

Effects of added absorption on the vocal exertions of talkers in a reverberant room

Michael K. Rollins, Timothy W. Leishman, Jennifer K. Whiting, Eric J. Hunter, and Dennis L. Eggett

Citation: *The Journal of the Acoustical Society of America* **145**, 775 (2019); doi: 10.1121/1.5089891

View online: <https://doi.org/10.1121/1.5089891>

View Table of Contents: <https://asa.scitation.org/toc/jas/145/2>

Published by the [Acoustical Society of America](#)

ARTICLES YOU MAY BE INTERESTED IN

[Empirical corrections for predicting the sound insulation of double leaf cavity stud building elements with stiffer studs](#)

The Journal of the Acoustical Society of America **145**, 703 (2019); <https://doi.org/10.1121/1.5089222>

[Experiment and estimation of the sound absorption coefficient for clearance of corrugated honeycomb](#)

The Journal of the Acoustical Society of America **145**, 724 (2019); <https://doi.org/10.1121/1.5089427>

[Lexical frequency effects in English and Spanish word misperceptions](#)

The Journal of the Acoustical Society of America **145**, EL136 (2019); <https://doi.org/10.1121/1.5090196>

[Modeling low-frequency vibration in light-weight timber floor/ceiling systems](#)

The Journal of the Acoustical Society of America **145**, 831 (2019); <https://doi.org/10.1121/1.5087706>

[Morphological characteristics of male and female hypopharynx: A magnetic resonance imaging-based study](#)

The Journal of the Acoustical Society of America **145**, 734 (2019); <https://doi.org/10.1121/1.5089220>

[Error patterns of native and non-native listeners' perception of speech in noise](#)

The Journal of the Acoustical Society of America **145**, EL129 (2019); <https://doi.org/10.1121/1.5087271>

Effects of added absorption on the vocal exertions of talkers in a reverberant room

Michael K. Rollins,^{a)} Timothy W. Leishman, and Jennifer K. Whiting
*Acoustics Research Group, Department of Physics and Astronomy, Brigham Young University,
N283 Eyring Science Center, Provo, Utah 84602, USA*

Eric J. Hunter
*Department of Communicative Sciences and Disorders, Michigan State University, 113 Oyer Speech and
Hearing Building, East Lansing, Michigan 48824, USA*

Dennis L. Eggett
*Department of Statistics, Brigham Young University, 223 Talmage Math Computer Building, Provo,
Utah 84602, USA*

(Received 8 May 2018; revised 13 January 2019; accepted 21 January 2019; published online 12 February 2019)

Occupational speech users such as schoolteachers develop voice disorders at higher rates than the general population. Previous research has suggested that room acoustics may influence these trends. The research reported in this paper utilized varying acoustical conditions in a reverberant room to assess the effects on vocal parameters of healthy talkers. Thirty-two participants were recorded while completing a battery of speech tasks under eight room conditions. Vocal parameters were derived from the recordings and the statistically significant effects of room acoustics were verified using mixed-model analysis of variance tests. Changes in reverberation time (T_{20}), early decay time (EDT), clarity index (C_{50}), speech transmission index (STI), and room gain (G_{RG}) all showed highly correlated effects on certain vocal parameters, including speaking level standard deviation, speaking rate, and the acoustic vocal quality index. As T_{20} , EDT, and G_{RG} increased, and as C_{50} and STI decreased, vocal parameters showed tendencies toward dysphonic phonation. Empirically derived equations are proposed that describe the relationships between select room-acoustic parameters and vocal parameters. This study provides an increased understanding of the impact of room acoustics on voice production, which could assist acousticians in improving room designs to help mitigate unhealthy vocal exertion and, by extension, voice problems.

© 2019 Acoustical Society of America. <https://doi.org/10.1121/1.5089891>

[FM]

Pages: 775–783

I. INTRODUCTION

Past studies have suggested that both adults¹ and children² adjust their voices in accordance with their communication environment. For example, the Lombard effect is the involuntary speech production response to changes in room background noise, where a talker may increase vocal effort to maintain a satisfactory signal-to-noise ratio for communication.³ While such adjustments are normal and reflex-like, they may also strain the voice, leading to vocal fatigue. An example of this in the workplace involves school teachers who often speak with loud voices in noisy environments for long periods of time; they also present with voice problems more often than the average person.⁴ This not only affects teacher health and quality of life,⁵ but increases costs to schools due to missed workdays and the hiring of substitute teachers,⁶ while also decreasing student learning and academic performance.⁷

In a limited sense, previous investigations have probed certain effects of room-acoustic properties (e.g., voice

support, reverberation time, and noise level) on voice production that may connect to vocal health. For example, using 184 individuals in groups of 23 talkers reading set phrases in eight rooms, Black⁸ showed that groups within a room read the phrases with different vocal intensities and rates depending on the rooms, which varied in size and reverberation time. Pelegrín-García and Brunskog⁹ found that for a talker addressing a listener 1.5 m away, increasing the room gain 1 dB decreased voice levels 1.6 dB on average. Puglisi *et al.*¹⁰ showed that teacher sound pressure level (SPL) exhibited both a linear increase with respect to background noise level and a parabolic relationship with classroom reverberation time in the range $0.4 \text{ s} \leq T_{30} \leq 1.4 \text{ s}$, with a minimum SPL at $T_{30} = 0.7 \text{ s}$. Bottalico *et al.*¹¹ suggested a correlation between vocal load and reverberation time. Bottalico¹² later showed in a laboratory experiment involving 20 talkers that speech was highly affected by variations in room reverberation, with more monotonous speech being produced in more reverberant environments. Bosker and Cooke¹³ showed that talkers produce more pronounced amplitude modulations when speaking in noise. While insightful, these studies presented only simple vocal measures such as talker voice level, fundamental frequency, rate, and amount (duration) of speaking in response to one or two

^{a)}Current address: MED-Otolaryngology, MSB 6303, 231 Albert Sabin Way, Cincinnati, OH 45267-0528, USA. Electronic mail: MichaelKentRollins@gmail.com

room-acoustic variables. The effects of these and other useful room-acoustic properties on a broader set of effort-related vocal measures, including voice parameters related to vocal quality and vocal health, have yet to be reported.

While the auditory-feedback manifestations of distinctive room-acoustic properties are likely to play a role in vocal accommodations (i.e., adapting the characteristics of voices to local circumstances) of occupational speech workers,^{3,12,14} more research is needed to understand the specific effects.¹⁵ The impact of classroom and other room acoustics on *listeners* has been thoroughly studied, especially regarding noise levels and speech intelligibility,^{16–18} resulting in listener-focused standards and recommendations.^{19,20} However, investigations regarding acoustic effects of classrooms and other settings on *talkers* are still limited. The present study was conducted to help address this deficit and increase understanding of key relationships between room acoustics and vocal accommodation.

