00-0.04	**0.031**	0.054
	Perceived Physical Condition	−0.27	0.13	−0.53-−0.00	**0.048**	0.072
Shimmer (dB)	(Intercept)	0.51	0.01	0.48-0.54	**< 0.001**	
Model info: Observations = 142	Sex [Female]	−0.09	0.02	−0.12-−0.05	**< 0.001**	**0.007**
R^2 / R^2 adjusted = 0.231 / 0.220	Group [Singer]	−0.04	0.02	−0.07-−0.01	**0.006**	**0.030**
AIC = −266.699						
LTAS Slope	(Intercept)	−18.94	0.47	−19.87-−18.01	**< 0.001**	
Model info: Observations = 135	PTA	0.00	0.03	−0.06-0.06	0.984	0.984
R^2 / R^2 adjusted = 0.003 / −0.012	GDS	−0.07	0.11	−0.28-0.14	0.524	0.669
AIC = 690.372						
LTAS Tilt	(Intercept)	−11.23	0.29	−11.79-−10.66	**< 0.001**	
	Sex [Female]	0.64	0.14	0.37-0.91	**< 0.001**	**0.007**
	Voice Use	−0.38	0.17	−0.72-−0.04	**0.029**	0.054

TABLE 2 (*Continued*)

Dependent Variable	*Predictors*	*b*	*SE*	*CI*	*P*	*P (FDR)*
Model info: Observations = 142	Age	0.01	0.00	−0.00-0.02	0.121	0.169
R^2 / R^2 adjusted = 0.186 / 0.156	Group [Singer]	0.85	0.41	0.04-1.66	**0.039**	0.063
AIC = 346.765	Age × Group [Singer]	−0.02	0.01	−0.03-−0.00	**0.028**	0.054
Habitual Intensity (SPL)	(Intercept)	68.88	1.66	65.60-72.16	**< 0.001**	
Model info: Observations = 131	GDS	−0.04	0.25	−0.54-0.46	0.873	0.873
R^2 / R^2 adjusted = 0.017 / −0.006	PTA	0.08	0.08	−0.07-0.23	0.267	0.337
AIC = 896.832	Social Participation	0.23	0.20	−0.17-0.63	0.262	0.337
Overall Severity	(Intercept)	0.17	0.04	0.09-0.24	**< 0.001**	
Model info: Observations = 140	Age	0.0016	$7x10^{-4}$	0.0003-0.0029	**0.016**	**0.019**
R^2 / R^2 adjusted = 0.231 / 0.214	Group [Singer]	−0.12	0.05	−0.23-−0.02	**0.019**	**0.019**
AIC = −229.438	Age × Group [Singer]	0.0023	0.0009	0.0004-0.0041	**0.016**	**0.019**

Notes: Significant P-values of predictors are bolded.
Abbreviations: AIC, Akaike Information Criterion; AVQI, Acoustic Voice Quality Index; b, estimate; CI, confidence interval of the estimate; CPPs, Cepstral peak prominence smoothed; DSI, Dysphonia Severity Index; GDS, Geriatric Depression Scale; HNR, Harmonic-to-noise ratio; ILow, Lowest intensity; f_0High, Highest fundamental frequency; ICC, Intraclass Correlation Coefficient; MPT, Maximum phonation time; N, Number of participants; P (FDR), P value after applying FDR correction; PTA, average of the best pure tone average thresholds; SE, standard error of the estimate.

TABLE 3.
Results Of The Linear Mixed-effects Regression Models For Vocal Motor Control Parameters (VOT, VOT-CoV)

Dependent Variable	*Predictors*	*b*	*SE*	*CI*	*P*	*P (FDR)*
VOT	(Intercept)	12.30	3.04	6.33-18.26	**< 0.001**	
model statistics:	/ta/ syllable	5.85	1.04	3.82-7.88	**< 0.001**	**0.006**
σ^2 = 11.79	/ka/ syllable	14.91	1.04	12.87-16.94	**< 0.001**	**0.006**
τ_{00} = 13.16	PTA	−0.03	0.05	−0.13-0.07	0.572	0.686
ICC = 0.53	Speech Rate	−0.36	0.22	−0.79-0.08	0.109	0.196
N =127	GDS	0.16	0.14	−0.12-0.44	0.268	0.438
Observations =1112	Perceived Physical Condition	0.12	0.39	−0.64-0.88	0.760	0.760
Marginal R^2	Social Participation	−0.08	0.11	−0.30-0.13	0.463	0.641
Conditional R^2 = 0.505 / 0.766	Block [Block 2]	0.46	0.25	−0.03-0.96	0.068	0.136
AIC = 6271.984	Block [Block 3]	1.22	0.26	0.71-1.73	**< 0.001**	**0.006**
	Group [Singer]	1.54	2.20	−2.77-5.85	0.483	0.641
	Age	0.01	0.03	−0.05-0.07	0.702	0.743
	/ta/ × Age	−0.02	0.02	−0.06-0.02	0.335	0.515
	/ka/ × Age	−0.06	0.02	−0.10-−0.02	**0.001**	**0.005**
	/ta/ × Group [Singer]	−3.42	1.49	−6.34-−0.50	**0.022**	**0.049**
	/ka/ × Group [Singer]	−3.62	1.49	−6.54-−0.69	**0.015**	**0.045**
	Group [Singer] × Age	−0.02	0.04	−0.10-0.06	0.622	0.699
	/ta/ × Group [Singer] × Age	0.06	0.03	0.01-0.12	**0.020**	**0.049**
	/ka/ × Group [Singer] × Age	0.08	0.03	0.03-0.13	**0.003**	**0.010**
VOT-CoV ($\sqrt{}$)	(Intercept)	4.46	0.41	3.67-5.26	**< 0.001**	
model statistics:	PTA	0.00	0.01	−0.01-0.01	0.804	0.869
Observations = 1111	Speech Rate	−0.06	0.04	−0.14-0.02	0.166	0.490
R^2 / R^2 adjusted = 0.028 / 0.019	GDS	−0.02	0.02	−0.06-0.01	0.196	0.490
AIC = 3834.844	/ta/ syllable	0.02	0.14	−0.25-0.30	0.869	0.869
	/ka/ syllable	−0.34	0.14	−0.61-−0.06	**0.015**	0.075
	Perceived Physical Condition	−0.01	0.05	−0.10-0.08	0.795	0.869
	Social Participation	−0.00	0.01	−0.03-0.02	0.866	0.869
	Group [Singer]	0.11	0.14	−0.16-0.39	0.416	0.693
	/ta/ × Group [Singer]	−0.50	0.20	−0.89-−0.11	**0.012**	0.075
	/ka/ × Group [Singer]	−0.20	0.20	−0.59-0.19	0.321	0.642

Notes: Significant P-values of predictors are bolded.
Abbreviations: AIC, Akaike Information Criterion; b, estimate; CI, confidence interval of the estimate; GDS, Geriatric Depression Scale; ICC, Intraclass Correlation Coefficient; N, Number of participants; P (FDR), P value after applying FDR correction; PTA, average of the best pure tone average thresholds; SE, standard error of the estimate; σ2, Residual variance; τ00, Random effect variance; VOT-CoV ($\sqrt{}$), square-root transformed VOT-CoV data.

