JASA ARTICLE

Singing in different performance spaces: The effect of room acoustics on vibrato and pitch inaccuracy

Pasquale Bottalico,[1,a)] Natalia Łastowiecka,[1,b)] Joshua D. Glasner,[2,c)] and Yvonne Gonzales Redman[3]
[1]*Department of Speech and Hearing Science, University of Illinois at Urbana-Champaign, Champaign, Illinois 61820, USA*
[2]*Department of Visual and Performing Arts, Clarke University, Dubuque, Iowa 52001, USA*
[3]*School of Music, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*

ABSTRACT:
Previous literature suggests that musical performers may be influenced to some extent by the acoustic environment in which they sing or play. This study investigates the influence of room acoustics on singers' voice production, by analyzing consecutive sung performances of classically trained students in five different performance spaces. The analyzed voice parameters were vibrato rate, extent, and pitch inaccuracy. Nine classically trained student-singers performed the same aria unaccompanied on a variable starting pitch that was consistent between spaces. Variance in vibrato rate and pitch inaccuracy was primarily explained by individual differences between singers. Conversely, the variance attributable to the rooms for the parameter of vibrato extent was larger compared to the variance attributable to the performers. Vibrato extent tended to increase with room clarity (C80) and was inversely associated with early decay time (EDT). Additionally, pitch inaccuracy showed a significant negative association with room support (ST_v). Singers seem to adjust their vocal production when performing in different acoustic environments. Likewise, the degree to which a singer can hear themself on stage may influence pitch accuracy.
 https://doi.org/10.1121/10.0011675

(Received 7 March 2022; revised 15 May 2022; accepted 25 May 2022; published online 22 June 2022)

[Editor: Francesco Martellotta] Pages: 4131–4139

I. INTRODUCTION

Singers tend to report that the acoustic environment in which they sing alters the way they perform a musical work (Jordan, 1980) regardless of how that repertoire was prepared. Indeed, spatial environments do modify sound and typically performance spaces are acoustically designed for the benefit of the listening audience (Beranek, 2004). Performing spaces can influence a singer's perception and affect their musical delivery; adding variability and possibly requiring adaptation during a performance (Ternström, 1993; Luizard and Bernardoni, 2020). Anecdotally, vocalists are sometimes encouraged by their instructors to perform in different spaces to familiarize themselves with different acoustic environments.

Past literature has explored the influence acoustic performance spaces may have upon the musicians playing within different environments (Amengual Gari *et al.*, 2019; Schärer Kalkandjiev and Weinzierl, 2015; Schärer Kalkandjiev and Weinzierl, 2013; Kato *et al.*, 2015; Bolzinger *et al.*, 1994). Investigations have primarily studied instrumental musicians performing in virtual sound environments that are auralized in real-time. This process allows systematic study of the effect of acoustic spaces upon musicians and a solution to the challenges in traveling and reserving access to multiple venues. These types of studies suggest that performers adjust tempo, articulation, and timbre as a result of the aural feedback they receive during a performance.

In a study of pianists, participants decreased their dynamic level and articulation when reverberation time increased (Bolzinger *et al.*, 1994). Likewise, trumpet players tended to lower dynamic level, decrease tempo, and darken their timbre (i.e., highlight lower frequency spectral energy) in more reverberant environments (Amengual Gari *et al.*, 2019). In a study including twelve instrumentalists, researchers determined that 8% of the variance in their musical interpretations, which included tempo and timbre, could be explained by room acoustic parameters (Schärer Kalkandjiev and Weinzierl, 2015). An interesting point, the authors of this study acknowledge, was that the physical space in which musicians performed was literally and experientially different than the virtual acoustic environments they were perceiving. Their subjects expressed putting "effort" into visualizing the spaces they were aurally experiencing. As such, the resulting perceptual disconnect may have influenced participants' instinctive responses. Kato *et al.* (2015) measured one operatic baritone in a sound field simulation system using five different room acoustic conditions; ranging from an anechoic space to a church-like space with high reverberation. The authors noted significant changes to tempo, loudness, pitch accuracy, and vibrato extent as a result of room conditions.

a)Electronic mail: pb81@illinois.edu
b)Also at: School of Music, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.
c)Also at: School of Graduate and Professional Studies, Delaware Valley University, Doylestown, PA 18901, USA.

Studies in live performance halls also suggest that performers alter their performance due to aural feedback. A field study with a renowned cellist performing the same concert in seven different concert halls showed that more than half (58%) of the variance in measured performance characteristics could be explained by room acoustic parameters (Schärer Kalkandjiev and Weinzierl, 2013). More closely related to singers, Ternström (1993) found that different choirs adapt their vocalism in more absorbent rooms by raising their loudness levels and formant peak frequencies.

Luizard and Bernardoni (2020) also studied the effect of acoustic conditions on four singer performances. They investigated the extent of variation within six different empty rooms representing eight acoustic conditions; two were equipped with banners to change reverberation times. Singers of each voice type (soprano, mezzo, tenor, bass) performed three classical pieces without accompaniment in their own repertoire and then all performed one common piece. There was no pattern among these four singers together; however, there were significant individual adaptations made to the changing acoustics in temporal, dynamic, and/or timbral features. These mostly included changes in loudness and timbre, which Ternström (1993) also observed. Luizard and Bernardoni (2020) further recommended that this experiment should be replicated with singers performing consecutively in different acoustic environments to lower the variability in singing due to factors other than acoustics. These include the time between recordings and changes in psychological and physical state. Scheduling and accessing several halls for use in one day is difficult to reproduce; however, on the campus of the University of Illinois Urbana-Champaign (UIUC), there are five professional performance spaces that make this approach possible.

The aim of this present study is to explore variations in vibrato rate, vibrato extent, and pitch inaccuracy that singers make in response to performing consecutively in five different acoustic environments. Measurements are taken from audio recordings made in each performance space.