A matter that should be of concern to architectural acousticians is whether published optimal ranges for listener-focused room-acoustic parameters adequately address the vocal health of talkers. With greater understanding of the relationships between room acoustics and vocal effort, they will be better equipped to answer this question and design rooms that help mitigate unhealthy vocal accommodations, and by extension, voice problems. The following sections explore the effects of varying room-acoustic parameters on vocal-effort-related parameters for typical talkers in a reverberant room.

II. METHODS

The study involved 32 subjects, individually recorded as they spoke without sound reinforcement to an interviewer in a 204 m³ reverberation chamber with an average mid-frequency (500 Hz and 1 kHz octave bands) T_{20} of 7.9 s when empty. The chamber was sequentially treated with varying amounts of absorption to alter its acoustical characteristics. For a control condition, each subject also moved to an adjacent laboratory room for a final recording. Vocal accommodations were analyzed for each case by studying changes in acoustically derived vocal parameters across changing room-acoustic parameters. This led to simple empirical relationships between the two sets of parameters.

The subjects were all university students of self-reported gender (16 males and 16 females) and self-reported to have no hearing aids, hearing disorders, speaking impediments, or vocal disorders. All were briefed on the purposes and means of the experiment and gave written consent to full participation. Human subject participation approval for this research was obtained via Brigham Young University's Institutional Review Board for Human Subjects.

A. Acoustic conditions

Seven of eight acoustic conditions were presented in the reverberation chamber, which had a reasonably diffuse field for most conditions²¹ and low ambient noise. Its acoustical characteristics were specifically altered through the addition of absorptive wedges to the floor. Each was cut from 32 kg/m³ open-cell polyether foam rubber with a 94.5 cm overall

depth, a 30.5 × 30.5 cm base, and a profile based roughly on those suggested by Beranek and Sleeper.²² When positioned away from the walls according to ISO 354,²¹ each had an average equivalent absorption area of 1.0 m² over the 500 Hz and 1 kHz octave bands. For any given trial, the room-acoustic condition was determined by the number of wedges in the chamber (0, 2, 4, 8, 16, 24, or 32). These conditions were presented in random sequence for each subject. The wedges were distributed quasi-uniformly about the floor at marked locations to ensure consistency and repeatability for each talker.

Because the chamber had such a long mid-frequency T_{20} when empty, reaching a desired control condition with a T_{20} below 0.8 s would have required an impractical number and arrangement of absorptive wedges. The control condition was consequently presented to each talker in the adjacent laboratory room. It was of similar volume (181 m³) and had typical laboratory furnishings (i.e., workbenches, countertops, cabinets, and overhead shelves with equipment), which collectively contributed to a relatively nondescript acoustical environment. The ambient noise level in the chamber was below NCB 10,²³ and for both rooms it was well below the maximum recommended for classroom settings.^{20,24}

Each of the eight conditions was characterized by five room-acoustic parameters: reverberation time (T_{20}),^{24–28} early decay time (EDT),^{25–27} speech clarity (C_{50}),^{24,26–28} speech transmission index (STI),^{24,28,29} and room gain (G_{RG}).³⁰ Both EDT and G_{RG} provide insights into the acoustic feedback from the participant's own voice, while C_{50} and STI provide insights into the acoustic transmission from the interviewer's voice. It was hypothesized that talkers would adjust their voices in response to feedback from their own voices, as well as the quality of transmitted speech from the interviewer. Hence, as a group, these measures present a coherent picture of acoustic parameters that could influence talkers' voices. Because of its wide use, T_{20} is included to compare the other room-acoustic parameters against.

The T_{20} values for each condition were spatially averaged from integrated impulse response measurements. They involved 12 independent source/receiver combinations as suggested by ISO 3382-2²⁷ and ISO 354²¹ for precision measurements characterizing reverberation in a room as a whole. A movable dodecahedron loudspeaker served as the source, while four precision microphones with random-incidence correctors served as the receivers. Critical distances were calculated based on the resulting measures of total equivalent absorption area.

For all source/receiver combinations, the source and receiver were positioned at least 1 m from any wall and 0.75 m from any stationary diffuser or absorptive wedge, with a height of 120 cm ± 2 cm. One of the 12 source/receiver combinations corresponded to the fixed interviewer and subject positions used consistently throughout the study, with a spacing of 185 cm ± 2 cm. This distance was chosen as a conversational distance that leveraged acoustical effects and articulation loss in the room (see Fig. 1). For most acoustic conditions, the mid-frequency T_{20} measured from this combination was within one standard deviation of the spatially averaged T_{20} . For all others, it was within two standard deviations. The frequency-dependent values of the spatially averaged T_{20} are shown in Fig. 2.

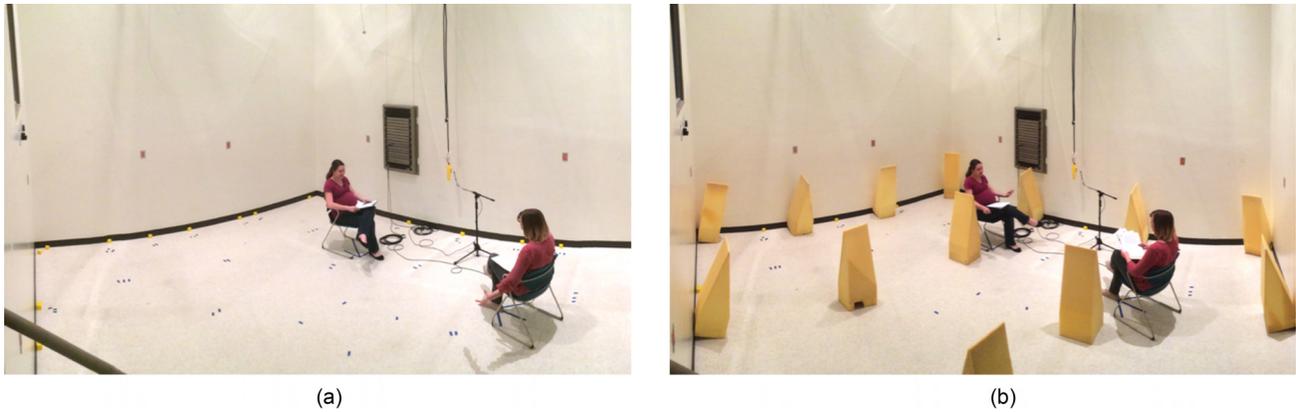


FIG. 1. (Color online) Interviewer (left) and subject (right) in the reverberation chamber with (a) 0 and (b) 16 acoustically absorptive wedges present.

