DOI: 10.61782/fa.2025.0392

FORUM ACUSTICUM
EURONOISE 2025

EXPLORING THE ACOUSTIC CHARACTERISTICS THAT DETERMINE PERFORMANCE QUALITY IN OPERA SINGING

Haruka Kondo[1*] **Sotaro Kondoh**[1,2,3] **Shigeto Kawahara**[4] **Shinya Fujii**[2]
[1] Graduate School of Media and Governance, Keio University, Kanagawa, Japan
[2] Faculty of Environment and Information Studies, Keio University, Kanagawa, Japan
[3] Japan Society for the Promotion of Science, Tokyo, Japan
[4] The Keio Institute of Cultural and Linguistic Studies, Tokyo, Japan

ABSTRACT

Which acoustic characteristics of opera singing impact the judgment of perceived performance? To address this question, the current study recorded ten female singers performing Caro mio ben, which were then evaluated by four professional voice instructors, who provided an overall impression score on a 100-point scale and six vocal attributes rated on a 7-point Likert scale. We extracted three acoustic features from each recorded singing: singing power ratio (SPR), which represents the ratio of spectral peaks between 2-4 kHz and 0-2 kHz; harmonics-to-noise ratio (HNR), which measures the ratio of harmonic sound energy to noise in a voice signal, reflecting vocal quality; and sound pressure level (SPL), which is an acoustic measure that represents the loudness or intensity of a sound signal. A linear mixed-effects regression model, with the vocal attribute score as the dependent variable, three acoustic characteristics as fixed effects, and evaluators as a random effect, revealed that SPR was a significant predictor for several of the vocal attributes. These findings suggest that SPR, which reflects resonance within the 2–4 kHz range influences the judgment of performance quality in opera singing.

Keywords: *voice, vocal performance, opera singing, evaluation, singing power ratio.*

[*]***Corresponding author***: haruka_kondo@keio.jp.

1. INTRODUCTION

The musical competition is used extensively to identify talent and to determine musicians' careers. Competition results are generally determined by the judges' subjective scoring. However, detailed judgment criteria are rarely disclosed, and evaluations inevitably reflect judges' individual preferences and experiences. Consequently, it remains unclear which specific elements of a singer's voice contribute the most strongly to the overall score.

Previous research has investigated the subjective attributes that contribute to superior singing evaluation. These attributes include the singing technique [1], [2], perceived potential or talent based on voice [3]–[5], and vocal quality [6]. A survey of 1,000 vocal instructors identified vocal quality, intonation, and musicality as the most important factors [5]. Notably, the study by [7] provides valuable insights into which subjective evaluation scale items might explain the overall score. In their research, experts repeatedly evaluated recorded singing performance, and the consistency of evaluations within and between judges was assessed. Their findings revealed correlations between the overall score and attributes such as vibrato, resonance, timbre, and diction. However, this study did not employ statistical modeling to investigate which acoustic features contribute to vocal evaluation, leaving the underlying determinants of such evaluations still unresolved.

In addition, because judges assign overall scores based on the sound of the voice, objective acoustic features must influence their evaluations. Quantitative studies of subjective singing evaluations began in the 1920s [8]. Since then, researchers have investigated acoustic features that characterize high-quality voices and their correlation with subjective evaluations [7], [9], [10]. One particularly important aspect of opera singing is the ability of the voice to resonate throughout a large hall without amplification [11], [12]. This phenomenon, known as the "singer's formant," refers to a strong peak of vocal energy at approximately 3 kHz that enables a singer's voice to be heard clearly over an orchestra [11], [13]. The singer's formant can be quantified using the Singing Power Ratio (SPR), which measures the harmonic balance of a voice by comparing the strongest harmonic peak in the 2–4 kHz range with that in the 0–2 kHz range [14]. Higher SPR values, indicating a smaller difference in power between the 2–4 kHz and 0–2 kHz ranges, are associated with a bright, ringing tone [14]. Trained singers generally exhibit higher SPR values than untrained singers [5]. These results suggest that SPR likely plays a critical role in determining overall scores in opera singing.