Early research on the topic of vocal vibrato provided the acoustic description of vibrato that centers on the "periodic pulsation, generally involving pitch, intensity, and timbre, which produces a pleasing flexibility, tenderness, and richness to tone" (Seashore, 1931). There are two common parameters to objectively describe vocal vibrato in modern singing voice science literature: vibrato rate and vibrato extent. Vibrato rate specifies the number of undulations per second and vibrato extent describes how far phonation frequency rises and falls around a mean value during a vibrato cycle (Sundberg, 1987). Typical values for vibrato rate in singers are 4.5 to 6.5 Hz in frequency. Vibrato extent is measured either as a percent or in cents and typical values are around 0 to $\pm 3\%$ or ± 50 cents, a quarter-tone, above and below the mean frequency (Titze, 1994). It is important to note that commonly used applications and algorithms have been reported to measure vibrato extent differently (Nestorova and Glasner, 2021) and that recent studies have found that modern professional opera singers can exhibit vibrato extent in excess of ± 100 cents (Ferrante, 2011; Glasner and Johnson, 2020).

Vocal intonation, also known as pitch accuracy, is defined as the precision of a singer's fundamental frequency (f_0) to a reference frequency that is maintained by self-monitoring and by corrections in the laryngeal muscles (Bottalico *et al.*, 2017). A singer must often sing in different environments in which acoustic feedback used in self-monitoring will differ. Thus, singers learn to rely not only on external feedback but also on internal feedback such as skeletal vibrations (Kleber *et al.*, 2017; Scotto Di Carlo, 1994). However, even trained singers show decreases in pitch accuracy in the absence of external feedback (Ward and Burns, 1978). Such auditory feedback has been shown to be monitored by the audio-vocal system (Hain *et al.*, 2001) and, more recently, has been shown to affect vibrato characteristics in classically trained singers (Lester-Smith *et al.*, 2022).

II. EXPERIMENTAL METHOD

A. Subjects

The use of human subjects for this research was approved by the Office for the Protection of Research Subjects at the University of Illinois Urbana-Champaign (IRB #18179). Six female and three male singers (average age 25.2 years) volunteered to take part in the experiment. The age, gender, and voice type of the nine participants are reported in Table I. The singers were predominantly graduate students in Western classical singing, with an average number of consistent classical singing lessons equal to 9.2 years.

B. Room descriptions

Participants were asked to sing in five different performing venues on the University of Illinois at Urbana-Champaign campus: the Smith Memorial Room, Smith Recital Hall, Colwell Playhouse, Krannert Foellinger Great Hall, and the Krannert Amphitheater. To describe the rooms, Impulse Responses (IRs) were recorded using an Impulsive Sound Source (model BAS006, Larson Davis, Provo, Utah) and analyzed with XL2 Audio and Acoustic Analyzer (NTi Audio, Tigard, Oregon). The microphone used was the

TABLE I. Characteristics of the sample, with age, gender, voice type, and number of years of experience.

Subject ID	Age	Gender	Voice : type	Years of experience
1	26	Female	Mezzo-soprano	11
2	21	Female	Soprano	5
3	24	Male	Tenor	18
4	26	Female	Soprano	12
5	30	Female	Soprano	16
6	24	Female	Soprano	8
7	24	Male	Baritone	10
8	29	Male	Bass	10
9	23	Female	Mezzo-soprano	7

https://doi.org/10.1121/10.0011675

M2211 (NTi Audio, Tigard, Oregon). The source was located in the center of the stage and the receiver was located in different positions in the audience. From the IRs reverberation time (T30), early decay time (EDT), and clarity (C80) were calculated with AURORA, a plug-in for AUDACITY (Farina, 2000). The parameters were averaged over the 500 Hz and the 1 kHz octave bands, following the indication of the standard ISO 3382–1 (ISO, 2008). Volume, number of seats, reverberation time (T30), early decay time (EDT), and clarity (C80) of the halls are listed in Table II. The table includes the average, the minimum, and the maximum values measured in different audience locations. The Smith Memorial Room and Smith Recital Hall are both located within the Tina Weedon Smith Memorial Hall, a historic building completed in 1920. The Smith Memorial Room is modeled after a drawing room with marble floors, plaster walls, crystal chandeliers, and a large rug covering almost the entirety of the floor. It seats 56 people and has a volume of about 400 m^3. It is used primarily for chamber music concerts and recordings. Its reverberation time (T30) is 1.04 s. T30 in this room is lower compared to the preferred values of T30 for chamber music halls (between 1.6 and 1.8 s) suggested by Beranek (2004). Also, the average EDT (0.98 s) is lower compared to the preferred values of EDT for chamber music halls (between 2.2 and 2.6 s). The C80 instead is within the preferred values (−2.0 to 2.0 dB). The Smith Recital Hall is a large concert venue seating 802 guests on the main floor and balcony. The interior is mainly mahogany wood and has an approximate volume of 6600 m^3. The T30 for this space is 1.75 s. T30 and EDT in this room are lower compared to the preferred values for concert halls (between 1.8 and 2.1 s for T30 and 2.2 and 2.6 s for EDT). The C80 instead is within the preferred values (−3.0 to 0 dB).

The remaining spaces are housed in the Krannert Center for Performing Arts, a 28 000 m^2 performing arts and academic facility home to five indoor stages and an outdoor amphitheater. The Colwell Playhouse serves as a venue for small-scale dance and spoken word performances. It seats 641 people and has an approximate volume of 11 600 m^3. It is occasionally used for music performances and is typically judged by the performers as a rather dry space. Its T30 is 1.42 s. Even if this room is designed for spoken word performances, T30, EDT, and C80 are within the range of preferred values for opera houses (between 1.4 and 1.6 s for T30; 1.5 and 1.9 s for EDT; and 1.0 and 3.0 dB for C80). The Krannert Foellinger Great Hall has a grand stage that is designed for large orchestral performances. The hall can seat 2059 people and has a volume of 16 500 m^3 and a T30 of 2.47 s. T30 in this room is higher compared to the preferred values for concert halls (between 1.8 and 2.1 s). The EDT and the C80 instead are within the preferred values (2.2 and 2.6 s for EDT; –3.0 to 0 dB for C80). The Krannert Amphitheater is a Greek-inspired outdoor performance venue used during temperate weather conditions and up to 560 people can sit on the curved concrete steps. In this space, the average T30 is 0.68 s, the average EDT is 0.22 s, and the average C80 is 18.30 dB. These values are similar to the ones measured in ancient open-air theaters such as the Theater of Tyndaris in Sicily, Italy (Astolfi *et al.*, 2020).