The EDT, C_{50} , and STI depended significantly upon direct, early reflected, and reverberant sound, whereas T_{20} was calculated following the first 5 dB of integrated impulse response decay, meaning the direct and early reflected sound were inherently neglected.^{21,26,27} To better characterize the temporal and spatial dependencies between the interviewer and subject and the impact of speech directivity on the room response, the EDT, C_{50} , and STI were measured with a single source/receiver combination involving a KEMAR head and torso simulator (HATS) at the fixed interviewer position and a microphone at the fixed subject position. The HATS incorporated a mouth simulator, which provided reliable and repeatable measurements germane to the study.³¹ The T_{20} , EDT, C_{50} , and STI values were calculated using EASERA software (Ahnert Feistel Media Group, version 1.2.13).

As suggested by ISO 3382-1 Sec. 8, for large venues with distributed stages and audience seating areas, some room-acoustical measurements may be reported for distinct regions of a room or for a room as a whole.²⁶ For this study, all interviewers and subjects were carefully and consistently located not just within fixed regions of the much smaller rooms, but at carefully fixed locations within the rooms. The distinct placements of the HATS and microphone were then justified on the grounds that the two positions were controlled and constant for each interviewer, subject, and condition,

meaning there was no appreciable spatial distribution. Furthermore, ISO 3382-2²⁷ Sec. 4.2.1 allows sources without specific directivities for engineering and survey measurements, meaning the use of the HATS at the interviewer position was suitable. Finally, since the circumstances for the present study were in some ways dissimilar to those intended by ISO 354,²¹ ISO 3382-1,²⁶ and ISO 3382-2,²⁷ the authors considered the measurements to be unique and thus not fully circumscribed by the standards.

The room gain G_{RG} may be defined as the gain, in decibels, introduced by the reflections from the room boundaries to the voice of the talker at his or her own ears. It was calculated for each acoustic condition using oral-binaural impulse responses (OBRIRs) from the mouth to ears of the HATS at the subject position³² and the formula

$$G_{RG} = L_E - L_D, \quad (1)$$

where L_E is the total level and L_D is the direct level (initial and early diffracted level, without room reflections) of the airborne sound.³⁰

The 500 Hz and 1 kHz octave-band values of T_{20} , EDT, and C_{50} were averaged to generate mid-frequency, single-number values for each condition.^{25,26,28} The results are presented in Fig. 3 as functions of both the number of wedges and the equivalent absorption area A (including air absorption) in the chamber. They are also listed in Table I, along with critical distances that assume unity directivity factors.

As indicated earlier, the average equivalent absorption area per wedge over the 500 Hz and 1 kHz octave bands was 1.0 m^2 when they were positioned away from the room boundaries as required by ISO 354,²¹ with four wedges acting as object specimens. However, for the speech tests, some wedges were placed in the dihedral and trihedral corners of the room (angled inward at a consistent angle) to maintain ample wedge spacing for all configurations and increase low-frequency and total equivalent absorption area as compared to that of closely spaced or clustered wedges. The positioning of some wedges in the room corners resulted in slightly reduced average mid and high-frequency equivalent absorption areas per wedge. The average for several configurations over the 500 Hz and 1 kHz octave bands was approximately 0.8 m^2 per wedge.

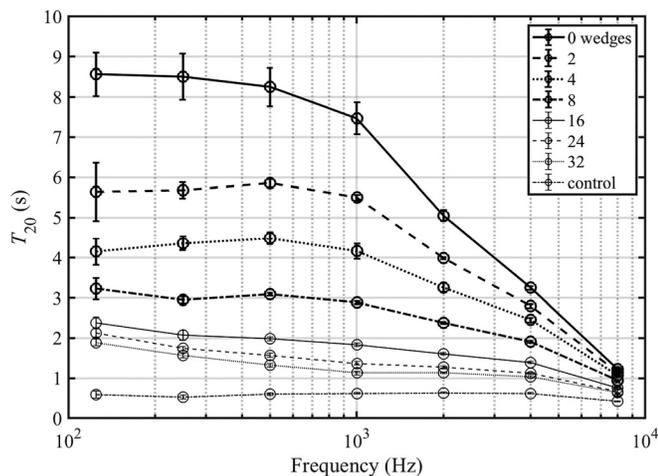


FIG. 2. Spatially averaged T_{20} values for all unoccupied experimental conditions, reported in octave bands.