Other indicators that may influence the overall score include the harmonic-to-noise ratio (HNR) and sound pressure level (SPL). HNR measures the amount of periodic (harmonic) energy in the voice and serves as an indicator of voice clarity [15], [16]. A higher HNR signifies a lower noise level, which listeners typically perceive as a clearer voice [17]. SPL is determined by the amplitude of the vocal signal and is directly related to the perceived loudness. Higher SPL values increase the presence of voice and tend to enhance the performance of other acoustic parameters, including HNR [18], [19].

Professional singers generally achieve high SPL without compromising voice quality [20]. Therefore, in addition to SPR, both HNR and SPL may play a significant role in determining the overall scores of opera singing. However, it remains unclear which acoustic features explain the vocal attribute score significantly.

This study aimed to identify the acoustic features that explain judges' evaluations of opera singing. To achieve this, we recorded opera performances, asked judges to assess them, and collected ratings for overall impression as well as six vocal characteristics: resonance, timbre, vibrato, diction, intonation, and expressiveness, based on previous research [7]. We then analyzed the acoustic features of the recordings, including SPR, HNR, and SPL. Seven linear mixed models were constructed to predict each of the evaluation scores—overall impression, resonance, timbre, vibrato, diction, intonation, and expressiveness—based on these acoustic features. Through these models, we sought to clarify the extent to which objective acoustic parameters contribute to judges' evaluations of opera singing.

2. MATERIALS AND METHODS

2.1 Participants

Ten female Japanese singers specializing in classical vocal music (mean age ± SD: 25.10 ± 4.41 years old) participated in this study. All participants were either currently enrolled in a music university, had graduated from a music university, or had received equivalent professional training.

The recordings of the ten singers were evaluated by four vocal instructors, all professional singers (4 females; mean age ± SD: 47.75 ± 12.26 years). The instructors had between 10 and 28 years of experience in vocal instruction (mean ± SD: 20.00 ± 8.12 years) and professional careers spanning 15 to 37 years (mean ± SD: 26.25 ± 10.44 years).

Ethical approval for the study was obtained from the Research Ethics Committee of Keio University Shonan Fujisawa Campus (Approval Number: 441). All participants were thoroughly informed of the experimental procedures and written consent was obtained prior to the experiment.

2.2 Procedure and data acquisition

The participants completed vocal exercises in a soundproof room before singing the assigned musical piece. The recorded data were used for acoustic analysis, and a separate evaluation session was conducted in which judges assessed the performance based on predefined criteria

2.2.1 Procedure for Singers

The recordings were conducted in a soundproof room at the Keio Institute of Cultural and Linguistic Studies in Tokyo, Japan. This room was specifically designed with an elevated floor to prevent the transmission of footstep vibrations, and silencers were installed in the air conditioning and ventilation ducts to eliminate ambient noise.

The recording setup included a computer (MacBook Retina 12-inch, 2017, macOS Monterey, Apple Inc., Cupertino, CA, USA) connected to an audio interface (M-TRACK 2X2M, M-AUDIO, Cumberland, RI, USA) and a microphone (AT2035, Audio-Technica, Tokyo, Japan). The microphone, which had a frequency response range of 20–20,000 Hz, was positioned 20 cm from the singer's mouth. Audio recordings were captured using the Audacity software (ver. 3.4.2) with a standardized sampling frequency of 19.2 kHz.

Before recording, the singers completed a vocal experience questionnaire detailing their vocal experience. They were then given 10 min of vocal warm-up in the soundproof room (AMG35, YAMAHA Corporation, Hamamatsu, Japan) to acclimatize to the recording environment. Following the warm-up, each singer performed the assigned piece, Caro mio ben, a cappella. The singers used music sheets placed on a stand during their performance, rather than singing from memory. The performance duration was set to two minutes, and each singer performed the piece only once. Caro mio ben, composed by Tommaso Giordani in 1859, was selected for its accessibility, manageable vocal range, and low technical difficulty, making it suitable for singers with varying levels of experience. In addition, this piece is commonly used by vocal students in Japan.