Another set of IRs was recorded through a Head and Torso Simulator (HATS, GRAS 45BB KEMAR, Holte, Denmark), to represent the acoustics experienced by the singers. An exponential sweep signal was emitted by the mouth and recorded by the ears of the HATS. The recorded sweep was deconvolved with the emitted sweep inverted on the time axes, obtaining the IR, as explained by Brunskog *et al.* (2009) and Pelegrín-García (2011). The impulse responses mouth-ears (IR_{ME}), early decay time (EDT) corresponding to the perceptual dimension of the reverberance, C80 corresponding to the music clarity, $IACC_{late}$ corresponding to the envelopment were analyzed with AURORA, a plug-in for AUDACITY (Farina, 2000). Voice support (ST_v), corresponding to the degree of amplification produced by the room on the talker's voice, as perceived by the talker himself, was calculated with MATLAB R2018a following the indication of Pelegrín-García (2011). The parameters EDT and C80 were averaged between 500 and 1000 Hz octave bands and between the two ears. $IACC_{late}$ and ST_v were averaged over the 125 Hz to 4 kHz octave bands. The average between the two ears is reported for ST_v. Figure 1 shows the five rooms. The parameters obtained from the IR_{ME} are reported in Table III.

C. Protocol

The experiment was conducted in one recording session within four performance halls and one outdoor amphitheater. Recordings were made in five locations consecutively beginning in Smith Memorial Room, followed by Foellinger Great Hall, Colwell Playhouse, Krannert Amphitheatre, and

TABLE II. Reverberation time (T30), early decay time (EDT), and clarity (C80) measured with an omnidirectional source located on the stage and the receiver in on the audience in the five spaces. All parameters were averaged over the 500 Hz and the 1 kHz octave bands. The table lists the average, the minimum, and the maximum values measured in the audience. The volume and the seat number are also included.

Room name	Volume (m^3)	Seats	Measurement point No.	T30 (s)	C80 (dB)	EDT (s)
Smith Memorial Room	400	56	9	1.04 (1.07–1.11)	2.00 (0.65–3.45)	0.98 (1.09–1.18)
Smith Recital Hall	6600	802	27	1.75 (1.56–1.87)	−0.83 (−2.65–2.87)	1.85 (1.72–2.15)
Krannert Great Hall	16 500	2059	21	2.47 (2.01–2.65)	−0.29 (−2.67–8.45)	2.18 (1.55–2.53)
Krannert Colwell Playhouse	11 600	641	18	1.42 (1.34–1.48)	2.48 (0.50–7.25)	1.43 (1.18–1.71)
Krannert Amphitheatre	—	560	9	0.68 (0.41–0.99)	18.30 (12.00–22.07)	0.22 (0.06–0.52)

https://doi.org/10.1121/10.0011675

FIG. 1. (Color online) The five halls analyzed in the study: (a) Smith Recital Hall, (b) Smith Recital Hall, (c),(d) Krannert Great Hall, (e) Krannert Amphitheatre, and (f) Krannert Colwell Playhouse.

concluding with Smith Recital Hall. The recordings were performed using an omnidirectional measurement microphone (M2211 NTI audio) connected to a portable recorder (TASCAM DR-40X portable recorder) with a sampling rate of 44.1 kHz at 32 bits.

In all of the recordings, the performance spaces were unoccupied, and their environments were unchanged. Singers entered and stood in the same location which was marked for them in the center of each stage. Each singer performed the A and B sections of the Italian art song, "Caro mio ben" by Tommaso Giordani once in each space and without musical accompaniment. The starting pitch was chosen according to the voice type and was maintained in all spaces. The starting pitches chosen were F_5 for the sopranos, Eb_5 for the mezzos, the tenor selected Eb_4, the baritone D_4, and the bass C_4. They were not directed to follow a tempo.

TABLE III. Early decay time (EDT), C80, late interaural cross-correlation (IACClate), and voice support (ST_v) measured on the state with a HATS (IR_{ME}) in the five space. All parameters were averaged over the 500 Hz and the 1 kHz octave bands, except for IACClate and ST_v, which were averaged over the 125 Hz to 4 kHz octave bands. All the parameters, except for IACClate, were averaged between the two ears.

Room name	C80 (dB)	EDT (s)	IACC_late	STv (dB)
Smith Memorial Room	16.16	0.40	0.499	−10.54
Smith Recital Hall	23.53	0.35	0.510	−13.28
Krannert Great Hall	26.82	0.35	0.487	−13.00
Krannert Colwell Playhouse	25.84	0.34	0.487	−4.97
Krannert Amphitheatre	22.82	0.38	0.499	−11.04

D. Voice Analysis

MATLAB version 2014 b and PRAAT version 5.4.01 were used for voice analysis. Three parameters were analyzed: vibrato rate, vibrato extent, and pitch inaccuracy. Each audio recording was manually segmented to isolate the notes/syllables highlighted in Fig. 2.