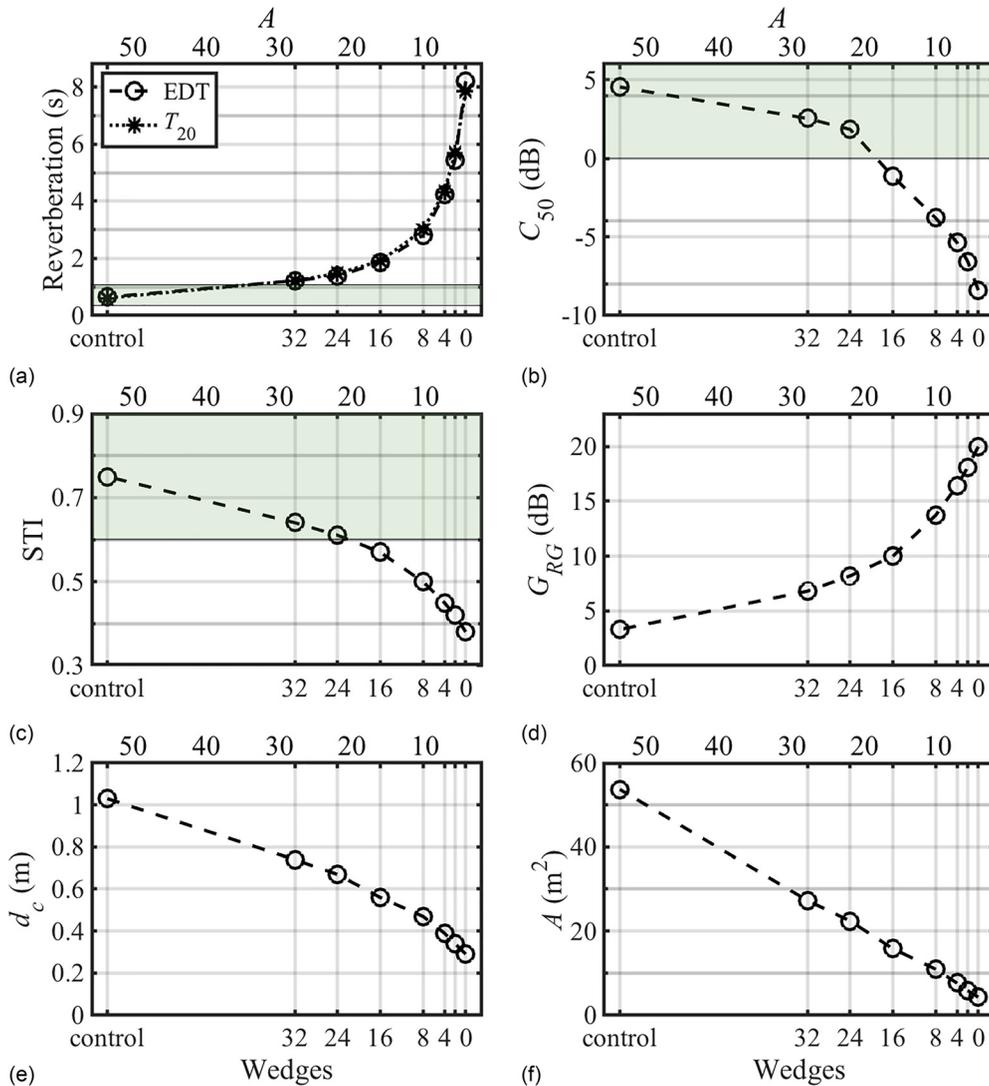


FIG. 3. (Color online) Relationships between room-acoustical parameters (a) EDT, T_{20} ; (b) C_{50} ; (c) STI; (d) G_{RG} ; (e) d_c ; and (f) A ; and the numbers of absorptive wedges (lower abscissa) and total equivalent absorption area A (upper abscissa) in the testing environments. Corresponding values can be found in Table I. The shaded regions represent approximate listener-oriented room parameter recommendations for optimal speech intelligibility (Refs. 24, 28, and 29). [Because Long (Ref. 24) and Ahnert and Tennhardt (Ref. 28) recommended slightly different ranges for the T_{20} , an average of the two was used.] Here and in Fig. 4, the abscissas represent progressively decreasing numbers of wedges and A toward the right, which correspond to increasing reverberation times.

The mid-frequency equivalent absorption area of two arbitrary people was also measured following ISO 354 for absorptive-object specimens and found to be approximately 0.7 m^2 . In principle, this would have a small but nonnegligible impact on total equivalent absorption area and reverberation time in the room, especially with no or few absorptive wedges present. However, because this additional absorption would likely vary with the size, clothing, and hair of each interviewer and subject, we herein report, for consistency, only the acoustical measurement values for unoccupied room conditions.

In Fig. 3, the total equivalent absorption area indicated by the upper abscissa does not depend upon the approximate 0.8 m^2 equivalent absorption area per wedge, but is based instead on actually measured total equivalent absorption areas for the various room conditions. These resulted from the spatially averaged T_{20} values and Sabine's equation as recommended by ISO 354. The total absorption area of the control room was 47.6 m^2 over the mid-frequency bands, but

this was scaled upward by a factor of 1.13 to account for the difference in room volume (181 m^3 vs 204 m^3). This provided a more equitable comparison via Sabine's equation, since the scaled equivalent absorption area in the reverberation chamber would produce the same spatially averaged mid-frequency T_{20} as that measured in the control room.

The various conditions of the reverberation chamber, with its ample volume, highly reflective surfaces, and stationary diffusers, involved reasonably diffuse fields for most configurations and frequencies of interest.^{21,33} Strong correlations between certain acoustic parameters were thus anticipated from a theoretical standpoint and were indeed found from the experimental data (see Table II). Yet because each measure has a distinct purpose for acoustical characterizations, acousticians typically exercise caution in choosing one to the exclusion of others. Moreover, a high average correlation of spatially varying parameters does not always imply a lack of dissimilarities over entire talker-listener regions.^{24,27,34,35}

TABLE I. Room-acoustic parameters for the eight acoustic conditions of the study. The T_{20} , EDT, C_{50} , and equivalent absorption area A values are averages over the 500 Hz and 1 kHz octave bands. The G_{RG} values in the reverberation chamber were originally reported by Whiting (Ref. 32), while the G_{RG} value for the control condition was extrapolated from the same report (indicated by square brackets) using volume-weighted equivalent absorption area of the room. The critical distance d_c is given for a unity directivity factor, as suggested by the T_{20} measurements taken with a dodecahedron loudspeaker at several positions in the rooms. To account approximately for the mid-frequency directivity of the human voice, one may multiply the d_c values by a fixed factor, e.g., 1.4. Plots of these values are presented in Fig. 3.

Wedges	T_{20} (s)	EDT (s)	C_{50} (dB)	STI	G_{RG} (dB)	d_c (m)	A (m ²)
control	0.61	0.65	4.55	0.75	[3.3]	1.03	53.7
32	1.22	1.22	2.55	0.64	6.8	0.74	27.3
24	1.47	1.39	1.85	0.61	8.2	0.67	22.4
16	1.91	1.86	-1.15	0.57	10.0	0.56	15.8
8	2.99	2.81	-3.80	0.50	13.7	0.47	10.9
4	4.33	4.24	-5.35	0.45	16.4	0.39	7.6
2	5.68	5.44	-6.55	0.42	18.1	0.34	5.8
0	7.86	8.20	-8.40	0.38	20.0	0.29	4.2

B. Speech elicitation

For speech evaluation under the various room conditions, each talker was fitted with a head-worn pre-polarized condenser microphone (DPA 4066) positioned 1 cm from the corner of his or her mouth. Its signal was routed to a digital audio interface (PreSonus Firepod) and recorded using Reaper Digital Audio Workstation (version 5) software. The recordings were later analyzed using custom MATLAB code controlling Praat (version 5.4) software.

For each condition, a researcher prompted the participant to perform several speech tasks. These included reading the first paragraph of the Rainbow Passage³⁶ in a conversational fashion, sustaining the low back unrounded vowel /a/ three times for five seconds each, describing a cartoon image from an image inventory, and answering an open-ended prompt (e.g., “Tell me about your favorite city” or “Describe your favorite dessert”) for about 60 s. One iteration of all speech tasks in a single acoustic condition constituted one trial. The presence of the interviewer in the room was intended to promote a conversational mode of communication.

Overall, each participant completed nine trials with the changing conditions. The first, which took place under the control condition, was only used to instruct the participant and was not included in the analysis. The participant and researcher then relocated to the reverberation chamber for the next seven

TABLE II. Coefficients of determination between all room-acoustic parameters and number of wedges in the experimental conditions. The correlations in the first four columns are all greater than 0.8.