2.2.2 Procedure for Judges

The evaluation sessions were conducted in the same soundproof room (AMG35, YAMAHA Corporation, Hamamatsu, Japan) used for the recordings. Audio recordings were played on a computer (MacBook Retina 12-inch, 2017, macOS Monterey, Apple Inc., Cupertino, CA, USA) connected to headphones (HD280pro, Sennheiser, Wedemark, Germany). Before the session, the judges adjusted the playback volume to ensure consistent listening conditions across all the recordings.

Before beginning the evaluations, the judges completed a questionnaire detailing their vocal experiences and professional careers. They then listened to the recordings of the 10 singers, presented in a randomized order, and evaluated the performances based on two criteria: (1) an overall score on a 100-point scale and (2) six vocal attributes—vibrato, resonance, timbre, diction, intonation, and expressiveness—rated on a 7-point Likert scale (1 = very low, 7 = very high). These attributes were selected based on the previous research by Wapnick and Ekholm (1997). The judges were all professional singers; therefore, we did not provide formal definitions of the vocal attributes. However, for clarity, the six vocal attributes are described as follows: resonance refers to vocal depth and richness, whereas timbre represents tonal qualities such as brightness and warmth. Vibrato is characterized by its consistency, speed, and depth. Diction reflects pronunciation clarity and intelligibility, whereas intonation reflects pitch accuracy and stability. Expressiveness captures the singer's ability to convey emotions, use dynamics, and shape phrases effectively.

2.3 Analysis

2.3.1 Acoustic Analysis

We analyzed the acoustic features of the recorded singing performance using three parameters: SPR, HNR, and SPL. Praat software (version 6.3.10) was used for this analysis [21]. To ensure the accuracy of the acoustic analysis, the recorded audio files were preprocessed to isolate the sung portions of the performances. Non-singing segments such as pauses and breaths were excluded.

The Singing Power Ratio (SPR) was calculated following the methodology described by Omori et al. (1996). SPR was calculated using the power spectrum (dB) obtained from a Fast Fourier Transform (FFT) with a time window of 1024 points and a bandwidth of 4000 Hz. In the power spectrum of each sung, the most prominent harmonic peak was identified within the 0–2 kHz range and the 2–4 kHz range. The peak within the 0–2 kHz range represents the lower-frequency harmonic content, while the peak in the 2–4 kHz range corresponds to the so-called “singer’s formant” region. The SPR was then computed as the logarithmic power ratio (in decibels, dB) between these two peaks:

$$SPR\text{ (dB)} = Power_{High} - Power_{low} \quad \text{(Eq.1)}$$

The Harmonics-to-Noise Ratio (HNR) was calculated using the autocorrelation method implemented by Praat. HNR quantifies the ratio of the harmonic energy to the noise energy in the voice signal. The calculation is based on the following formula:

$$HNR\text{ (dB)} = Power_{Harmonics} - Power_{Noise} \quad \text{(Eq.2)}$$

where $Power_{Harmonics}$ represents the power of the harmonic component, and $Power_{Noise}$ represents the power of the noise component. The analysis was conducted using a frame-based window and the average HNR was calculated across the entire performance.

The Sound Pressure Level (SPL) was computed using Praat (version 6.3.10; [21]), based on the digital waveform of the recorded singing voices. SPL was calculated in decibels (dB) using the following formula:

$$SPL\text{(dB)} = 20log_{10}\left(\frac{P}{P_0}\right) \quad \text{(Eq.3)}$$

where P is the root mean square (RMS) pressure of the signal and P_0 is the reference pressure, set to 20 μPa in air.