Vibrato primarily consists of a periodic f_0 modulation (Dejonckere *et al.*, 1995). As mentioned, traditionally, the main parameters used to characterize the vibrato are the rate and the extent. Vibrato rate (V_{rate}) represents the number of f_0 oscillations per second. It is evaluated as the reciprocal of the mean time difference between two subsequent f_0 maxima,

$$V_{rate} = \frac{1}{N}\sum_{i=1}^{N-1}\left|\frac{1}{t^{i+1}_{f_{0max}} - t^{i}_{f_{0max}}}\right|, \tag{1}$$

where i is the cycle identifier, N is the f_0 maxima identifier, and $t^i_{f_{0max}}$ is the time instant of the i th cycle of f_0 maximum. It is important to define limits of pulsation rate that may be considered as vibrato. Relying on the observations of Ekholm *et al.* (1998) and Ferrante (2011), the extreme range may be defined as 4.2–8.1 Hz in females and 4.8–6.6 Hz in males (mean ± 2 standard deviations [SD]). Following the indications of Manfredi *et al.* (2015), the range 4.2–8.1 Hz was considered for both genders.

Vibrato extent (V_{ext}) is the difference in frequency between a maximum and a minimum of f_0 within a cycle (i.e., the amplitude of the cycle). Here, it has been computed as the mean of the differences between f_0 maximum and f_0 minimum in each cycle,

https://doi.org/10.1121/10.0011675

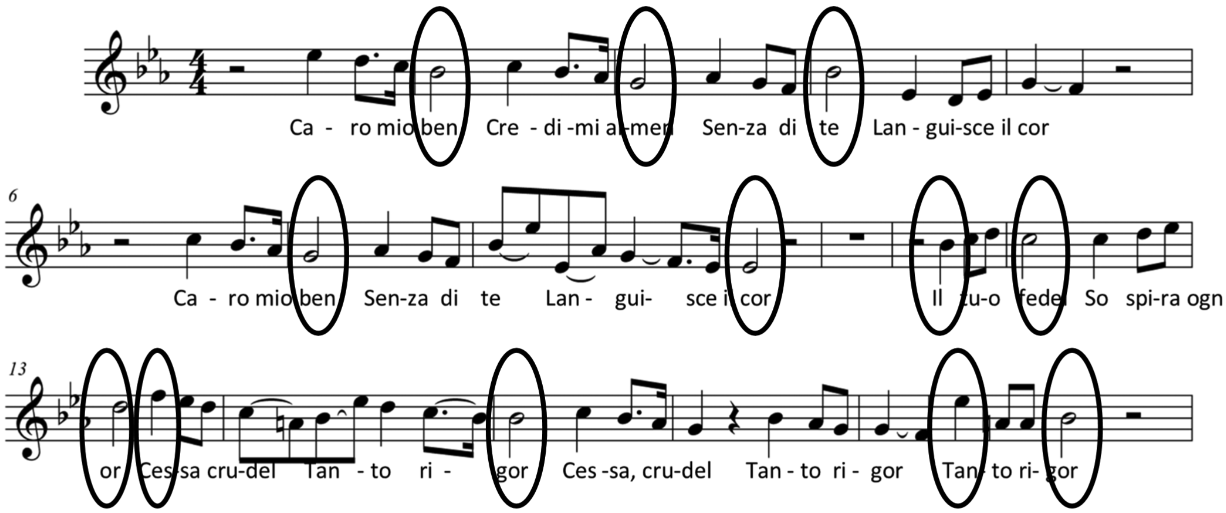

FIG. 2. Score of the aria performed by the participants. The highlighted notes were analyzed for voice parameters.

$$V_{ext} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{f^i_{0_{max}} - f^i_{0_{min}}}{2}\right), \quad (2)$$

where $f^i_{0_{max}} - f^i_{0_{min}}$ is the difference between the maximum and minimum of fundamental frequency on each cycle. To compare results among participants, the measurements in Hz were converted in cents, by means of

$$C_{ent(i)} = 1200 * 3.322 * \left(\frac{f^i_{0_{max}}}{f^i_{0_{min}}}\right). \quad (3)$$

In the same way as for rate, limits need to be defined for extent values that may be considered compatible with vibrato. According to reports by Ferrante (2011) and Anand *et al.* (2012), the extremes should be ±17–145 cents for females and ±38–130 cents for males. For vibrato extent, the relevant limit is the lower one. Once again, following the indications of Manfredi *et al.* (2015), the threshold of ±17 cents was taken for all singers. Henceforth, V_{ext} is computed excluding those cycles that exhibit an extent less than ±17 cents.

Pitch inaccuracy was evaluated by extracting the central portion of each note to exclude voice attack and release effects. Fundamental frequency was estimated by means of the autocorrelation method in PRAAT, using Hanning windows with a temporal length of 3 divided by the value of the pitch floor, with pitch limits of ±123 cents from the reference value, with a 0.05-time step, an octave cost of 0.0025 per octave, and a voiced/unvoiced cost of 0.20. All other parameters had default values. Following Ward and Burns (1978), the distance (absolute value of the difference) in cents between the produced note and the reference note is given by

$$\Delta c = \left|1200 \cdot \log_2(fo/f_{ref})\right|, \quad (4)$$

where f_0 is the produced note and f_{ref} is the reference note in Hz. Reference notes were established for both equal temperament and pure and just intonation; however, the results were not statistically different. To compare the present results with those of previous studies, results are reported for the equal temperament.

E. Statistical analysis

Linear mixed models (LMM) and generalized linear mixed models (GLMM) were used for the statistical analyses, using the software R3.6.0 and the lme4 (version 1.1–10) package (Bates *et al.*, 2014). Different models were built for the three response variables (vibrato rate, vibrato extent, and pitch inaccuracy). Vibrato rate and extent followed a normal distribution and linear mixed models were used to fit these two response variables. In agreement with analytical methods used for generic accuracy data, the distribution of the response variable (the distance in cents between the note produced by the singer and the reference note) was most consistent with the Gamma distribution, thus a GLMM with a binomial distribution (Laplace approximation) was used for this response variable. With eight singers singing in five spaces, the data are structured in different levels of variance. Mixed-effects models are particularly suited to analyze a hierarchical data structure (Hox *et al.*, 2017). The models were computed for each of the three voice parameters, where both fixed effects (the room acoustical parameters) and random effects (the difference among signers) could be simultaneously taken into account.