	EDT	C_{50}	STI	G_{RG}	d_c	A	Wedges
T_{20}	0.996	0.887	0.853	0.901	0.743	0.582	0.579
EDT	1	0.855	0.818	0.868	0.706	0.545	0.855
C_{50}		1	0.973	0.989	0.976	0.892	0.892
STI			1	0.988	0.921	0.798	0.799
G_{RG}				1	0.936	0.820	0.821
d_c					1	0.966	0.966
A						1	0.999

trials (see Fig. 1) before returning to the control condition again for the final trial. As indicated earlier, the presented order of acoustic conditions in the reverberation chamber was randomized for each participant. The control condition was always last. Between each trial, research assistants entered the chamber to add or remove absorptive wedges, during which time (from 2 to 3 min) the participant’s voice was given a chance to rest.

C. Data exploration

The recordings were segmented into separate .wav files for each condition and task. Two of the tasks, sentences 2 and 3 of the Rainbow Passage and 3 s of the sustained vowel /a/, were concatenated into a single audio file from which Praat, under MATLAB control, extracted several vocal parameters.

These vocal parameters included Lombard-related measures: speech fundamental frequency (F_0), intensity of voiced speech (dB_v), and speaking rate.³⁷ They also included voice parameters which have been shown to have a relationship with vocal quality and vocal health: pitch strength, a measure capturing the salience of pitch presence; harmonics-to-noise ratio (HNR), a component of pitch strength but more widely used; smoothed cepstral peak prominence (CPPs), a quantity highly correlated with dysphonia severity;³⁸ shimmer dB, a measure of local change in amplitude; and acoustic voice quality index (AVQI).³⁹ The latter is a weighted combination of CPPs, HNR, shimmer dB, and three additional parameters: local shimmer, slope, and tilt. The AVQI was proposed by Maryn *et al.*^{38,39} as a measure of voice quality to capture any of several voice dysphonias and has been shown to be clinically feasible as a measure of dysphonia severity.⁴⁰ The AVQI was calculated using a specialized Praat script.³⁸ The pitch strength was determined via a MATLAB implementation of Aud-SWIPE-P.⁴¹

All voice parameters exhibited a normal distribution with the exception of F_0 standard deviation. Therefore, a log-base-10 transform was applied to the F_0 standard deviation values to obtain a normal distribution. Hereafter, this vocal parameter will be referred to in writing as “ F_0 standard deviation,” but in figures it will be indicated in its log-base-10 transform. Each vocal parameter was subjected to a mixed-model analysis of variance (ANOVA) test to evaluate the influences of three independent variables: room-acoustic parameters (T_{20} , EDT, etc.), trial number, and gender. For each independent variable, pair-wise comparisons were then made between all levels within the independent variable. These levels were grouped according to common differences, using a significance threshold of $p < 0.01$. The Tukey-Kramer adjusted p value was used. All statistical analyses were performed using SAS (version 9.4). All two-way interactions of independent variables were found to be statistically insignificant. The effects of gender are not included in this report. Vocal parameters significantly influenced by the room-acoustic parameters are presented in Sec. III.

III. RESULTS

While F_0 mean and dB_v mean were nearly uniform over EDT, Fig. 4 shows that F_0 standard deviation, dB_v standard

deviation, pitch strength mean, speaking rate, and AVQI had more dynamic relationships to EDT (left column). The vertically shaded areas in the figures indicate recommended T_{60} and STI ranges for optimal speech intelligibility in a room with a volume of approximately 200 m³.^{24,28,29} Because the T_{60} ranges recommended by Long²⁴ and Ahnert and Tennhardt²⁸ differ slightly, an average of the two was used in the figures of this report. The F_0 standard deviation, dB_v standard deviation, speaking rate, and pitch strength mean all tended to decrease with increasing reverberation time, while AVQI tended to increase. The right column of Fig. 4 features the same vocal parameters as does the left, but they are plotted against STI to give the reader a better sense of the relationships between the vocal parameters and speech intelligibility. While speaking rate was affected by the room-acoustic parameters, it was also significantly and independently influenced by the trial number.

Based on the quasi-linear relationships between F_0 standard deviation, dB_v standard deviation, pitch strength mean, pitch strength standard deviation, speaking rate, AVQI, and the room-acoustic parameters, a linear fit was found using the ordinary least squares method (Fig. 4). The empirically derived equations and the closeness of their fit are included in Table III. The R^2 goodness-of-fit values in Table III show that, in general, EDT was the best predictor of the listed vocal parameters compared to C_{50} , STI, and G_{RG} .

When treated as categorical levels, the acoustic conditions can be grouped according to common differences. Such a comparison is shown in Table IV, where the wedge number in the left column can be replaced by the value for any other room-acoustic parameter in the corresponding row in Table I. In Table IV, A, B, C, D, and E refer to groups of acoustic conditions such that the difference between any two members of a given group are statistically insignificant for the vocal parameter indicated in the column header. For example, the pitch strength mean was statistically similar for all acoustic conditions indicated in the top four rows, and independently similar for the 8- and 4-wedge conditions; the 4- and 2-wedge conditions; and the 2- and 0-wedge conditions.

IV. DISCUSSION

Most studies of room effects on speech have focused specifically on noise effects (Lombard) with speech intensity (dB) as the primary outcome. However, some have shown that males may also adjust F_0 in conjunction with dB, but it is not the primary effect.^{42,43} For the cases of this study involving more extreme reverberation, which we hypothesized would be somewhat similar to noise, F_0 mean and dB_v mean did not change significantly. While it may not be surprising that F_0 mean did not change much, as it is a minor effect in Lombard, we did expect that voice level (dB_v mean) would decrease with increased reverberation as has been seen for smaller reverberation changes.⁴⁴ In our results, the dB_v mean curve did have some initial trend to decrease (similar to that reported in Ref. 44), but this was followed by a trend to increase with more reverberance.