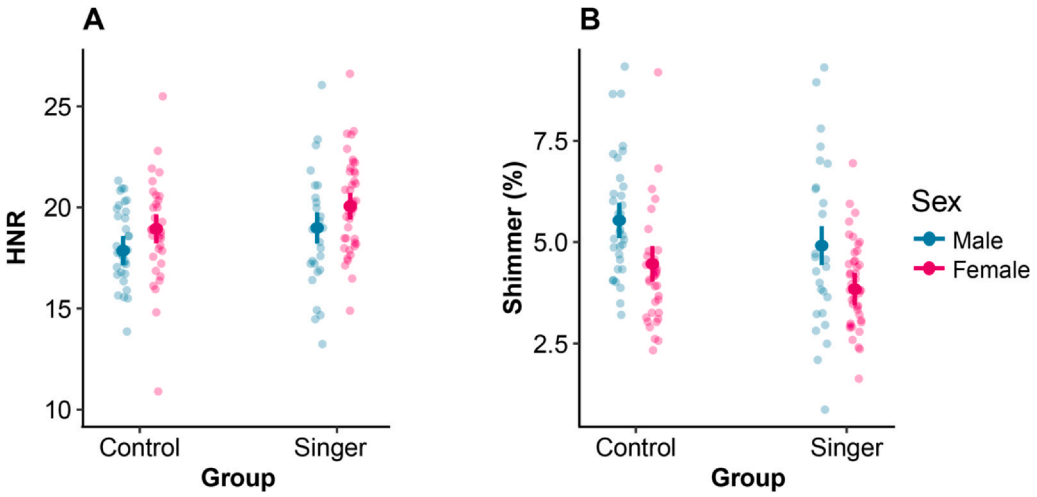


FIGURE 1. *Note*. Results from linear regression models, shown with individual data points color-coded by sex, for **A.** HNR across groups and **B.** Shimmer across groups. The larger central point indicates the mean for each group and sex; error bars represent the 95% confidence intervals. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

assumptions, buildmer linear regression models, FDR corrections), using the DSI and AVQI composite scores and their respective parameters as dependent variables. Results indicated that normalized singing experience was not associated with any DSI or AVQI measures. These analyses and results are reported in Supplementary Material 1.

Subjective assessment of vocal quality

Results of interrater reliability assessment for CAPE-V overall severity ratings showed moderate reliability (ICC [3,2] = 0.731, 95% CI [0.354, 0.887]).[78] For intrarater reliability, judge 1 demonstrated moderate reliability (ICC [3,1] = 0.693, 95% CI [0.408, 0.856]), while Judge 2 demonstrated excellent reliability (ICC [3,1] = 0.935, 95% CI [0.854, 0.972]).[78] Judge 2's ratings were therefore used in the analyses in all overlapping cases (ie, for samples rated by both judges).

Results of the linear regression model for subjective assessment of vocal quality are shown in Table 2. We analyzed the effects of age and group, and their interaction, on overall severity ratings (objectives 1B and 2). Three participants were excluded from this analysis as their data were identified as outliers. These individuals showed extreme overall severity values relative to the median of the full distribution.

The analysis revealed a main effect of group, with singers exhibiting lower severity ratings compared to controls (b = −0.12, SE = 0.05, P = 0.019). Additionally, a significant main effect of age was observed, with increasing age associated with higher severity ratings (b = 0.0016, SE = 0.0007, P = 0.019). A significant age-by-group interaction indicated that the relationship between age and severity differed across groups (b = 0.002, SE = 0.0009, P = 0.019) (Figure 2A). Slope analysis revealed two significant slopes, indicating that age was positively associated with severity ratings in both groups, but overall severity increased more sharply with age among singers (b = 0.0038, SE = 0.0006, P < 0.0001) when compared to controls (b = 0.0016, SE = 0.0006, P = 0.021). Figure 2A shows a trend toward lower severity ratings in younger and middle-aged singers, with this apparent advantage diminishing and reversing in older singers.

Vocal motor control

Results of the VOT-CoV interrater reliability assessment indicated moderate to good reliability (ICCs [3,1] > 0.737).[78] The intrarater reliability assessment showed excellent reliability (ICCs [3,1] > 0.964).[78]

Results of the mixed-effects regression models for vocal motor control are shown in Table 3. We examined the effects of group and age, and their interaction, on vocal motor control using VOT and VOT-CoV measurements (objectives 1C and 2). For these analyses, data from up to 16 participants (ie, up to 1.9% of all VOT/VOT-CoV data) were excluded due to incorrect execution of the /pataka/ syllable sequence for at least one block of the task.

VOT

Regression analysis revealed that both /ta/ (b = 5.85, SE = 1.04, P = 0.006) and /ka/ (b = 14.91, SE = 1.04, P = 0.006) syllables were associated with significantly longer VOT durations compared to the reference syllable /pa/. Significant three-way interactions between syllable, group, and age were found for both /ta/ and /ka/ (P = 0.049 and P = 0.010, respectively) (Figure 2B). *Post hoc* tests revealed no significant differences between groups when the slopes of VOT by age were compared for each syllable individually (all P > 0.362), although distinct trends were noted: slightly longer VOT durations in singers with age for /ta/ and /ka/ compared to slightly lower or stable VOT

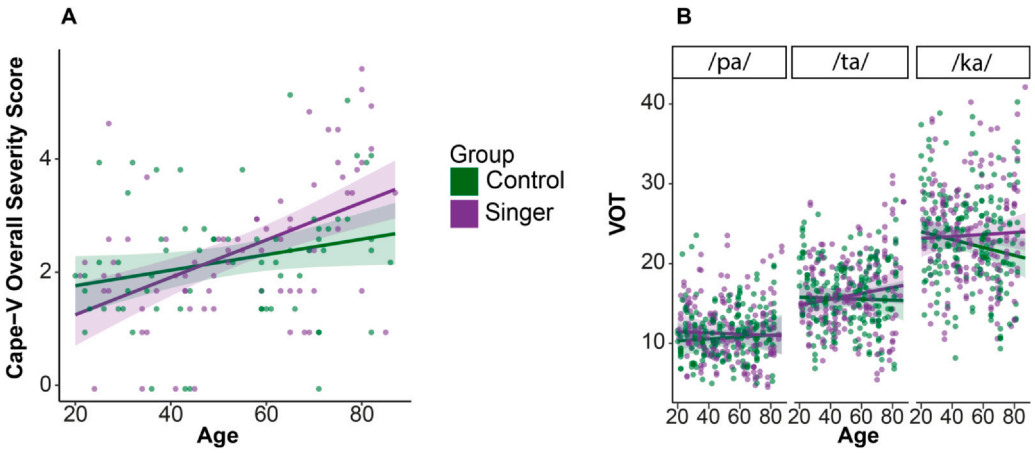


FIGURE 2. *Note*. Interaction plots with individual data points color-coded by group, where shaded area around the lines represents the 95% confidence interval of the regression line. **A**. Interaction between age (in years) and group (control, singer); **B.** Interaction between age (in years), group (control, singer), and syllable (/pa/, /ta/ and /ka/). VOT values are expressed in milliseconds. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

durations with age in non-singers. For /pa/, both groups showed similar trends. None of the individual slopes reached significance (all $P > 0.701$).

VOT-CoV

The analysis for VOT-CoV revealed a significant main effect of the /ka/ syllable and a significant two-way interaction between syllable and group. However, the model explained little variance ($R^2 = 0.028$, adjusted $R^2 = 0.019$), and none of these effects survived FDR correction. See Table 3 for details. See also Figure 3 for summarized results.

DISCUSSION

This study aimed to examine whether regular amateur singing practice influences the adult speaking voice and whether it moderates any age-related effects across *(1)* objective measures of vocal function and quality; *(2)* subjective assessment of vocal quality; and *(3)* vocal motor control. Our first hypothesis was that adults who engage in amateur singing would demonstrate enhanced vocal function, vocal quality, and articulatory-laryngeal coordination compared to non-singers. Our second hypothesis was that singers would show an attenuated association with age in these domains compared to non-singers. This section is structured around these hypotheses: group effects are discussed first, followed by age and age-by-group interaction effects. For the sake of concision, only effects of interest and significant effects that survived FDR correction are discussed.