III. RESULTS

First, the effect of sex on the three parameters measured was assessed. Three mixed effect models were run with vibrato rate, vibrato extent, and pitch inaccuracy as response variables and sex as a predictor. Room and singer id were considered as random factors. In all the three models, the differences introduced by the singer sex were not statistically significant, for this reason, sex was not included as a factor in the following analyses.

TABLE IV. Means and standard errors of the three parameters vibrato rate (V_{rate}), vibrato extent (V_{ext}), and pitch inaccuracy (Δc) by sex and by rooms.

	V_{rate} (Hz)		V_{ext} (cents)		Δc (cents)	
	Mean	Std. error	Mean	Std. error	Mean	Std. error
Sex						
Females	6.1	0.03	87.9	1.5	32.1	1.4
Males	5.7	0.03	87.6	1.9	33.7	1.9
All	5.9	0.02	87.8	1.2	32.8	1.2
Rooms						
Smith Memorial Room	6.06	0.05	58.4	2.1	30.1	2.4
Smith Recital Hall	5.89	0.05	84.9	2.4	32.6	2.6
Krannert Great Hall	5.91	0.05	98.7	2.4	33.7	2.6
Krannert Colwell Playhouse	5.90	0.05	96.2	2.3	27.1	2.2
Krannert Amphitheatre	5.89	0.05	100.8	2.2	40.5	3.0

Intraclass correlation coefficients (ICCs) based on random intercept only models were computed to estimate the relative contribution of respective random effects in the explanation of the overall variance in the data. Regarding the vibrato rate (V_{rate}) and pitch inaccuracy (Δc), the variance attributable to the rooms ($ICC_{Room} = 2\%$ and 2%, respectively) was smaller in contrast to the variance attributable to the performers ($ICC_{Singer} = 50\%$ and 43%, respectively). For the parameter vibrato extent (V_{ext}), the variance attributable to the rooms ($ICC_{Room} = 33\%$) was bigger compared to the variance attributable to the performers ($ICC_{Singer} = 23\%$). The mean values and the standard error of the three parameters in the different rooms are reported in Table IV.

The associations between the parameters, both voice (V_{rate}, V_{ext}, and Δc) and room acoustics (C80, EDT, $IACC_{late}$, and ST_{voice}), were assessed through the univariate linear mixed effect models. Participant id was used as a random factor. Table V lists the estimate and the standard error of each model. Bonferroni adjusted alpha levels of 0.0125 per test were used to control for type I error (0.05/4 acoustical parameters). The following associations were statistically significant. The vibrato rate was negatively associated with the C_{80} while it was positively associated with EDT. The vibrato extent was positively associated with the C_{80} while it was negatively associated with EDT and $IACC_{late}$. The pitch inaccuracy was negatively associated with ST_{voice}.

IV. DISCUSSION

The present study aims to investigate the effects of the room acoustic environment—as described by specific acoustic parameters—on the performance and voice production of Western Classical singers. It analyzes the performances of nine singers in five different acoustic environments. The measurements of those acoustic environments are detailed in Table III. Overall, the results of this study indicate that variance in performance characteristics can be attributed to both the individual singer and the room in which a singer performs, depending upon the acoustic variable. This overall finding partially confirms the observations of Luizard and Bernardoni (2020), but the inclusion of vibrato parameters in this study demonstrates that vibrato characteristics specifically are altered—though likely not volitionally—when a singer performs in a different room. This finding seems to indicate that singers may adjust their vocal production when confronted with differing acoustic environments.

A. Effect of room and individual singers

Most directly, this study followed in part the design of Luizard and Bernardoni (2020); however, there were two key differences that distinguish how these results should be interpreted. First, the participants in this project sang an identical excerpt. This decision allows the results of this study to be interpreted without consideration of different singing styles, vocal ranges, or other differences that arise as a result of varying pitch, dynamic, and rhythmic content in multiple song excerpts. Second, the current study specifically focused on vibrato characteristics as indicators of vocal performance and, indirectly, of vocal function. As opposed to Luizard and Bernardoni (2020), who observed changes in closed quotient (CQ) measurements of singers—which has been described as a correlate to vocal effort (Huang *et al.*, 1995)—the authors chose to study vibrato extent and vibrato rate. These two parameters have been shown in singing voice research literature and suggested in historical voice pedagogy texts to be indicators of vocal function (Prame, 1994; Hakes *et al.*, 1987) and even technical "flaws" (Nix *et al.*, 2016; McKinney, 2005). As such, it seems reasonable to conclude—based on the results of the analysis of the intraclass correlation coefficient—that the room acoustic environment does alter in some way a singer's vocal production. Specifically, the ICC_{Room} with room as a random effect and vibrato extent as a dependent

TABLE V. Associations between the three voice parameters (V_{rate}, V_{ext}, and pitch inaccuracy), and the room acoustics parameters from the oral-binaural impulse response (C80, EDT, $IACC_{late}$, and ST_v). Bonferroni adjusted alpha levels of 0.0125 per test were used (0.05/4) to control for type I error. The bold values are statistically significant assuming a p-value lower than 0.0125.

	C80		EDT		$IACC_{Late}$		ST_v	
Parameters	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
V_{rate}	**−0.016**	**0.004**	**2.32**	**0.68**	0.178	1.80	0.001	0.005
V_{ext}	**3.77**	**0.24**	**−467**	**43**	**−689**	**120**	0.37	0.35
Pitch inaccuracy	−0.021	0.311	66.86	50.88	149.91	131.25	**−0.88**	**0.37**

https://doi.org/10.1121/10.0011675

variable indicated that a higher amount of the variance could be explained by the singers performing in a different room ($ICC_{Room} = 33\%$) than by individual differences between singers ($ICC_{Singer} = 23\%$). This finding is in contrast with those related to vibrato rate and pitch inaccuracy, for which variance in the models was primarily explained by individual differences between singers.