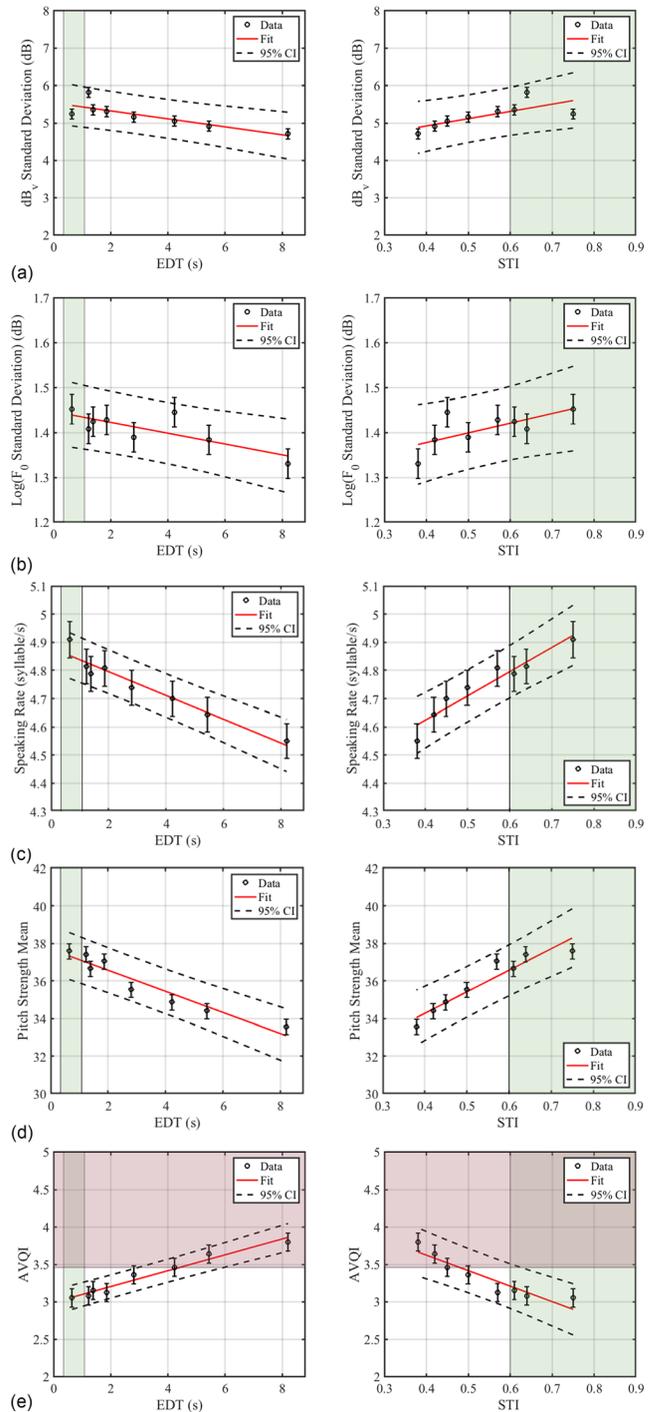


FIG. 4. (Color online) Plots of the population average for (a) dB_v standard deviation, (b) Log(F_0 standard deviation) (indicating diminished prosodic variation), (c) speaking rate, (d) pitch strength mean, and (e) AVQI (indicating increased voice pathology) across the range of EDT (left column) and STI (right column) for the various acoustic conditions. Linear fitted models with upper and lower 95% prediction bounds (observational, nonsimultaneous) are included. Error bars on the data points indicate ± 1 standard error. The vertical shaded zones represent listener-oriented room parameter value recommendations for optimal speech intelligibility (Refs. 24, 28, and 29). Values with error bars that do not overlap have a statistically significant difference. In (e), horizontal shaded zones above AVQI = 3.46 suggest a pathological voice (Ref. 45).

This trend was accompanied by a significant lowering of dB_v standard deviation in response to greater reverberance [Fig. 4(a)]. This, coupled with a similar trend for the F_0 standard deviation [Fig. 4(b)], equates to decreasing inflections

TABLE III. Empirically derived linear-fit equations for influence of room-acoustic parameters on select vocal parameters.

Vocal parameter	Linear fitted models	R^2 of fitted model
$\log_{10}(F_0$ standard deviation)	$\log_{10}(F_{0_{st.d.}}) = -0.012 \cdot EDT + 1.45$	0.62
	$\log_{10}(F_{0_{st.d.}}) = 0.0057 \cdot C_{50} + 1.42$	0.46
	$\log_{10}(F_{0_{st.d.}}) = 0.22 \cdot STI + 1.29$	0.47
dB_v standard deviation)	$\log_{10}(F_{0_{st.d.}}) = -0.0046 \cdot G_{RG} + 1.46$	0.47
	$dB_{v_{st.d.}} = -0.11 \cdot EDT + 5.54$	0.69
	$dB_{v_{st.d.}} = 0.057 \cdot C_{50} + 5.31$	0.65
Speaking rate	$dB_{v_{st.d.}} = 1.94 \cdot STI + 4.15$	0.54
	$dB_{v_{st.d.}} = -0.044 \cdot G_{RG} + 5.73$	0.62
	$SR = -0.042 \cdot EDT + 4.88$	0.94
Pitch strength mean	$SR = 0.023 \cdot C_{50} + 4.79$	0.89
	$SR = 0.86 \cdot STI + 4.28$	0.92
	$SR = -0.018 \cdot G_{RG} + 4.97$	0.93
	$PS_m = -0.56 \cdot EDT + 37.7$	0.92
AVQI	$PS_m = 0.31 \cdot C_{50} + 36.5$	0.92
	$PS_m = 11.5 \cdot STI + 29.7$	0.90
	$PS_m = -0.25 \cdot G_{RG} + 38.8$	0.94
	$AVQI = 0.11 \cdot EDT + 2.99$	0.96
AVQI	$AVQI = -0.056 \cdot C_{50} + 3.22$	0.89
	$AVQI = -2.07 \cdot STI + 4.45$	0.86
	$AVQI = 0.046 \cdot G_{RG} + 2.79$	0.91

in speech and prosodic interest (e.g., talkers using less vocal pitch and loudness variation in their speech) as a response to increasing reverberance and decreasing speech clarity. This also suggests an accommodation that sacrifices speech appeal or interest to increase intelligibility through more monotone production under progressively extreme acoustic conditions. Both of these results correspond with previous work that showed a more monotonous speech style when speaking in more reverberant conditions.¹² Not surprisingly, speaking rate tended to decrease slightly with increasing reverberation time, which may have served the same purpose [Fig. 4(c)].

The pitch strength mean, which at a basic level can be thought of as the inverse of speech “breathiness,” decreased with increasing reverberation (i.e., speech “breathiness” increased with increasing reverberation) [Fig. 4(d)]. This might likewise reflect the tendency toward decreasing expressiveness described earlier. However, overall estimated voice quality clearly tended toward dysphonia, as depicted by the increasing AVQI in response to increasing reverberation time [Fig. 4(e)]. In fact, the two most reverberant conditions

resulted in the total population average falling above a proposed AVQI threshold for pathological speech evaluation, with the third most reverberant condition corresponding to a total population average falling on the said pathological threshold [upper shaded area in Fig. 4(e)].⁴⁵

The groupings in Table IV show two potential room-acoustic thresholds for influencing multiple voice parameters: the boundary between 16 and 8 wedges, and the boundary between 4 and 2 wedges. These correspond to T_{20} ranges of approximately 1.9–3.0 s and 4.3–5.7 s, respectively, and reflect potential thresholds for eliciting a significant change in F_0 standard deviation, dB_v standard deviation, speaking rate, pitch strength mean, and AVQI. These and successive findings could help determine room-acoustic standards that account for vocal health.