The effect of singing status

For objective measures of vocal function and quality, no main effect of group emerged for the global DSI and AVQI scores, but analyses revealed that singers exhibited higher

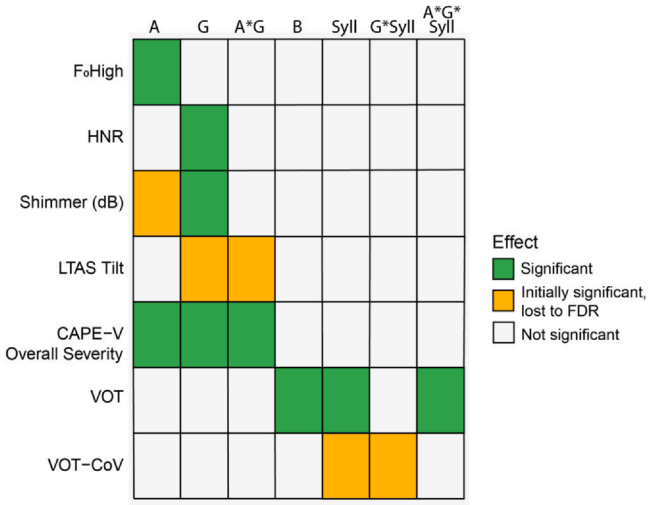


FIGURE 3. *Note.* Summary of main findings. A = age, G = group, B = block, Syll = syllable (/pa/, /ta/ or /ka). Green indicates a significant effect with FDR correction, orange indicates a significant effect prior to FDR correction only, and light gray indicates the absence of an effect. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

HNR and lower shimmer values (both in % and dB) compared to non-singers (Figure 1). A limited number of studies have compared multiparametric acoustic measures between singers and non-singers,[5,6] and one has examined differences between amateur singers (including choir singers) and non-singers.[14] Collectively, these studies report higher DSI scores in singers compared to controls. While amateur singers in our sample showed an advantage on some individual acoustic measures, differences in overall vocal quality and function compared to non-singers may not have been important enough to show through multiparametric measures. The lack of a clear and consistent

definition of amateur singing may partly explain discrepancies between studies. For example, in the study by Awan and Ensslen,[14] amateur singers reportedly engaged in more frequent and intensive singing practice compared to the amateur singers in our sample. This may suggest a dose-response association of singing with voice quality and function, as detectable through multiparametric measures, although our exploratory analyses did not reveal any advantages for singers with more experience. Discrepancies between our findings and previous studies reporting higher DSI scores in singers and amateur singers may be attributed to differences in participants' sociodemographic characteristics, including age range, which should be consistently accounted for, and from limited information on whether singer and non-singer groups were matched on other relevant sociodemographic factors (eg, physical condition, depression symptoms). Nonetheless, the higher HNR and lower shimmer values in our sample of amateur singers are in line with expectations of enhanced vocal quality in singers and suggestive of a clearer and more stable voice. These results are also consistent with previous studies reporting lower shimmer[13,39] and jitter values (ie, reduced period instability) in amateur singers[14,37] during a sustained vowel task. However, it is worth noting that other studies have found no differences in shimmer values between singers and non-singers during speech tasks.[37,88]

Regarding the subjective assessment of vocal quality, a group effect was found for the overall severity rating of the CAPE-V. Consistent with results from the acoustic measures, the analysis revealed that singers exhibited lower ratings of overall severity compared to non-singers. However, since it also revealed a significant age-by-group interaction, these findings will be addressed in the Moderating Effects section.

No group effect was found for the habitual intensity measure, suggesting that amateur choir singing may not be associated with differences in vocal intensity within the context of spontaneous speech. The literature assessing intensity measures in choir singers is limited. Moon et al[38] found no improvement in mean intensity during sustained /a/ phonation in a "general singing group" receiving no specific instructions and training on posture, breath control, and pitch modulation, compared to a "therapeutic singing group" that completed a 12-session program including vocal and respiratory exercises, and did improve. This may point to a dose-response effect or a training-related effect of singing, and therefore vocal intensity may not represent the most sensitive marker of vocal changes in amateur choir singers without formal training or without reported vocal deficits (eg, Parkinson disease, symptomatic presbyphonia).

Finally, no main effect of singing was found on vocal motor control measures in voiceless plosives in our sample. However, three-way interactions between syllable, group, and age were found and will be addressed in the next section.

Taken together, our results suggest that amateur singing may be associated with subtle differences in specific acoustic features, including increased HNR and reduced shimmer, partly supporting our group-effect hypothesis. However, inconsistencies in the literature (eg, different age ranges) and variability in singing practice (ie, frequency, intensity, and nature of vocal training) highlight the challenges of isolating the effect of amateur singing on voice measures. By identifying the most sensitive measures in a relatively large sample, our study contributes to a more nuanced understanding of how vocal practice may shape acoustic outcomes in amateur singers.

The effect of age and the moderating effect of singing

Our analyses revealed no age effect or age-by-group interaction on overall DSI and AVQI scores in a sample of healthy adults with high self-perceived physical condition and no depressive symptoms. This result contrasts with prior literature reporting age-related declines in DSI scores[19,89,90] and AVQI scores.[91] Although sample sizes and age ranges were comparable across these studies and ours, differences in eligibility criteria (eg, presence and type of voice screening) and statistical approaches (eg, regression models with age as a continuous variable vs. comparison of categorical age groups) may have contributed to the noted discrepancies. Conversely, our findings align with previous literature reporting limited or inconsistent evidence regarding the impact of age on the AVQI (for a review, see Jayakumar and Benoy[92]).

Despite the lack of an age effect on acoustic measures of vocal quality, perceptual judgments of overall severity ratings were found to increase with age in both groups, which was expected considering documented age-related increases in strain, breathiness, and hoarseness.[21,22,93] The discrepancy between objective and subjective measures in this study may partly be explained by the fact that perceptual judgments were based on the full reading of a standardized passage, whereas acoustic measures were extracted only from the passage's first sentence, as well as from single vowels. The full passage provides a more dynamic and variable voice sample (ie, more prosodic variations, respiratory management, and articulatory transitions) which may increase sensitivity to more subtle alterations in non-disordered voices. This sensitivity interpretation aligns with the findings of Batthyany et al[94] who reported that the AVQI and CPPs did not discriminate between perceptually normophonic and very slightly dysphonic participants without self-reported voice complaints.

A moderating effect of singing was found for ratings of overall severity, showing a trend toward better perceptual vocal quality in younger and middle-aged singers, which was not observed in older singers. Therefore, while severity ratings increased with age in both groups, the association between age and severity was steeper among singers (Figure 2A). Interestingly, analyses indicated a similar interaction pattern for the LTAS tilt, but this effect did not survive FDR correction. The lack of an apparent advantage in vocal quality in older singers may reflect a combination of contextual factors, such as vocal hygiene, singing practices,

and high vocal demands, as well as age-related physiological changes. Indeed, we lack information on how amateur singers in our sample cared for their voice. Studies indicate that many choir singers receive little to no training in vocal hygiene,[95–97] and that vocal warm-up and cooldown exercises are often absent or inadequate in community choir settings (eg, not long enough/too long, vocal rest as a cooldown).[97,98] Additionally, choir singers seemed more susceptible to vocal complaints or signs of vocal effort after singing,[99,100] and results from Moura et al[101] and Cohen et al[102] showed that age is an important factor influencing difficulties reported by singers on the Singing Voice Handicap Index (SVHI). In both studies, older age was associated with greater self-reported vocal difficulties, like increased vocal fatigue, decreased vocal stability, and greater hoarseness compared to younger singers. These features correspond to key perceptual dimensions contributing to overall vocal severity ratings on the CAPE-V scale. Therefore, the steeper association between overall severity and age observed in singers compared with non-singers in our sample is consistent with the nature of the vocal difficulties reported by aging singers. It is possible that physiological changes in the laryngeal and respiratory systems affect the ability of older amateur singers to meet high vocal demands effectively, especially if combined with limited vocal hygiene knowledge and suboptimal vocal practices. Bearing that in mind, a longitudinal study by Stager et al[41] compared less-experienced and more-experienced older singers enrolled in a highly structured choral program: weekly 90-minute rehearsals for 15 weeks, standardized vocal warm-ups, consistent repertoire, professional conductors, and guided home practice. Several aerodynamic and acoustic measures improved and were maintained over time, leading the authors to conclude that regular singing might be beneficial for preserving speaking voice over time in older adults. Notably, in the present study, the interaction suggests a trend towards lower severity ratings (ie, better) in singers compared to non-singers prior to the crossover point (ie, in younger and middle-aged participants)—consistent with the group effects found for subjective and objective measures of vocal quality. This pattern might reflect a protective effect of amateur singing on vocal quality, at least up to midlife, and raises the possibility that early singing-related advantages may not be found with aging in all singing contexts. The present study focuses on community-based choral singing, a more ecological context (ie, what older adults would typically encounter when joining a local community choir), where rehearsal structures and vocal hygiene habits may not have been standardized or monitored. This context would be consistent with broader reports of limited access to education on vocal hygiene and good practices within community choirs.[95–97] Therefore, this difference in regimen may contribute to the divergent outcomes between our study and Stager et al's.[41]