The analysis of ICC reveals how a singer's performance might be impacted or otherwise altered by the room acoustics environment, but it does not clearly reveal why a singer might involuntarily or voluntarily change their performance or vocal function. As such, it is necessary to consider the relationship between the dependent variables of vibrato rate (V_{rate}), vibrato extent (V_{ext}), and pitch inaccuracy (Δc) as well as the independent variables of early decay time (EDT), late interaural cross correlation coefficient ($IACC_{late}$), clarity (C_{80}), and voice support (ST_V) (see Table IV). Note that for the purposes of this study, only comparisons between vocal variables and room acoustic variables are relevant to the present discussion.

B. Vibrato extent and room acoustics

Of the significant associations listed in Table V, V_{ext} was found to have a positive association with C80. That is to say, both the clarity of the room environment at the position of the singer and the amplitude modulation of the singer's vibrato tended to increase together. This phenomenon bears exploration. Given that C_{80} was measured from the IR representing the path of the mouth-ears of the HATS, an increase in C_{80} is equivalent to saying that the proportion of sound arriving at the singer's ear within 80 ms was high compared to the sound arriving at the singer's ear after 80 ms. This point of clarification is in contrast to studies that analyze room acoustics from the perspective of the audience member. As such, the singers in this study tended to exhibit an increase in vibrato extent (i.e., a "wider" vibrato f_0 range) when a higher proportion of their sound resided in the early sound field than in the late sound field. It is possible that this result could be caused by participants reacting to the auditory feedback (or lack thereof) while singing.

A recent study suggests that loud speaking is self-reported as being less effortful in a reverberant room than in a semi-reverberant or anechoic room (Bottalico *et al.*, 2017). There is also evidence that singers adjust loudness, dynamic range, and fullness—as well as other timbral characteristics—when performing in different spaces (Luizard and Bernardoni, 2020). While the current study did not measure SIL or SPL of the singer-participants, past literature demonstrates that professional opera singers create sounds in excess of 97 dB (Björkner, 2008) (i.e., rather loud sounds). It is possible that the participants from this study increased their vocal load as a result of varying amounts of feedback while singing and, as such, altered their vocal function in a way that increased the amplitude modulation of their vibrato (V_{ext}). There is anecdotal evidence that heavier vocal production while singing could result in a slower vibrato rate and larger vibrato extent (Nix *et al.*, 2016). While it is not yet clear from the literature the extent to which auditory feedback impacts vibrato or vocal function (Schultz-Coulon, 1978; Grillo *et al.*, 2010; Mürbe *et al.*, 2007), there is a long-standing acknowledgment in the voice research community that the antagonism between the cricothyroid and the thyroarytenoid muscles is one possible explanation for the existence of vocal vibrato (Sundberg, 1987). As such, any adjustment—conscious or subconscious—that alters the function of either of those two muscles or the acoustic feedback within the vocal tract could result in a measurable difference in vibrato extent.

The models also found that vibrato extent was negatively associated with EDT. That is to say, as the perceived reverberance of a room (EDT) decreased, vibrato extent increased—and vice versa. This result can be considered similar to that of the association between vibrato extent and clarity. A low EDT indicates that a room is not particularly reverberant. As such, a room with a low EDT at the point of the singer's head—as would be the case in a semi-reverberant or anechoic room—might result in increased self-perceived vocal effort and could deleteriously affect vocal function.

As mentioned previously, Luizard and Bernardoni (2020) suggested that a significant and positive relationship between pitch and $IACC_{late}$ might be interpreted as an increase in vocal effort with a decrease in spatial envelopment or intimacy (given that $IACC_{late}$ and subjective spatial envelopment are inversely related). Furthermore, Huang *et al.* (1995) found that vocal effort tends to increase with CQ. Indeed, the results of this study show a negative association between $IACC_{late}$ and vibrato extent. Extrapolating from the conclusions of Luizard and Bernardoni (2020), it is possible that a decrease in a singer's sense of envelopment could produce an increase in CQ and, potentially, an increase in vocal effort. The culminating result of those functional and perceptual changes is a decrease in vibrato extent. Of course, the causes of changes in vibrato rate and vibrato extent are not well-understood; however, it seems from these results that the room acoustic environment does impact singers' vocal performance and, potentially, their vocal function.

C. Vibrato rate and room acoustics

Vibrato rate was found to have a negative association with C80 and a positive association with EDT. This relationship seems to indicate that when the feedback perceived by the singer is clearer and less reverberant, singers may change their vocal function in a way that produces a slower vibrato. Most importantly, it was found that a higher proportion of the variance in vibrato rate could be explained by individual singers than by the room in which those singers performed. In other words, measured differences in vibrato rate seem to be more a function of the singer recorded than an effect of the acoustic environment.

While the room condition did not account for a high degree of variance in vibrato rate between samples, it is notable that the vibrato rate results from this study adhered to trends reported in the literature. Singers who are assigned male at birth (AMAB) tended to have a slower vibrato rate than singers assigned female at birth (AFAB); putting this study in agreement with past reports (Guzman *et al.*, 2012; Nix *et al.*, 2016; Sundberg, 1994; Glasner and Johnson, 2020) (see Table III). Additionally, the average vibrato rate of the participants (5.9 Hz) fits within reported normative measures in previous studies (Nix, 2014; Hakes *et al.*, 1988; Titze, 1994; Sundberg, 1994) and, as such, situates the results from this study within current singing voice research literature.