Since no calibration with an external SPL meter was performed prior to data collection, the SPL values used in this study represent relative rather than absolute intensity levels. All recordings were made under consistent recording conditions (e.g., fixed microphone gain and distance) across participants, allowing for valid comparisons of relative loudness between performances. While absolute SPL values are not available, the relative SPL values serve as a meaningful acoustic indicator for within-study comparisons.

2.3.2 Statistics

To analyze the effects of acoustic features on vocal evaluations, seven linear mixed-effects models were constructed. The analysis was conducted using R software (version 4.4.2) with the lme4 and lmerTest packages [22], [23], which facilitated linear mixed-effects modeling with p-value estimation.

The models examined the contribution of three acoustic features—SPR, HNR, and SPL—to seven vocal evaluation scores: overall impression, resonance, timbre, vibrato, diction, intonation, and expressiveness. In the model, acoustic features were treated as fixed effects, and judge variabilities were included as a random effect. The model formula is as follows:

$$Vocal\ Evaluations\ for\ Judges \sim SPR + HNR + SPL + (1|\,Judge\ ID) \quad \text{(Eq.4)}$$

To assess multicollinearity among the acoustic predictors, the variance inflation factor (VIF) was calculated using the car package [24]. The significance level α was set at 0.05 for all analyses.

3. RESULTS

3.1 Effects of Acoustic Features on Overall Score

The results of the linear mixed-effects model (Eq.4) assessing the influence of acoustic features on overall scores are summarized in Table 1. Among the three acoustic features, SPR had a significant effect on the overall impression scores (β = 1.86, p = 0.034; Figure 2A). In contrast, HNR (β = 0.95, p = 0.580; Figure 2B) and SPL (β = 0.88, p = 0.445; Figure 2C) did not exhibit statistically significant effects. All VIF values were below 10, indicating no concerns regarding multicollinearity.

Both $R^2{}_m$ and $R^2{}_C$ were 0.183 (95% CI: 0.062–0.479 for R2 and 0.044–0.420 for $R^2{}_C$). The small difference between $R^2{}_m$ and R^2 suggests that variability among judges contributes minimally to the model's explanatory power.

Table 1. Estimation of Linear Mixed-Effects Models Fitted to Overall Score (Fixed Effects: Acoustic Features). Asterisk (*) shows statistical significance.

Variable	*β*	*SE*	*df*	*t-value*	*p-value*	*VIF*
SPR	1.86	0.84	36.00	2.21	0.034*	1.84
HNR	0.95	1.69	36.00	0.56	0.580	2.91
SPL	0.88	1.14	36.00	0.77	0.445	2.01

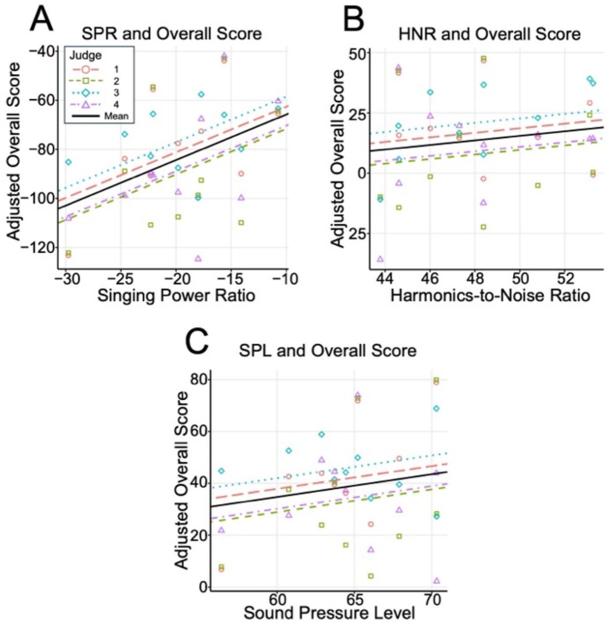

Figure 2. Scatter plots of acoustic features versus adjusted overall scores.