To further characterize how aging influences vocal function, we tested the effect of age and the moderating effect of singing on selected measures of frequency and intensity. These analyses revealed an age effect on f_0High, consistent with previous studies reporting a decrease in f_0High among older adults,[5,89] although others found no significant age-related differences.[19,39] Contrary to our hypothesis, this effect was not moderated by singing. Increasing pitch to its highest level requires fine laryngeal muscle adjustments and precise neuromuscular control, both of which become less efficient as age-related changes in the larynx affect muscle and neuromuscular junction properties.[16,103] Our results suggest that amateur choir singing may not provide sufficient training to moderate the effects of aging on this specific skill. As for habitual intensity, no age-related or moderating effects were observed. Previous literature on age-related changes in vocal intensity and loudness is mixed; some studies have reported lower habitual intensity in older adults during a reading passage task[19] and a sustained syllable production,[104] whereas others have found no age-related differences.[105] Although reduced loudness and breathiness are commonly reported complaints among older adults with age-related vocal symptoms,[22] it is important to note that voice problems do not affect all older adults; estimates of their prevalence suggest that just under one-fifth of the aging population presents with vocal symptoms.[106] In the present study, participants did not report any voice complaints prior to testing, which may partly explain the absence of age effects on habitual intensity.

In terms of voice motor control, we found no main effect of age on VOT duration or variability, suggesting that age alone does not predict changes in articulatory-laryngeal timing, at least in French voiceless plosives. This aligns with several studies in English and French reporting no overall age-related differences in VOT means when considering all voiceless plosives together.[29–31,107] However, our analysis revealed three-way interactions between syllable, group, and age, suggesting that the effect of singing on VOT in aging may be syllable-specific. In fact, distinct trends across groups were noted. For /ta/ and /ka/, VOTs tended to be slightly longer with age in singers, whereas they tended to be stable or slightly shorter in non-singers. For /pa/, VOT remained relatively stable with age across both groups (Figure 2B). The relationship between VOT and age remains unclear in the literature, but a previous study examining VOT in French speakers did report age-related decreases in voiceless plosives.[108] Age-related changes in VOT are coherent with broader evidence on age-related changes in speech motor control, likely due to reduced neuromotor control with age.[23,26,28,109] Moreover, the distinct group trends observed only for the syllables /ta/ and /ka/, but not /pa/, may reflect the articulatory demands specific to alveolar and velar plosives. Both /t/ and /k/ require precise tongue movements, which have been shown to slow with age and have been proposed as contributors to age-related intelligibility decline.[110] Furthermore, some studies have reported differences in articulatory-laryngeal timing between singers and non-singers, at least for certain

plosives. However, the direction of these differences remains unclear: a study found shorter VOTs in singers,[9] whereas another reported longer VOTs,[10] and another observed no difference in VOT durations.[12] Despite these discrepancies, the distinct trends noted across groups in the present study are consistent with earlier findings showing articulatory differences between groups in a DDK task—specifically, greater articulatory accuracy in amateur singers compared with non-singers.[15] Since articulation relies on finely timed movements,[111] singing may confer an advantage in precise temporal coordination, but additional studies are needed to clarify the extent and conditions under which such an advantage might emerge.

In summary, our findings revealed an advantage for singers for specific vocal features, like HNR and shimmer, but also age-related effects for both groups, including a gradual decline in f_0High and an increase in overall severity judgments of voice quality. These results suggest that singing and aging influence certain aspects of vocal production, though not uniformly across all measures or contexts. Overall severity judgments suggest that amateur singing may be associated with perceptual benefits in early to mid-adulthood, but these advantages were not observed in older singers. This may reflect a combination of age-related physiological changes and contextual factors (eg, vocal demands, vocal hygiene), a hypothesis that warrants empirical investigation. Additionally, results revealed syllable-specific and group-dependent trends in VOTs across age, suggesting differences in articulatory-laryngeal coordination between singers and non-singers, at least in some phonetic contexts. More evidence is needed to corroborate these trends and clarify their implications for intelligibility. While most research on singing focuses on acoustic outcomes, incorporating a motor control dimension provides a more comprehensive picture of its potential benefits on communicative functions.

LIMITATIONS

Some limitations of this study should be acknowledged. First, limited information regarding the regimen of participants' choir rehearsals and home practice restricts our ability to evaluate our hypothesis about vocal hygiene and practices within choir settings (eg, warm-up/cool-down routines, if any). Another limitation is the absence of objective health assessments, given, for example, the key role of the respiratory system in phonation and its influence on vocal performance. Although we included a self-reported physical condition rating, no objective measures of respiratory function were collected. Finally, although we included self-reported symptom screening and perceptual assessments of voice quality, no laryngeal examination was conducted to verify whether participants were free of vocal pathology. The absence of such an examination limits our ability to confirm the vocal status of the sample, including the presence or degree of vocal fold atrophy, which can occur even in the absence of voice symptoms.

CONCLUSION

This study provides new insights into the complex relationships between amateur singing, vocal aging, and vocal motor control, using a well-characterized, relatively large sample of participants without vocal complaints, and with high self-perceived physical condition and no depressive symptoms. Our results show that, while amateur singers showed specific acoustic and perceptual voice advantages, some of these benefits were not found in older amateur singers. Such a reduced advantage may be due to physiological changes in the laryngeal and respiratory systems, combined with high vocal demands in the context of potentially suboptimal vocal hygiene practices. This underscores the importance of documenting participants' singing regimen and vocal habits in future studies, as such information could provide valuable context for interpreting observed patterns. Notably, the National Association of Teachers of Singing (NATS)[112] includes vocal hygiene and good vocal health practices as key components in its syllabi for Science-Informed Voice Pedagogy courses, highlighting the relevance of these variables for understanding vocal behavior. Beyond vocal quality outcomes, our results show that amateur singing may be associated with distinct syllable-specific trends in articulatory-laryngeal coordination, a dimension that has received limited attention in previous research. Our results emphasize the need for more systematic investigations into the effects of singing across the lifespan, including large-scale studies that assess and compare singing parameters and vocal hygiene practices, as well as randomized singing training studies. Identifying strategies to optimize the potential benefits of amateur singing could help inform evidence-based guidelines to promote community-based singing as a preventive tool against age-related vocal decline.

CRediT authorship contribution statement

Amélie Nadeau: Investigation, Software, Formal Analysis, Visualization, Writing—Original Draft. **Pascale Tremblay**: Conceptualization, Data Curation, Funding Acquisition, Investigation, Methodology, Project Administration, Supervision, Validation, Resources, Writing—Review & Editing. **Maude Desjardins**: Conceptualization, Formal Analysis, Methodology, Project Administration, Resources, Software, Supervision, Validation, Writing—Review & Editing.

Data Availability

Individual data will not be publicly available, as the informed consent obtained from participants did not include permission for public data sharing. However, aggregated behavioral data will be available on Borealis, the Canadian Dataverse Repository (https://borealisdata.ca/fr/), upon publication, along with analytical scripts.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this article, the first author used Microsoft Copilot to assist with the translation of specialized terminology, grammar and spelling corrections, and the creation and verification of parts of the analytical scripts. All authors reviewed and edited the content and take full responsibility for the content of the published article.

Acknowledgments

This research was supported by P.T.'s grants from the Drummond Foundation [2016RFA proposal #27], the Social Sciences and Humanities Research Council [#435-2023-0578], and the Canadian Foundation for Innovation [#31,408]. P.T. holds a Canada Research Chair in the Neurobiology of Speech and Hearing [CRC #2022-00090]. M.D. grants from the Réseau provincial de recherche en adaptation-réadaptation (REPAR) and from the Institut de recherche Robert-Sauvé en santé et en sécurité du travail (IRSST) provided support to A.N, who was also supported by a master's research scholarship from the Canadian Institutes of Health Research.