D. Pitch inaccuracy

The control of fundamental frequency has been the subject of an impressive amount of scholarly discourse. While it is apparent that control of f_0 is of import to the professional singer, it is unclear to what extent auditory feedback or motor control training play in pitch accuracy (Titze, 1994). It also seems likely that auditory feedback has at least some influence on pitch accuracy and loudness (Grillo *et al.*, 2010; Mürbe *et al.*, 2007; Bauer *et al.*, 2006; Larson *et al.*, 2007, Bottalico *et al.*, 2017). The results of the current study show a negative association between pitch inaccuracy and ST_v. That is to say, as the support for the voice from the room (ST_v) increases, singers are more accurate. This finding represents a somewhat obvious conclusion. It stands to reason that the degree to which a singer can hear themself on stage may positively influence pitch accuracy.

E. Practical application

Practically speaking, this study confirms the existence of a phenomenon that performers have noted anecdotally for at least a century: singers tend to perform differently as a result of the acoustic environment in which they sing (Jordan, 1980). As such, a pedagogical consideration based on these findings is that singers should train in a variety of spaces so as to become accustomed to different acoustical environments. Singing teachers might also work to develop pedagogies and methods that prioritize not auditory feedback, but rather kinesthetic or somatosensory awareness to limit the possible negative implications of singing in a "live" or "dry" hall. Such methods could then be studied, analyzed, and optimized to enable singers to attain repeatable and functional results on stage.

V. CONCLUSIONS

The present study aims to investigate the effect of room acoustics environment on the vibrato extent, vibrato rate, and pitch inaccuracy of nine Western Classical singers while singing in five different acoustic environments. Vibrato extent was found to have (i) a positive association with C80, (ii) a negative association with EDT, and (iii) a negative association with $IACC_{late}$. That is to say, (i) both the clarity of the room environment at the position of the singer, and the amplitude modulation of the singer's vibrato (i.e., vibrato extent) tended to increase together, (ii) as the perceived reverberance of a room decreased, vibrato extent increased, and (iii) as the spatial envelopment or intimacy decreased, vibrato extent increased suggesting an increase in vocal effort. Vibrato rate was found to have (i) a negative association with C80 and (ii) a positive association with EDT. This means that when the feedback perceived by the singer is clearer and less reverberated, singers tend to produce a slower vibrato. Regarding pitch inaccuracy, the results of the current study show a negative association between pitch inaccuracy and ST_v. That is to say, as the support for the voice from the room (STv) increases, singers are more accurate. A pedagogical consideration based on these findings is that singers should train in a variety of spaces so as to become accustomed to different acoustical environments.

ACKNOWLEDGMENT

The kind cooperation of the singers has made this work possible and the support of the University of Illinois Urbana-Champaign College of Fine and Applied Arts. Special thanks to Lisa Lillig, Co-Director of Patron Services at Krannert Center for the Performing Arts for providing access to and use of the performance spaces.

Amengual Gari, S. V., Kob, M., and Lokki, T. (**2019**). "Analysis of trumpet performance adjustments due to room acoustics," in *Proceedings of the International Symposium on Room Acoustics*, Amsterdam, Netherlands (15–17 September), pp. 65–73.

Anand, S., Windgate, J. M., Smith, B., and Shrivastav, R. (**2012**). "Acoustic parameters critical for an appropriate vibrato," J. Voice **26**(6), e19–e25.

Astolfi, A., Bo, E., Aletta, F., and Strepi, L. (**2020**). "Measurements of acoustical parameters in the ancient open-air theatre of Tyndaris (Sicily, Italy)," Appl. Sci. **10**(16), 5680.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (**2014**). "Fitting linear mixed-effects models using lme4," J. Stat. Software **67**, 1–48.

Bauer, J. J., Mittal, J., Larson, C. R., and Hain, T. C. (**2006**). "Vocal response to unanticipated perturbations in voice loudness feedback: An automatic mechanism for stabilizing voice amplitude," J. Acoust. Soc. Am. **119**(4), 2363–2371.

Beranek, L. (**2004**). *Concert Halls and Opera Houses: Music, Acoustics, and Architecture* (Springer-Verlag, New York), pp. 15–18.

Björkner, E. (**2008**). "Musical theatre and opera singing—Why so different? A study of subglottal pressure, voice source, and formant frequency characteristics," J. Voice **22**(5), 533–540.

Bottalico, P., Graetzer, S., and Hunter, E. J. (**2017**). "Effect of training and level of external auditory feedback on the singing voice: Pitch inaccuracy," J. Voice **31**(1), 122-e9–122-e16

Bolzinger, S., Warusfel, O., and Kahle, E. (**1994**). "A study of the influence of room acoustics on piano performance," J. Phys. IV. France **04**(C5), C5-617–C5-620.

Brunskog, J., Gade, A., Paya-Ballester, G., and Reig-Calbo, L. (**2009**). "Increase in voice level and speaker comfort in lecture rooms," J. Acoust. Soc. Am. **125**, 2072–2082.

Dejonckere, P. H., Hirano, M., and Sundberg, J. (**1995**). *Vibrato* (Singular Publishing Group, San Diego, CA), pp. 9–34.

Ekholm, E., Papagiannis, G. C., and Chagnon, F. (**1998**). "Relating objective measurements to expert evaluation of voice quality in western classical singing: Critical perceptual parameters," J. Voice **12**, 182–196.

Farina, A. (**2000**). "Simultaneous measurement of impulse response and distortion with a swept-sine technique," in *Audio Engineering Society Convention 108*, Audio Engineering Society.

Ferrante, I. (**2011**). "Vibrato rate and extent in soprano voice: A survey on one century of singing," J. Acoust. Soc. Am. **130**, 1683–1688.

https://doi.org/10.1121/10.0011675

Glasner, J. D., and Johnson, A. M. (**2020**). "Effects of historical recording technology on vibrato in modern-day opera singers," J. Voice (published online).

Grillo, E., Verdolini Abbott, K., and Lee, T. D. (**2010**). "Effects of masking noise on laryngeal resistance for breathy, normal, and pressed voice," J. Speech. Lang. Hear. Res. **53**, 850–861.

Guzman, M. A., Dowdall, J., Rubin, A. D., Maki, A., Levin, S., Mayerhoff, R., and Jackson-Menaldi, M. C. (**2012**). "Influence of emotional expression, loudness, and gender on the acoustic parameters in classical singers," J. Voice **26**(5), 675.e5–675.e11.