3.2 Effects of Acoustic Features on Six Vocal Attributes

To examine the effects of acoustic features on vocal evaluations, seven linear mixed-effects models were constructed with SPR, HNR, and SPL as fixed effects and random intercepts for each judge (Vocal Evaluation ~ SPR + HNR + SPL + (1 | Judge ID)).

For the Resonance score (Table 2), SPR was found to be a significant predictor (*β* = 0.18, *p* = 0.003), indicating that higher SPR values were associated with higher resonance evaluations. HNR and SPL were not significant predictors (p = 0.765 and p = 0.365, respectively). In the model predicting Vibrato scores (Table 3), SPR was also a significant predictor (*β* = 0.12, *p* = 0.043), suggesting that higher SPR values corresponded to higher vibrato ratings. Neither HNR nor SPL reached significance (*p* > 0.3 for both). For Timbre evaluations (Table 4), SPR again significantly predicted scores (*β* = 0.18, *p* = 0.004). HNR (*β* = 0.22, *p* = 0.071) and SPL (*β* = 0.15, *p* = 0.070) showed marginal trends toward significance, suggesting potential relevance with a larger sample size. In the Diction model (Table 5), SPR was a significant predictor (*β* = 0.24, *p* = 0.002), while HNR and SPL were not significant (*p* = 0.365 and *p* = 0.930, respectively). No significant predictors were observed in the models for Intonation (Table 6) or Expressiveness (Table 7), with all p-values exceeding.

Overall, SPR consistently showed significant positive associations with multiple vocal evaluation scores, particularly resonance, vibrato, timbre, and diction, while HNR and SPL did not demonstrate significant effects across any of the models.

Table 2. Estimation of Linear Mixed-Effects Models Fitted to Resonance Score (Fixed Effects: Acoustic Features). Asterisk (*) shows statistical significance.

Variable	*β*	*SE*	*df*	*t-value*	*p-value*
SPR	0.18	0.06	36.00	3.14	0.003*
HNR	0.04	0.12	36.00	0.30	0.765
SPL	0.07	0.08	36.00	0.92	0.365

Table 3. Estimation of Linear Mixed-Effects Models Fitted to Vibrato Score (Fixed Effects: Acoustic Features). Asterisk (*) shows statistical significance.

Variable	*β*	*SE*	*df*	*t-value*	*p-value*
SPR	0.12	0.06	36.00	2.10	0.043*
HNR	-0.12	0.12	36.00	-1.03	0.312
SPL	-0.07	0.08	36.00	-0.90	0.371

Table 4. Estimation of Linear Mixed-Effects Models Fitted to Timbre Score (Fixed Effects: Acoustic Features). Asterisk (*) shows statistical significance.

Variable	*β*	*SE*	*df*	*t-value*	*p-value*
SPR	0.18	0.06	33.00	3.12	0.004*
HNR	0.22	0.12	33.00	1.87	0.071
SPL	0.15	0.08	33.00	1.87	0.070

Table 5. Estimation of Linear Mixed-Effects Models Fitted to Diction Score (Fixed Effects: Acoustic Features). Asterisk (*) shows statistical significance.

Variable	*β*	*SE*	*df*	*t-value*	*p-value*
SPR	0.24	0.07	33.00	3.30	0.002*
HNR	0.13	0.14	33.00	0.92	0.365
SPL	-0.01	0.10	33.00	-0.09	0.930

Table 6. Estimation of Linear Mixed-Effects Models Fitted to Intonation Score (Fixed Effects: Acoustic Features). Asterisk (*) shows statistical significance.

Variable	*β*	*SE*	*df*	*t-value*	*p-value*
SPR	0.09	0.06	33.00	1.63	0.113
HNR	-0.06	0.11	33.00	-0.55	0.586
SPL	-0.08	0.07	33.00	-1.12	0.271

Table 7. Estimation of Linear Mixed-Effects Models Fitted to Expressiveness Score (Fixed Effects: Acoustic Features). Asterisk (*) shows statistical significance.