We thank all the participants for their valuable contributions. We also thank Pr. Josée Vaillancourt for characterizing singers, and Maxime Perron, Julie Poulin, Lisa-Marie Deschênes, Anne-Christine Bricaud, Elena Vaccaro, and Antoine Halbaut for their contributions to recruitment and testing. We also thank Rosalie Gagné and Billie-Tanisha Moar for CAPE-V assessments of overall severity, Theodora Nestorova for the training in VOT segmentation, Florence Caron for VOT segmentation and reliability assessment, as well as all other research assistants and research interns who contributed to previous phases of the project. Finally, we would like to acknowledge the contribution of Youri Maryn for guidance related to the DSI and AVQI scripts, and of Timothy Pommée, for reviewing a version of this manuscript.

Appendix A

Details on questionnaires and composite score of social participation

1) The sociodemographic questionnaire included items on social participation—such as involvement in group activities (eg, bingo, museum tours) and informal social interactions (eg, coffee with friends, social gatherings). The questionnaire also included questions about current or former occupation as well as perceived physical condition, rated on a scale from 1 (poor) to 7 (excellent);
2) A questionnaire assessed past and present musical and singing experiences, which was used to determine group assignments;
3) The long version of the Geriatric Depression Scale (GDS) is a 30-item yes/no questionnaire where each "negative" response is scored as one point, and higher scores suggest more depressive symptoms (0–9 = normal, 10–19 = suggest mild depression, 20–30 = suggest severe depression)[48];
4) The Montreal Cognitive Assessment (MoCA) is a brief screening tool scored out of 30, where higher scores reflect better cognitive performance[49];
5) The Quebec French version of the Voice Handicap Index (VHI-QF) is a self-reported 30-item questionnaire scored on a 0 to 4-point scale assessing the functional impact of one's voice on daily life, with higher scores indicating a greater perceived voice-related handicap[50,51];
6) The composite score of social participation ranged from 0 to 17. It was derived from four criteria: living situation (alone vs. with others), employment status (retired/unemployed vs. employed/student), weekly frequency of participation in organized group activities (eg, bingo, museum tours), and weekly frequency of social interactions lasting one hour or more (eg, coffee meetings, social gatherings). Higher scores indicate greater social engagement (Table A.1 for more details).

TABLE A1.

Components and Scoring of the Social Participation Composite Score

Factor	Response Options	Number of Points
Living arrangement	Lives alone / Lives with others	0 / 1
Employment status	Retired, unemployed / Employed, student	0 / 1
Weekly interactions ≥ 1 h	Less than once every 2 weeks	0
	Once every 2 weeks	0.5
	≥1 per week	Number of interactions[a]
Weekly group activities ≥ 1 h	Less than once every 2 weeks	0
	Once every 2 weeks	0.5
	≥1 per week	Number of interactions[a]
Total		

Note: In our sample, scores ranged from 0 to 17, with higher scores reflecting greater social engagement.

[a] Number of interactions = the exact number of interactions reported by the participants.

Appendix B

Results of reliability assessments (interrater and intrarater) using intraclass correlation coefficients (ICCs) and their 95% confident intervals for CAPE-V overall severity ratings and VOT segmentations.

TABLE B1.

Results Of Interrater Reliability Assessment of the CAPE-V Overall Severity Ratings

	Intraclass Correlation	95% Confidence Interval		*F* Test			
		Lower Bound	Upper Bound	Value	df1	df2	Sig.
Mean ratings	0.731	0.354	0.887	3.6	22	22	0.002

Note: Results of ICC calculation assessing interrater reliability of the CAPE-V overall severity ratings using mean-rating (*k*=2), absolute agreement, two-way mixed-effects model (ICC [3,2]).

TABLE B2.

Results Of Intrarater Reliability Of The CAPE-V Overall Severity Ratings

	Intraclass Correlation	95% Confidence Interval		*F* Test			
		Lower Bound	Upper Bound	Value	df1	df2	Sig.
Juge 1							
Single rater	0.693	0.408	0.856	5.76	22	22.4	0.000
Juge 2							
Single rater	0.935	0.854	0.972	30.1	22	22,9	0.000

Note: Results of ICC calculation assessing intrarater reliability of the CAPE-V overall severity ratings for both judges using single-rater, absolute agreement, 2-way mixed-effects model (ICC [3,1]).

TABLE B3.

Results Of Interrater Reliability Of The VOT-CoV

	Intraclass Correlation	95% Confidence Interval		*F* Test			
		Lower Bound	Upper Bound	Value	df1	df2	Sig.
VOT-CoV /pa/							
Single rater	0.737	0.465	0.882	6.6	21	21	0.000
VOT-CoV /ta/							
Single rater	0.853	0.679	0.936	12.6	21	21	0.000
VOT-CoV /ka/							
Single rater	0.913	0.802	0.963	21.9	21	21	0.000
VOT-CoV							
Single rater	0.752	0.491	0.889	7.06	21	21	0.000

Note: Results of ICC calculation assessing interrater reliability of the VOT-CoV using single rater, consistency, two-way mixed-effects model (ICC [3,1]).

TABLE B4.
Results of Intrarater Reliability of the VOT-CoV

	Intraclass Correlation	95% Confidence Interval		*F* Test			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
VOT-CoV /pa/							
Single rater	0.975	0.940	0.989	7404	21	21	0.000
VOT-CoV /ta/							
Single rater	0.979	0.951	0.991	94.3	21	21	0.000
VOT-CoV /ka/							
Single rater	0.964	0.914	0.959	51.4	21	21	0.000
VOT-CoV							
Single rater	0.975	0.940	0.989	75.2	21	21	0.000

Note: Results of ICC calculation assessing intrarater reliability of the CoV calculations using single rater, absolute agreement, two-way mixed-effects model (ICC [3,1]).

Appendix C

Significant sex differences in DSI parameters: DSI global scores, MPT, jitter, ILow and *f*₀High

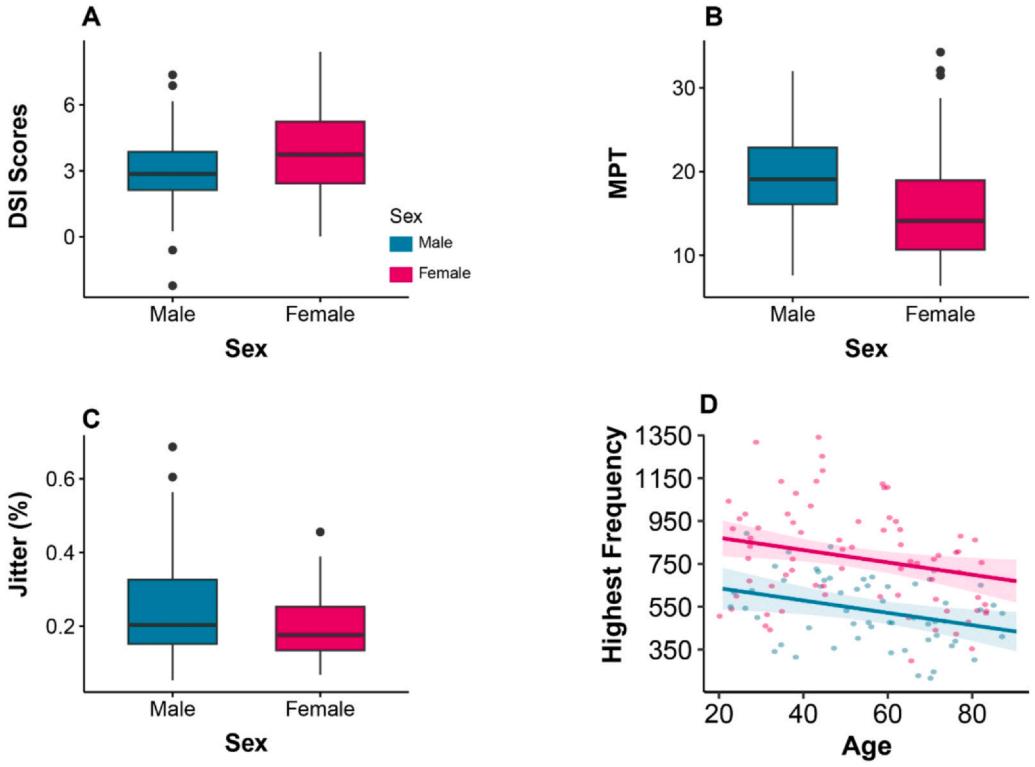


Figure C1. *Note*. **A-C.** Boxplots illustrate sex differences in vocal function parameters: overall Dysphonia Severity Index (DSI) score, maximum phonation time (MPT) (in seconds), and jitter (%). The boxes show the interquartile range, in which the middle horizontal line is the median (50th percentile); **D.** The scatter plot shows the relationship between age and highest attainable fundamental frequency (in Hz) (*f*₀High) separately for men and women; the shaded area around the lines represents the 95% confidence interval of the regression line. Each dot represents one participant.