Hain, T. C., Burnett, T. A., Larson, C. R., and Kiran, S. (**2001**). "The effects of masked and delayed auditory feedback on fundamental frequency modulation in vocal vibrato," J. Acoust. Soc. Am. **109**(5), 2146–2152.

Hakes, J., Shipp, T., and Doherty, E. T. (**1987**). "Acoustic properties of straight tone, vibrato, trill, and trillo," J. Voice **1**(2), 148–156.

Hakes, J., Shipp, T., and Doherty, E. T. (**1988**). "Acoustic characteristics of vocal oscillations: Vibrato, exaggerated vibrato, trill, and trillo," J. Voice **1**(4), 326–331.

Hox, J. J., Moerbeek, M., and van de Schoot, R. (**2017**). *Multilevel Analysis: Techniques and Applications* (Routledge, New York).

Huang, D. Z., Minifie, F. D., Kasuya, H., and Lin, S. X. (**1995**). "Measures of vocal function during changes in vocal effort level," J. Voice **9**(4), 429–438.

ISO (**2008**). 3382-2:2008(E): "Acoustics—Measurement of room acoustic parameters, Part 2: Reverberation time in ordinary rooms" (International Organization for Standardization, Geneva, Switzerland).

Jordan, V. L. (**1980**). *Acoustical Design of Concert Halls and Theatres: A Personal Account* (Applied Science, London), pp. 81–83.

Kato, K., Ueno, K., and Kawai, K. (**2015**). "Effect of room acoustics on musicians' performance. Part II: Audio analysis of the variations in performed sound signals," Acta Acust. Acust. **101**(4), 743–759.

Kleber, B., Friberg, A., Zeitouni, A., and Zatorre, R. (**2017**). "Experience-dependent modulation of right anterior insula and sensorimotor regions as a function of noise-masked auditory feedback in singers and nonsingers," Neuroimage **147**, 97–110.

Larson, C. R., Sun, J., and Hain, T. C. (**2007**). "Effects of simultaneous perturbations of voice pitch and loudness feedback on voice FO and amplitude control," J. Acoust. Soc. Am. **121**(5), 2862–2872.

Lester-Smith, R. A., Hilger, A., Dunne-Platero, K. E., Kim, J. H., Chan, C. L., and Larson, C. R. (**2022**). "The effects of masked and delayed auditory feedback on fundamental frequency modulation in vocal vibrato," J. Voice (published online).

Luizard, P., and Bernardoni, N. H. (**2020**). "Changes in the voice production of solo singers across concert halls," J. Acoust. Soc. Am. **148**(1), EL33–EL39.

Manfredi, C., Barbagallo, D., Baracca, G., Orlandi, S., Bandini, A., and Dejonckere, P. H. (**2015**). "Automatic assessment of acoustic parameters of the singing voice: Application to professional western operatic and jazz singers," J. Voice **20**(4), 517-e1–517.e9.

McKinney, J. C. (**2005**). *The Diagnosis & Correction of Vocal Faults: A Manual for Teachers of Singing & for Choir Directors* (Waveland Press, Long Grove, IL), pp. 198–204.

Mürbe, D., Zahnert, T., Kuhlische, E., and Sundberg, J. (**2007**). "Effects of professional singing education on vocal vibrato—A longitudinal study," J. Voice **21**(6), 683–688.

Nestorova, T., and Glasner, J. D. (**2021**). "What about extent?: Examining current vibrato extent metrics," PAVA InFormant **2**(3), 3–11.

Nix, J. (**2014**). "Shaken, not stirred: Practical ideas for addressing vibrato and nonvibrato singing in the studio and the choral rehearsal," J. Singing **70**, 411–418.

Nix, J., Perna, N., James, K., and Allen, S. (**2016**). "Vibrato rate and extent in college music majors: A multicenter study," J. Voice **30**(6), 756.E31–756.E41.

Prame, E. (**1994**). "Measurements of the vibrato rate of ten singers,' " J. Acoust. Soc. Am. **96**(4), 1979–1984.

Pelegrín-García, D. (**2011**). "Comment on 'Increase in voice level and speaker comfort in lecture rooms,' " J. Acoust. Soc. Am. **129**(3), 1161–1164.

Schärer Kalkandjiev, Z., and Weinzierl, S. (**2013**). "The influence of room acoustics on solo music performance: An empirical case study," Acta Acust. Acust. **99**(3), 433–441.

Schärer Kalkandjiev, Z., and Weinzierl, S. (**2015**). "The influence of room acoustics on solo music performance: An experimental study," Psychomusicol.: Music Mind Brain **25**(3), 195–207.

Schultz-Coulon, H. J. (**1978**). "The neuromascular control system and vocal function," Acta Oto-Laryngol. **86**, 142–153.

Scotto Di Carlo, N. (**1994**). "Internal voice sensitivities in opera singers," Folia Phoniatr. Logop. **46**(2), 79–85.

Seashore, C. (**1931**). "The natural history of the vibrato," Proc. Natl. Acad. Sci. U.S.A. **17**(12), 623–626.

Sundberg, J. (**1987**). *The Science of the Singing Voice* (Northern Illinois University Press, DeKalb, IL), pp. 163–165.

Sundberg, J. (**1994**). "Acoustic and psychoacoustic aspects of vocal vibrato," Quart. Prog. Rep. **35**(2-3), 45–68.

Ternström, S. (**1993**). "Long-time average spectrum characteristics of different choirs in different rooms," Speech Transm. Lab.: Quart. Prog. Rep. **30**(3), 15–31.

Titze, I. (**1994**). *Principles of Voice Production* (Prentice Hall, Iowa City, IA).

Ward, W. D., and Burns, E. M. (**1978**). "Singing without auditory feedback," J. Res. Singing **1**(2), 4–44.