Variable	*β*	*SE*	*df*	*t-value*	*p-value*
SPR	0.07	0.07	36.00	1.16	0.256
HNR	-0.05	0.13	36.00	-0.34	0.738
SPL	0.08	0.09	36.00	0.94	0.353

4. DISCUSSION

This study investigated how specific acoustic features—Singing Power Ratio (SPR), Harmonic-to-Noise Ratio (HNR), and Sound Pressure Level (SPL)—influence judges' evaluations of opera singing. By analyzing recordings of trained vocalists performing the classical Italian song Caro mio ben, and by modeling both overall impression scores and six specific vocal attributes using linear mixed-effects models, we aimed to uncover which objective acoustic parameters most strongly contribute to the subjective evaluation of operatic performance.

4.1 Influence of SPR on Overall Score, Resonance, Vibrato, Timbre, and Diction

A central characteristic of opera singing is the "singer's formant" cluster, often reflected by a strong spectral peak between approximately 2–4 kHz [11], [13]. SPR, which compares the power in the 2–4 kHz range to the power in the 0–2 kHz range [14], is believed to be integral to the perceived "ring" or "resonance" in trained opera voices. Higher SPR values, indicating a more prominent high frequency "singer's formant," can help the voice project over an orchestra without amplification, which in turn may be perceived by listeners (including trained judges) as improved resonance, clarity, and overall vocal quality [5].

4.1.1 Resonance

The strong association between SPR and resonance is consistent with prior work emphasizing the role of the singer's formant in generating the impression of a resonant, projecting sound [25]. A higher SPR directly supports the voice's ability to ring in a large concert hall, leading to higher perceived resonance scores.

4.1.2 Vibrato

While vibrato is often characterized by oscillations in fundamental frequency, the fact that SPR significantly predicted vibrato scores may reflect how a well-defined singer's formant can enhance the perceived steadiness and richness of vibrato. Listeners may hear a more prominent vibrato "color" when the overtones in the 2–4 kHz range are strong, as these partials can contribute to the perception of vibrato depth and warmth.

4.1.3 Timbre

Timbre is closely tied to the spectral energy distribution of the voice. A higher SPR suggests a bright, ringing quality typically associated with an operatic timbre [14]. It is thus unsurprising that SPR was a consistent predictor of perceived timbre, as altering the balance of energy in the higher formant region changes the color or "ring" of the voice.

4.1.4 Diction

Although SPR significantly predicted diction scores, this association may represent a spurious correlation, as SPR does not directly capture articulatory precision or phoneme clarity. It is possible that voices exhibiting a stronger singer's formant were perceived as more intelligible overall, thereby receiving higher diction ratings, even though SPR is not a direct acoustic measure of diction.

4.2 Limitations of SPR in Predicting Intonation and Expressiveness

4.2.1 Intonation

Intonation depends primarily on pitch accuracy and control, which is not directly captured by SPR. The singer's formant relates to how the harmonic structure above the fundamental is amplified, rather than how stably the fundamental is produced at the correct frequency. Singers who excel at pitch accuracy may do so regardless of how their voice resonates in the 2–4 kHz range. Thus, variability in intonation scores may be driven more by fundamental frequency control rather than spectral energy distribution.

4.2.2 Expressiveness

Expressiveness in opera singing can involve dynamic contrast, emotional interpretation, phrasing, and stylistic nuances beyond purely spectral or amplitude-based measures. For instance, singers might manipulate tempo rubato, vary vibrato rate or extent, and use subtle dynamic shifts for emotional effect—none of which are directly captured by SPR, HNR, or SPL. Consequently, the judges' expressiveness ratings may reflect interpretive and artistic choices rather than the presence of a strong singer's formant.