Appendix D. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jvoice.2026.09.019.

References

1. Lee SH, Kwon HJ, et al. The singer's formant and speaker's ring resonance: a long-term average spectrum analysis. *Clin Exp Otorhinolaryngol.* 2008;1:92–96. https://doi.org/10.3342/ceo.2008.1.2.92.
2. Oliveira Barrichelo VM, Heuer RJ, et al. Comparison of singer's formant, speaker's ring, and LTA spectrum among classical singers and untrained normal speakers. *J Voice.* 2001;15:344–350. https://doi.org/10.1016/s0892-1997(01)00036-4.
3. Mendes AP, Rothman HB, et al. Effects of vocal training on the acoustic parameters of the singing voice. *J Voice.* 2003;17:529–543. https://doi.org/10.1067/s0892-1997(03)00083-3.
4. Sreeparvathi A, Sheela S, et al. Influence of carnatic vocal training on voice measures in males. *Folia Phoniatr Logop.* 2025;77:193–204. https://doi.org/10.1159/000541215.
5. Maruthy S, Ravibabu P. Comparison of Dysphonia Severity Index between younger and older carnatic classical singers and nonsingers. *J Voice.* 2015;29:65–70. https://doi.org/10.1016/j.jvoice.2014.05.001.
6. Hosseiniheshmatian M, Darouie A, et al. A comparison of the Dysphonia Severity Index and Singing Voice Handicap Index between trained and untrained male Iranian traditional singers. *J Voice.* 2024;40:1239.e1–1239.e10. https://doi.org/10.1016/j.jvoice.2024.02.032.
7. Pestana PM, Vaz-Freitas S, et al. Prevalence of Voice disorders in singers: systematic review and meta-analysis. *J Voice.* 2017;31:722–727. https://doi.org/10.1016/j.jvoice.2017.02.010.
8. Abramson AS, Whalen DH. Voice onset time (VOT) at 50: theoretical and practical issues in measuring voicing distinctions. *JPhon.* 2017;63:75–86. https://doi.org/10.1016/j.wocn.2017.05.002.
9. McCrea CR, Watts C. Relations of singing talent with voice onset time of trained and untrained female singers. *Percept Mot Skills.* 2007;105:133–142. https://doi.org/10.2466/pms.105.1.133-142.
10. McCrea CR, Morris RJ. Comparisons of voice onset time for trained male singers and male nonsingers during speaking and singing. *J Voice.* 2005;19:420–430. https://doi.org/10.1016/j.jvoice.2004.08.002.
11. McCrea CR, Morris RJ. Effects of vocal training and phonatory task on voice onset time. *J Voice.* 2007;21:54–63. https://doi.org/10.1016/j.jvoice.2005.05.002.
12. McCrea CR, Morris RJ. Voice onset time for female trained and untrained singers during speech and singing. *J Commun Disord.* 2007;40:418–431. https://doi.org/10.1016/j.jcomdis.2006.10.004.
13. Ravi SK, Shabnam S, et al. Acoustic and aerodynamic characteristics of choral singers. *J Voice.* 2019;33:803.e801–803.e805. https://doi.org/10.1016/j.jvoice.2018.03.018.
14. Awan SN, Ensslen AJ. A comparison of trained and untrained vocalists on the Dysphonia Severity Index. *J Voice.* 2010;24:661–666. https://doi.org/10.1016/j.jvoice.2009.04.001.
15. Tremblay P, Gagnon L, et al. Speech production in healthy older adults with or without amateur singing experience. *J Speech Lang Hear Res.* 2023;66:4332–4352. https://doi.org/10.1044/2023_JSLHR-23-00126.
16. Mallick AS, Garas G, et al. Presbylaryngis: a state-of-the-art review. *Curr Opin Otolaryngol Head Neck Surg.* 2019;27:168–177. https://doi.org/10.1097/MOO.0000000000000540.
17. Lortie CL, Thibeault M, et al. Effects of age on the amplitude, frequency and perceived quality of voice. *Age (Dordr).* 2015;37:117. https://doi.org/10.1007/s11357-015-9854-1.
18. Dehqan A, Scherer RC, et al. The effects of aging on acoustic parameters of voice. *Folia Phoniatr Logop.* 2012;64:265–270. https://doi.org/10.1159/000343998.
19. Goy H, Fernandes DN, et al. Normative voice data for younger and older adults. *J Voice.* 2013;27:545–555. https://doi.org/10.1016/j.jvoice.2013.03.002.
20. Hodge FS, Colton RH, et al. Vocal intensity characteristics in normal and elderly speakers. *J Voice.* 2001;15:503–511. https://doi.org/10.1016/S0892-1997(01)00050-9.
21. Gorham-Rowan MM, Laures-Gore J. Acoustic-perceptual correlates of voice quality in elderly men and women. *J Commun Disord.* 2006;39:171–184. https://doi.org/10.1016/j.jcomdis.2005.11.005.
22. Rojas S, Kefalianos E, et al. How does our voice change as we age? A systematic review and meta-analysis of acoustic and perceptual voice data from healthy adults over 50 years of age. *J Speech Lang Hear Res.* 2020;63:533–551. https://doi.org/10.1044/2019_JSLHR-19-00099.
23. Bilodeau-Mercure M, Tremblay P. Age differences in sequential speech production: articulatory and physiological factors. *J Am Geriatr Soc.* 2016;64:e177–e182. https://doi.org/10.1002/hbm.23558.
24. Fougeron C, Guitard-Ivent F, et al. Multi-dimensional variation in adult speech as a function of age. *Languages.* 2021;6:176. https://doi.org/10.3390/languages6040176.
25. Ben-David BM, Icht M. Oral-diadochokinetic rates for Hebrew-speaking healthy ageing population: non-word versus real-word repetition. *Int J Lang Commun Disord.* 2017;52:301–310. https://doi.org/10.1111/1460-6984.12272.
26. Tremblay P, Deschamps I, et al. Aging of speech production, from articulatory accuracy to motor timing. *Psychol Aging.* 2018;33:1022. https://doi.org/10.1037/pag0000306.
27. Tremblay P, Deschamps I. Structural brain aging and speech production: a surface-based brain morphometry study. *Brain Struct Funct.* 2016;221:3275–3299. https://doi.org/10.1007/s00429-015-1100-1.
28. Sadagopan N, Smith A. Age differences in speech motor performance on a novel speech task. *J Speech Lang Hear Res.* 2013;56:1552–1566. https://doi.org/10.1044/1092-4388(2013/12-0293).
29. Neiman GS, Klich RJ, et al. Voice onset time in young and 70-year-old women. *J Speech Lang Hear Res.* 1983;26:118–123. https://doi.org/10.1044/jshr.2601.118.
30. Petrosino L, Colcord RD, et al. Voice onset time of velar stop productions in aged speakers. *Percept Mot Skills.* 1993;76:83–88. https://doi.org/10.2466/pms.1993.76.1.83.
31. Sweeting PM, Baken RJ. Voice onset time in a normal-aged population. *J Speech Lang Hear Res.* 1982;25:129–134. https://doi.org/10.1044/jshr.2501.129.
32. Larson GW, Hayslip B, et al. Changes in voice onset time in young and older men. *Educ Gerontol.* 1992;18:285–297. https://doi.org/10.1080/0360127920180310.
33. Bóna J. Voice onset time and speakers' age: data from Hungarian. *Clin Linguist Phon.* 2014;28:366–372. https://doi.org/10.3109/02699206.2013.875593.
34. Torre P, Barlow JA. Age-related changes in acoustic characteristics of adult speech. *J Commun Disord.* 2009;42:324–333. https://doi.org/10.1016/j.jcomdis.2009.03.001.
35. Kempster GB, Gerratt BR, et al. Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol. *Am J Speech Lang Pathol.* 2009;18:124–132. https://doi.org/10.1044/1058-0360(2008/08-0017).
36. Aquino FSd, Silva MAA, et al. Aspects of the speaking voice of elderly women with choral singing experience. *CoDAS.* 2016;28:446–453. https://doi.org/10.1590/2317-1782/20162015109.
37. Prakup B. Acoustic measures of the voices of older singers and nonsingers. *J Voice.* 2012;26:341–350. https://doi.org/10.1016/j.jvoice.2011.05.007.
38. Moon S, Park J, et al. The effects of therapeutic singing on vocal functions of the elderly: a study on Korean elderly. *J Voice.* 2022;36:437 e431–437 e439. https://doi.org/10.1016/j.jvoice.2020.06.007.
39. Lortie CL, Rivard J, et al. The moderating effect of frequent singing on voice aging. *J Voice.* 2017;31:112 e111–112 e112. https://doi.org/10.1016/j.jvoice.2016.02.015.
40. Stager SV, Sparks AD, et al. The role of choral singing in speaking voice preservation of aging adults. *J Speech Lang Hear Res.* 2020;63:2099–2114. https://doi.org/10.1044/2020_JSLHR-19-00347.