Measurements such as these in a relatively diffuse field have certain advantages. In ideal reverberant-field models, the room-acoustic parameters T_{20} , EDT, and C_{50} are related analytically. However, it remains to be shown that STI and G_{RG} are also related analytically. Measurements in a reverberation chamber thus provide a reasonable first step toward increasing understanding of relationships between room-acoustic parameters and vocal parameters. From this viewpoint, it appears that existing listener-oriented recommended values for T_{20} , C_{50} , and STI (corresponding to all shaded areas in Fig. 3 and the vertical shaded areas in Fig. 4)^{24,28,29} are sufficient guidelines for short-term voice use. However, vocal fatigue, vocal loading, noise, and other factors must be treated separately. Relationships for other types of rooms should also be explored.

V. CONCLUSIONS

The results presented in this work demonstrate basic relationships between several room-acoustic parameters and speech parameters. As T_{20} , EDT, and G_{RG} increased, and C_{50} and STI decreased, the speech parameters showed simultaneous tendencies toward diminished prosodic variations (F_0 and dB_v standard deviations), slower speaking rate, and more dysphonic phonation (pitch strength mean and AVQI). Empirical linear relationships between some of these vocal parameters and room-acoustic parameters were set forth. The study further found that formerly published optimal ranges of selected room-acoustic parameters,^{24,28} historically based on listener needs, tend to accommodate talker needs as well. However, additional research is needed to further

TABLE IV. Acoustic condition groups according to common differences. The condition indicated by the number of wedges in the leftmost column corresponds to the values of all other room-acoustic parameters in the corresponding row in Table I.

Wedges	F_0 standard deviation		dB_v standard deviation		Speaking rate			Pitch strength mean		AVQI	
Control	A		A	B	A			A		A	
32	A	B			C	A	B		A		A
24	A		B	C	A	B	C	D	A		A
16	A		B	C	A	B	C		A		A
8	A	B	A	B	A	B	C	D	B		A
4	A	B	A	B		B	C	D	B	C	B
2	A		A	B				D	E	C	C
0		B	A					E		D	D

validate this encouraging observation for various types of rooms.

The research provides insights on how talkers adjust their voice production in a wide range of reverberant conditions, particularly in reducing inflections commonly found in everyday oral communication, while also reducing speech rate. These insights may be useful to help architectural acousticians better understand and communicate about room designs and treatments for unamplified speech. It should also enable them to improve talker-friendly classrooms and other occupational settings to help mitigate unhealthy vocal accommodations due to room design.

Despite the relatively diffuse conditions of the reverberation chamber used for most measurements in the study and the correspondingly strong correlations between T_{20} , EDT, C_{50} , STI, and G_{RG} , the results serve as a basis for future investigations that may not involve such conditions. Similar experiments could be conducted using finer T_{20} increments, e.g., within the 1.9–3.0 s range, to better characterize a threshold at which vocal parameters begin to change considerably. Experiments could likewise be conducted for subjects with documented voice disorders. In addition, gender differences in vocal effort due to room-acoustic conditions could be explored in relationship to differences in gender vocal health. The authors recommend additional work in these areas to better establish ideal room-acoustic conditions for the speaking voice.

ACKNOWLEDGMENTS

The authors would like to thank M. Berardi, K. Smith, S. Markham, and R. Rasband for lending their time and expertise as part of this research effort. This research was funded in part by NIH Award No. R01DC012315. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors have no conflicts of interest to declare.

¹P. Bottalico, S. Graetzer, and E. J. Hunter, “Effects of voice style, noise level, and acoustic feedback on objective and subjective voice evaluations,” *J. Acoust. Soc. Am.* **138**(6), EL498–EL503 (2015).

²E. J. Hunter, “A comparison of a child’s fundamental frequencies in structured elicited vocalizations versus unstructured natural vocalizations: A case study,” *Int. J. Pediatr. Otorhinolaryngol.* **73**(4), 561–571 (2009).

³H. Lane and B. Tranel, “The Lombard sign and the role of hearing in speech,” *J. Speech Lang. Hear. Res.* **14**, 667–709 (1971).

⁴M. Angelillo, G. D. Maio, G. Costa, N. Angelillo, and U. Barillari, “Prevalence of occupational voice disorders in teachers,” *J. Prev. Med. Hyg.* **50**(1), 26–32 (2009).

⁵N. Roy, R. M. Merrill, S. Thibeault, S. D. Gray, and E. M. Smith, “Voice disorders in teachers and the general population: Effects on work performance, attendance, and future career choices,” *J. Speech Lang. Hear. Res.* **47**, 542–551 (2004).

⁶L. C. Cantor Cutiva and A. Burdorf, “Medical costs and productivity costs related to voice symptoms in colombian teachers,” *J. Voice* **29**(6), 776.e15–776.e22 (2015).

⁷D. R. Morton, V. Watson, and D. R. Watson, “The impact of impaired vocal quality on children’s ability to process spoken language,” *Logop. Phoniater. Voco.* **26**(1), 17–25 (2001).

⁸J. W. Black, “The effect of room characteristics upon vocal intensity and rate,” *J. Acoust. Soc. Am.* **22**(2), 174–176 (1950).

⁹D. Pelegrín-García and J. Brunskog, “Speakers’ comfort and voice level variation in classrooms: Laboratory research,” *J. Acoust. Soc. Am.* **132**(1), 249–260 (2012).

¹⁰G. E. Puglisi, A. Astolfi, L. C. Cantor Cutiva, and A. Carullo, “Four-day-follow-up study on the voice monitoring of primary school teachers: Relationships with conversational task and classroom acoustics,” *J. Acoust. Soc. Am.* **141**(1), 441–452 (2017).

¹¹P. Bottalico, A. Astolfi, and E. J. Hunter, “Teachers’ voicing and silence periods during continuous speech in classrooms with different reverberation times,” *J. Acoust. Soc. Am.* **141**(1), EL26–EL31 (2017).

¹²P. Bottalico, “Speech adjustments for room acoustics and their effects on vocal effort,” *J. Voice* **31**(3), 392.e1–392.e12 (2017).

¹³H. R. Bosker and M. Cooke, “Talkers produce more pronounced amplitude modulations when speaking in noise,” *J. Acoust. Soc. Am.* **143**(2), EL121–EL126 (2018).