4.3 Limited Predictive Power of HNR and SPL in Opera Singing Evaluation

4.3.1 HNR

This parameter measures the ratio of periodic (harmonic) energy to non-periodic (noise) energy in the voice signal [15]. HNR is often linked to perceived clarity or breathiness. However, in highly trained opera singers, HNR values may be sufficiently high and relatively uniform across performers, yielding little variation for statistical detection. Additionally, the judge's perception of "clarity" might overlap more with resonance or timbre, both of which appear better captured by SPR.

4.3.2 SPL

While SPL reflects vocal loudness and is crucial for projection, it may not be the deciding factor once a singer consistently achieves a certain threshold of loudness. Most trained opera singers can produce adequate volume, potentially minimizing SPL-related differences. Moreover, the singer's formant often plays a larger role than sheer loudness in enabling the voice to be heard over an orchestra [11]. Therefore, although SPL might help a voice carry, it may not contribute as decisively to specific vocal evaluations once all singers exceed a baseline loudness.

4.4 Limitations and Future Directions

While the present study provides valuable insights into the relationship between acoustic features and vocal evaluations in opera singing, several limitations should be acknowledged. First, the number of participants, both singers and judges, was relatively small. A limited sample size can reduce statistical power and may constrain the generalizability of the findings. Future studies should aim to recruit a larger and more diverse group of vocalists and evaluators, including singers of different voice types, experience levels, and stylistic backgrounds, to explore whether the observed relationships hold across broader populations.

Second, all performances were recorded in a soundproof room and evaluated through recorded playback rather than in a live performance setting. While such controlled environments ensure consistency in acoustic measurement, they may not fully capture the complexities of live opera singing, including spatial acoustics, visual cues, and real-time audience interaction, all of which can influence both vocal production and evaluative judgment. Moreover, playback through headphones or speakers may alter the perceived balance of frequencies—particularly in the higher range, where SPR is most relevant—potentially affecting how judges perceive vocal attributes like resonance or projection.

Taken together, these methodological constraints highlight the need for future research that combines controlled acoustic analysis with ecologically valid performance conditions. Studies incorporating live performance evaluations, larger and more diverse samples, and additional acoustic and perceptual parameters will further clarify the multidimensional nature of operatic vocal assessment.

5. FUNDINGS

This work was supported by the Taikichiro Mori Memorial Research Grant, the JST SPRING program (Grant No. JPMJSP2123) awarded to HK, and the JSPS KAKENHI (Grant No. 24H02199) awarded to SF.

6. ACKNOWLEDGMENTS

We sincerely thank Dr. Patrick Savage for his insightful feedback on the experimental design and data analysis. Additionally, we extend our heartfelt appreciation to Ms. Yuna Sakakibara for her valuable advice on statistical analysis and to Ms. Aiko Watanabe for her helpful suggestions on visualizing the results. Finally, we express our sincere gratitude to the judges and singers who generously contributed their time and expertise to this study.

7. REFERENCES

[1] R. F. Subotnik, "Adolescent pathways to eminence in science: lessons from the music conservatory. In P. Csermely & L. M. Lederman (Eds.)," *Science Education. Talent Recruitment and Public Understanding*, pp. 295-301. Amsterdam: IOS Press., 2003.

[2] R. F. Subotnik, "Transforming elite level musicans into professional artists: a view of the talent development process at the Juilliard School. In L. V. Shavinina & M. Ferrari (Eds.)," *Beyond knowledge: extra cognitive aspects of developing high ability*, pp. 137-167. Mahwah, MJ, London: Erlbaum Associates, 2004.

[3] H. Hollien, "That golden voice--talent or training?," *J. Voice*, vol. 7, no. 3, pp. 195–205, Sep. 1993.

[4] J. W. Davidson and D. Da Costa Coimbra, "Investigating performance evaluation by assessors of singers in a music college setting," *Music Sci.*, vol. 5, no. 1, pp. 33–53, Mar. 2001.

[5] C. Watts, K. Barnes-Burroughs, M. Andrianopoulos, and M. Carr, "Potential factors related to untrained singing talent: a survey of singing pedagogues," *J. Voice*, vol. 17, no. 3, pp. 298–307, Sep. 2003.