41. Stager SV, Goudelias D, et al. Effects of participation in a structured choral program on the voice of older adults. *J Voice.* 2025;39:1093–1101. https://doi.org/10.1016/j.jvoice.2023.02.018.
42. Yuen CW, Ma EP. Systematic review: singing-based interventions to improve physical functions related to aging voice in older adults. *J Speech Lang Hear Res.* 2024;67:2139–2158. https://doi.org/10.1044/2024_JSLHR-23-00641.
43. Tremblay P, Perron M. Auditory cognitive aging in amateur singers and non-singers. *Cognition.* 2023;230:105311. https://doi.org/10.1016/j.cognition.2022.105311.
44. Zhang X, Tremblay P. Aging of amateur singers and non-singers: from behavior to resting-state connectivity. *J Cogn Neurosci.* 2023;35:2049–2066. https://doi.org/10.1162/jocn_a_02065.
45. Marczyk A, Belley E, et al. Learning transfer from singing to speech: Insights from vowel analyses in aging amateur singers and non-singers. *Speech Commun.* 2022;141:23–39. https://doi.org/10.1016/j.specom.2022.05.001.
46. Perron M, Vaillancourt J, et al. Amateur singing benefits speech perception in aging under certain conditions of practice: behavioural and neurobiological mechanisms. *Brain Struct Funct.* 2022;227:943–962. https://doi.org/10.1007/s00429-021-02433-2.
47. Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia.* 1971;9:97–113. https://doi.org/10.1016/0028-3932(71)90067-4.
48. Yesavage JA, Brink TL, et al. Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res.* 1982;17:37–49. https://doi.org/10.1016/0022-3956(82)90033-4.
49. Nasreddine ZS, Phillips NA, et al. The montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc.* 2005;53:695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x.
50. Jacobson BH, Johnson A, et al. The Voice Handicap Index (VHI). *Am J Speech Lang Pathol.* 1997;6:66–70. https://doi.org/10.1044/1058-0360.0603.66.
51. Bourque JM, Defoy L, et al. Cross-cultural adaptation and validation of the Voice Handicap Index in the Quebec French population (VHI-QF). *J Voice.* 2020;34:811.e811–811.e816. https://doi.org/10.1016/j.jvoice.2019.04.010.
52. Mirza N, Ruiz C, et al. The prevalence of major psychiatric pathologies in patients with voice disorders. , 812, 814 *Ear Nose Throat J.* 2003;82:808–810. https://doi.org/10.1177/014556130308201015.
53. Shuji S, Mitsuteru N, et al. Depressive mood assessment method based on emotion level derived from voice: comparison of voice features of individuals with major depressive disorders and healthy controls. *Int J Environ Res Public Health.* 2021;18:5435. https://doi.org/10.3390/ijerph18105435.
54. Willinger U, Volkl-Kernstock S, et al. Marked depression and anxiety in patients with functional dysphonia. *Psychiatry Res.* 2005;134:85–91. https://doi.org/10.1016/j.psychres.2003.07.007.
55. Saito T, Murata C, et al. Influence of social relationship domains and their combinations on incident dementia: a prospective cohort study. *J Epidemiol Community Health.* 2018;72:7–12. https://doi.org/10.1136/jech-2017-209811.
56. Tremblay P, Gagnon L, et al. Exploring factors affecting verbal fluency in healthy aging. *Ann N Y Acad Sci.* 2025;1553:319–333. https://doi.org/10.1111/nyas.70064.
57. Perron M, Theaud G, et al. The frontotemporal organization of the arcuate fasciculus and its relationship with speech perception in young and older amateur singers and non-singers. *Hum Brain Mapp.* 2021;42:3058–3076. https://doi.org/10.1002/hbm.25416.
58. Association phonétique internationale. *The Principles Of The International Phonetic Association: Being A Description Of The International Phonetic Alphabet And The Manner Of Using It, Illus. By Texts In 51 Languages.* London: International Phonetic Association; 1949.
59. Boersma P, Van Heuven V. Speak and unSpeak with PRAAT. *Glot Int.* 2001;5:341–347.
60. Maryn Y, De Bodt M, et al. The value of the acoustic voice quality index as a measure of dysphonia severity in subjects speaking different languages. *Eur Arch Otorhinolaryngol.* 2014;271:1609–1619. https://doi.org/10.1007/s00405-013-2730-7.
61. Wuyts FL, De Bodt MS, et al. The Dysphonia Severity Index: an objective measure of vocal quality based on a multiparameter approach. *J Speech Lang Hear Res.* 2000;43:796–809. https://doi.org/10.1044/jslhr.4303.796.
62. Maryn Y, Corthals P, et al. Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels. *J Voice.* 2010;24:540–555. https://doi.org/10.1016/j.jvoice.2008.12.014.
63. M.S. Bodt Dysphonia Sev Index (DSI) Basics Clin Interpret 2010.
64. Maryn Y, Morsomme D, et al. Measuring the Dysphonia Severity Index (DSI) in the program Praat. *J Voice.* 2017;31:644.e629–644.e640. https://doi.org/10.1016/j.jvoice.2017.01.002.
65. Phonanium. Dysphonia Severity Index (v.02.03). v.02.03 ed2017.
66. Phonanium. Acoustic Voice Quality Index (v.02.06). v.02.06 ed2015.
67. Y. Maryn P. Corthals Acoust Voice Qual Index (V 02 06) Vol 2024 Phonanium 2014.
68. Maryn Y, Weenink D. Objective dysphonia measures in the program Praat: smoothed cepstral peak prominence and Acoustic Voice Quality Index. *J Voice.* 2015;29:35–43. https://doi.org/10.1016/j.jvoice.2014.06.015.
69. Neus Calaf All-Voice 2026.https://www.all-voiced.com/.
70. Awan SN, Lawson LL. The effect of anchor modality on the reliability of vocal severity ratings. *J Voice.* 2009;23:341–352. https://doi.org/10.1016/j.jvoice.2007.10.006.
71. Benoy JJ, Jayakumar T. Effect of anchor voices and listener expertise on auditory-perceptual judgments of voice quality using the GRBAS Scale. *J Voice.* 2024;40:694–705. https://doi.org/10.1016/j.jvoice.2023.12.011.
72. Pommée T, Renaud S-E, et al. Reliability and task effects in CAPE-V auditory-perceptual voice assessments: insights from the PVQD30 Subset. *J Voice.* 2025. https://doi.org/10.1016/j.jvoice.2025.02.020.In press.
73. Zraick RI, Kempster GB, et al. Establishing validity of the consensus auditory-perceptual evaluation of voice (CAPE-V). *Am J Speech Lang Pathol.* 2011;20:14–22. https://doi.org/10.1044/1058-0360(2010/09-0105).
74. Heller Murray E. Conducting high-quality and reliable acoustic analysis: a tutorial focused on training research assistants. *J Acoust Soc Am.* 2024;155:2603–2611. https://doi.org/10.1121/10.0025536.
75. Lisker L, Abramson AS. A cross-language study of voicing in initial stops: acoustical measurements. *Word.* 1964;20:384–422. https://doi.org/10.1080/00437956.1964.11659830.
76. Volaitis LE, Miller JL. Phonetic prototypes: Influence of place of articulation and speaking rate on the internal structure of voicing categories. *J Acoust Soc Am.* 1992;92:723–735. https://doi.org/10.1121/1.403997.
77. Posit Team. *RStudio: Integrated Development Environment for R. Version 2025.5.0.496*. Boston, MA: Posit Software; 2025.
78. Koo TK, Li MY. A Guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med.* 2016;15:155–163. https://doi.org/10.1016/j.jcm.2016.02.012.
79. Gamer M., Lemon J., et al. irr: Various Coefficients of Interrater Reliability and Agreement. Version 0.84.1. 2019.
80. C. Cesko Voeten. Buildmer: Stepwise Elimination and Term Reordering for Mixed-Effects Regression Version 2 11 2023.
81. Douglas B, Martin M, et al. Fitting linear mixed-effects models using lme4. *J Stat Softw.* 2015;67:1–48. https://doi.org/10.18637/jss.v067.i01.
82. J. Long jtools Anal Present Social Sci Data 2022.
83. Barr DJ. Random effects structure for testing interactions in linear mixed-effects models. *Front Psychol.* 2013;4:328. https://doi.org/10.3389/fpsyg.2013.00328.