¹⁴G. Fairbanks, “Selective vocal effects of delayed auditory feedback,” *J. Speech Hear. Disorders* **20**(4), 333–346 (1955).

¹⁵M. Kob, A. Kamproff, C. Neuschaefer-Rube, O. Goldschmidt, and G. Behler, “Experimental investigations of the influence of room acoustics on the teacher’s voice,” *Acoust. Sci. Technol.* **29**(1), 86–94 (2008).

¹⁶B. M. Shield and J. E. Dockrell, “The effects of noise on children at school: A review,” *Build. Acoust.* **10**(2), 97–116 (2003).

¹⁷R. C. Dreossi and T. M. Momensohn-Santos, “Noise and its interference over students in a classroom environment: Literature review,” *Pró-Fono R. Atual. Cient.* **17**(2), 251–258 (2005).

¹⁸D. M. Howard and J. A. S. Angus, “Room acoustics: How they affect vocal production and perception,” in *Occupational Voice: Care and Cure*, edited by P. Dejonckere (Kugler Publications, The Hague, The Netherlands, 2001), pp. 29–46.

¹⁹L. L. Doelle, *Environmental Acoustics* (McGraw-Hill Book Company, New York, 1972).

²⁰ANSI S12.60-2010: *American National Standard Acoustical Performance Criteria, Design Requirements, and Guidelines for Schools, Part 1: Permanent Schools* (American National Standards Institute, New York, 2010).

²¹ISO 354:2003(E): *Acoustics—Measurement of Sound Absorption in a Reverberation Room* (International Organization for Standardization, Geneva, Switzerland, 2003).

²²L. Beranek and H. Sleeper, Jr., “The design and construction of anechoic sound chambers,” *J. Acoust. Soc. Am.* **18**(1), 140–150 (1946).

²³ANSI S12.2-1995, *American National Standard Criteria for Evaluating Room Noise* (American National Standards Institute, New York, 1995).

²⁴M. Long, *Architectural Acoustics* (Elsevier Academic Press, Waltham, MA, 2014), Chaps. 3 and 7.

²⁵L. Beranek, *Concert Halls and Opera Houses: Music, Acoustics, and Architecture* (Springer, New York, 2004), Chap. 4 and Appendix 1.

²⁶ISO 3382-1:2009(E): *Acoustics—Measurement of Room Acoustic Parameters – Part 1: Performance Spaces* (International Organization for Standardization, Geneva, Switzerland, 2009).

²⁷ISO 3382-2:2008(E): *Acoustics—Measurement of Room Acoustic Parameters—Part 2: Reverberation Time in Ordinary Rooms* (International Organization for Standardization, Geneva, Switzerland, 2008).

²⁸W. Ahnert and H.-P. Tennhardt, “Room-acoustical fundamentals for auditoriums and concert halls,” in *Handbook for Sound Engineers*, edited by Glen M. Ballou (Focal Press, Burlington, MA, 2015), Sec. 9.2.

²⁹IEC 60268-16:2011: *Sound System Equipment – Part 16: Objective Rating of Speech Intelligibility by Speech Transmission Index* (International Electrotechnical Commission, for Electrotechnical Standardization, Geneva, Switzerland, 2011).

³⁰D. Pelegrín-García, J. Brunskog, V. Lyberg-Åhlander, and A. Löfqvist, “Measurement and prediction of voice support and room gain in school classrooms,” *J. Acoust. Soc. Am.* **131**(1), 194–204 (2012).

³¹T. Halkosaari, M. Vaalgamaa, and M. Karjalainen, “Directivity of artificial and human speech,” *J. Audio Eng. Soc.* **53**(7/8), 620–631 (2005).

³²J. Whiting, “Development of a real-time auralization system for assessment of vocal effort in virtual-acoustic environments,” Master’s thesis, Brigham Young University, Provo, UT, 2018, Chap. 3.

³³D. Nutter, “Sound absorption and sound power measurements in reverberation chambers using energy density methods,” Master’s thesis, Brigham Young University, Provo, UT, 2006.

³⁴J. S. Bradley, “Review of objective room acoustics measures and future needs,” *Appl. Acoust.* **72**(10), 713–720 (2011).

³⁵S. Rollins, “Acoustics of the Salt Lake Tabernacle: Characterization and study of spatial variation,” Master’s thesis, Brigham Young University, Provo, UT, 2005.

- ³⁶G. Fairbanks, *Voice and Articulation Drillbook* (Harper & Row, New York, 1960).
- ³⁷D. Crystal, *Prosodic Systems and Intonation in English* (Cambridge University Press, Cambridge, UK, 1976), pp. 152–156.
- ³⁸Y. Maryn and D. Weenink, “Objective dysphonia measures in the program Praat: Smoothed cepstral peak prominence and acoustic voice quality index,” *J. Voice* **29**(1), 35–43 (2015).
- ³⁹Y. Maryn, P. Corthals, P. Van Cauwenberge, N. Roy, and M. De Bodt, “Toward improved ecological validity in the acoustic measurement of overall voice quality: Combining continuous speech and sustained vowels,” *J. Voice* **24**(5), 540–555 (2010).
- ⁴⁰F. Núñez-Batalla, E. Díaz-Fresno, A. Álvarez-Fernández, G. M. Cordero, and J. L. L. Pendás, “Application of the acoustic voice quality index for objective measurement of dysphonia severity,” *Acta Otorrinolaringol.* **68**(4), 204–211 (2017).
- ⁴¹A. Camacho and J. G. Harris, “A sawtooth waveform inspired pitch estimator for speech and music,” *J. Acoust. Soc. Am.* **124**(3), 1638–1652 (2008).
- ⁴²A. Mcallister, E. Sederholm, and J. Sundberg, “Perceptual and acoustic analysis of vocal registers in 10-year-old children,” *Logop. Phoniatr. Voco.* **25**(2), 63–71 (2000).
- ⁴³T. Letowski, T. Frank, and J. Caravella, “Acoustical properties of speech produced in noise presented through supra-aural earphones,” *Ear Hear.* **14**(5), 332–338 (1993).
- ⁴⁴P. Bottalico, S. Graetzer, and E. J. Hunter, “Effects of speech style, room acoustics, and vocal fatigue on vocal effort,” *J. Acoust. Soc. Am.* **139**(5), 2870–2879 (2016).
- ⁴⁵V. Reynolds, A. Buckland, J. Bailey, J. Lipscombe, E. Nathan, S. Vijayasekaran, R. Kelly, Y. Maryn, and N. French, “Objective assessment of pediatric voice disorders with the acoustic voice quality index,” *J. Voice.* **26**(5), 627.e1–627.e7 (2012).