[6] J. M. Geringer and C. K. Madsen, "Musicians' ratings of good versus bad vocal and string performances," *J. Res. Music Educ.*, vol. 46, no. 4, pp. 522–534, Dec. 1998.

[7] J. Wapnick and E. Ekholm, "Expert consensus in solo voice performance evaluation," *J. Voice*, vol. 11, no. 4, pp. 429–436, Dec. 1997.

[8] C. E. Seashore and M. Metfessel, "Deviation from the regular as an art principle," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 11, no. 9, pp. 538–542, Sep. 1925.

[9] C. W. Robison, B. Bounous, and R. Bailey, "Vocal Beauty: A study proposing its acoustical definition and relevant causes in classical baritones and female belt singers," *Journal of Singing*, vol. 51, pp. 19–30, 1994.

[10] E. Ekholm, G. C. Papagiannis, and F. P. Chagnon, "Relating objective measurements to expert evaluation of voice quality in Western classical singing: critical perceptual parameters," *J. Voice*, vol. 12, no. 2, pp. 182–196, Jun. 1998.

[11] J. Sundberg, "What's so special about singers?," *J. Voice*, vol. 4, no. 2, pp. 107–119, Jan. 1990.

[12] R. T. Sataloff, *Vocal health and pedagogy: Science, assessment, and treatment*. Plural Publishing, 2017.

[13] W. T. Bartholomew, "A physical definition of 'good voice-quality' in the male voice," *J. Acoust. Soc. Am.*, vol. 6, no. 1, pp. 25–33, Jul. 1934.

[14] K. Omori, A. Kacker, L. M. Carroll, W. D. Riley, and S. M. Blaugrund, "Singing power ratio: quantitative evaluation of singing voice quality," *J. Voice*, vol. 10, no. 3, pp. 228–235, Sep. 1996.

[15] Y. Qi and R. E. Hillman, "Temporal and spectral estimations of harmonics-to-noise ratio in human

voice signals," *J. Acoust. Soc. Am.*, vol. 102, no. 1, pp. 537–543, Jul. 1997.

[16] P. J. Murphy, K. G. McGuigan, M. Walsh, and M. Colreavy, "Investigation of a glottal related harmonics-to-noise ratio and spectral tilt as indicators of glottal noise in synthesized and human voice signals," *J. Acoust. Soc. Am.*, vol. 123, no. 3, pp. 1642–1652, Mar. 2008.

[17] C. T. Ferrand, "Harmonics-to-noise ratio: an index of vocal aging," *J. Voice*, vol. 16, no. 4, pp. 480–487, Dec. 2002.

[18] P. Bottalico, S. Graetzer, and E. J. Hunter, "Effect of training and level of external auditory feedback on the singing voice: Volume and quality," *J. Voice*, vol. 30, no. 4, pp. 434–442, Jul. 2016.

[19] M. Brockmann-Bauser, J. E. Bohlender, and D. D. Mehta, "Acoustic perturbation measures improve with increasing vocal intensity in individuals with and without voice disorders," *J. Voice*, vol. 32, no. 2, pp. 162–168, Mar. 2018.

[20] M. Yadav and D. Cabrera, "Autophonic loudness of singers in simulated room acoustic environments," *arXiv [eess.AS]*, 17-Jun-2023.

[21] P. Boersma and D. Weenink, "Praat: doing phonetics by computer [Computer program]," *http://www.praat.org/*, 2024.

[22] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models Usinglme4," *J. Stat. Softw.*, vol. 67, no. 1, pp. 1–48, 2015.

[23] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "LmerTest package: Tests in linear mixed effects models," *J. Stat. Softw.*, vol. 82, no. 13, 2017.

[24] J. Fox and S. Weisberg, *An R Companion to Applied Regression (Third). Sage*. 2019.

[25] J. Sundberg, *Science of the singing voice*. DeKalb, IL: Northern Illinois University Press, 1987.