ARTICLE IN PRESS

84. D. Lüdecke sjPlot Data Visualization Statistics Social Science. Version 2. 8. 16 2024.
85. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer-Verlag; 2016.
86. J. Long Inter Compr Use-Friendly Toolkit Probing Inter 2024.
87. R Core Team. *R: A Language and Environment for Statistical Computing. Version 2025.05.0*. Vienna, Austria: R Foundation for Statistical Computing; 2024.
88. Brown WS, Rothman HB, et al. Perceptual and acoustic study of professionally trained versus untrained voices. *J Voice.* 2000;14:301–309. https://doi.org/10.1016/S0892-1997(00)80076-4.
89. Hakkesteegt MM, Brocaar MP, et al. Influence of age and gender on the dysphonia severity index: a study of normative values. *Folia Phoniatr Logop.* 2006;58:264–273. https://doi.org/10.1159/000093183.
90. Barsties v. Latoszek B, Ulozaitė-Stanienė N, et al. The influence of gender and age on the Acoustic Voice Quality Index and Dysphonia Severity Index: a normative study. *J Voice.* 2019;33:340–345. https://doi.org/10.1016/j.jvoice.2017.11.011.
91. Jayakumar T, Benoy JJ, et al. Effect of age and gender on Acoustic Voice Quality Index across lifespan: a cross-sectional study in Indian population. *J Voice.* 2022;36:436.e431–436.e438. https://doi.org/10.1016/j.jvoice.2020.05.025.
92. Jayakumar T, Benoy JJ. Acoustic Voice Quality Index (AVQI) in the measurement of voice quality: a systematic review and meta-analysis. *J Voice.* 2024;38:1055–1069. https://doi.org/10.1016/j.jvoice.2022.03.018.
93. Tunç Songur E, Hafizoğlu M, et al. Analysis of the auditory-perceptual voice quality in Older and younger adults without self-reported voice complaints. *J Voice.* 2025. https://doi.org/10.1016/j.jvoice.2024.12.022.In press.
94. Batthyany C, Maryn Y, et al. A case of specificity: how does the Acoustic Voice Quality Index perform in normophonic subjects? *Appl Sci.* 2019;9:2527. https://doi.org/10.3390/app9122527.
95. Braun-Janzen C, Zeine L. Singers' interest and knowledge levels of vocal function and dysfunction: survey findings. *J Voice.* 2009;23:470–483. https://doi.org/10.1016/j.jvoice.2008.01.001.
96. Ravall S, Simberg S. Voice disorders and voice knowledge in choir singers. *J Voice.* 2020;34:157.e151–157.e158. https://doi.org/10.1016/j.jvoice.2018.07.005.
97. Sharma V, Nayak S, et al. A survey of vocal health in church choir singers. *Eur Arch Otorhinolaryngol.* 2021;278:2907–2917. https://doi.org/10.1007/s00405-021-06770-0.
98. Onofre F, Prado Yd.A, et al. Measurements of the acoustic speaking voice after vocal warm-up and cooldown in choir singers. *J Voice.* 2017;31:129.e9–129.e14. https://doi.org/10.1016/j.jvoice.2015.12.004. 129.e129-129.e114.
99. Meerschman I, D'Haeseleer E, et al. Voice quality of choir singers and the effect of a performance on the voice. *J Voice.* 2025;39:561.e1–561.e10. https://doi.org/10.1016/j.jvoice.2022.08.017. 561.e561-561.e510.
100. Onofre F, Ricz H, et al. Vocal resistance among choir singers. *Eur Arch Otorhinolaryngol.* 2021;278:159–165. https://doi.org/10.1007/s00405-020-06238-7.
101. Moura WER, Gadenz CD, et al. Analysis of the singing voice handicap index for elderly choristers. *Codas.* 2021;34:e20200302. https://doi.org/10.1590/2317-1782/20212020302.
102. Cohen SM, Noordzij JP, et al. Factors associated with perception of singing voice handicap. *Otolaryngol Neck Surg.* 2008;138:430–434. https://doi.org/10.1016/j.otohns.2007.12.041.
103. Périé S, St Guily JL, et al. Innervation of adult human laryngeal muscle fibers. *J Neurol Sci.* 1997;149:81–86. https://doi.org/10.1016/s0022-510x(97)05395-1.
104. Baker KK, Ramig LO, et al. Control of vocal loudness in young and old adults. *J Speech Lang Hear Res.* 2001;44:297–305. https://doi.org/10.1044/1092-4388(2001/024.
105. Mazzetto de Menezes KS, Master S, et al. Differences in acoustic and perceptual parameters of the voice between elderly and young women at habitual and high intensity. *Acta Otorrinolaringol (Engl Ed).* 2014;65:76–84. https://doi.org/10.1016/j.otoeng.2013.11.012.
106. Wang L-H, Doan T-N, et al. Prevalence of voice disorders in older adults: a systematic review and meta-analysis. *Am J Speech Lang Pathol.* 2023;32:1758–1769. https://doi.org/10.1044/2023_AJSLP-22-00393.
107. R. Singh J. Keshet et al. The relationship of voice onset time and Voice Offset Time to physical age (Shanghai, China)IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2016 5390 5394.(Shanghai, China).
108. Ryalls J, Cliche A, et al. Voice-onset time in younger and older French-speaking Canadians. *Clin Linguist Phon.* 1997;11:205–212. https://doi.org/10.3109/02699209708985191.
109. Bilodeau-Mercure M, Kirouac V, et al. Movement sequencing in normal aging: speech, oro-facial, and finger movements. *Age.* 2015;37:1–13. https://doi.org/10.1007/s11357-015-9813-x.
110. Kuruvilla-Dugdale M, Dietrich M, et al. An exploratory model of speech intelligibility for healthy aging based on phonatory and articulatory measures. *J Commun Disord.* 2020;87:105995. https://doi.org/10.1016/j.jcomdis.2020.105995.
111. Tremblay P, Deschamps I, et al. *Neuromotor organization of speech production*. Oxford University Press; 2019.
112. National Association of Teachers of Singing Sci-Inf Pedagog Inst 2026.