
Acoustic Differences between Pop and Musical Theater Singing in Chinese Original Musicals

Zheng Lujing¹, Sumathi Maniam Raj^{2*}, Ter Wei Shean³

Faculty of Creative Design, City University Malaysia

Abstract

Purpose

This study investigates the acoustic and emotional expression mechanisms of pop and musical theater singing within the context of Chinese original musicals. It aims to fill a gap in current research by systematically analyzing how vocal production strategies relate to emotional communication on stage.

Method

Using a comparative acoustic analysis approach, the study examines key parameters including vibrato rate and extent, formant frequency distribution, and harmonic structure. Data were collected from two professional singers performing the same excerpt from Jinsha, a representative Chinese original musical. Tools such as Praat, MATLAB, and Origin were used to extract and visualize acoustic data.

Results

Findings reveal that pop singing demonstrates higher vibrato rates (6.10 Hz), greater vibrato extent (7.53 Hz), and lower formant placement ($F1 = 710.4$ Hz), contributing to a warmer and more emotionally flexible vocal quality. Musical theater singing, in contrast, features more centralized vibrato (rate = 5.4 Hz, extent = 16.04 Hz) and higher formant frequencies ($F1 = 648.9$ Hz), promoting vocal clarity and projection. These distinctions reflect different acoustic strategies for emotional delivery and audience reception.

Discussion

Pop singing was found to be more effective in conveying high-arousal emotions and psychological complexity, while musical theater singing supports textual intelligibility and narrative coherence. The results align with prior models in emotional acoustics and voice perception, supporting the functional divergence of singing styles in theatrical performance.

Conclusion

The study reveals that vocal timbre, pitch dynamics, and resonance distribution play critical roles in shaping emotional expression across styles. Findings offer theoretical insights for cross-genre vocal training and practical implications for original musical theater productions in China.

Keywords: *Acoustic analysis; Emotional expression; Pop singing; Musical theater; Vibrato; Fundamental frequency; Formants; Chinese original musicals*

1. Introduction

In recent years, the rapid rise of Chinese original musicals within the cultural industry has drawn increasing scholarly attention to their localized expressions and the artistic integration of Eastern and Western elements. As a multidisciplinary art form combining music, drama, and dance, musical theatre not only expands the expressive dimension of narrative but also enriches the channels for emotional communication (Feng, 2024). Notably, in Chinese original musicals, creators often incorporate indigenous cultural imagery and Eastern aesthetics into musical structure and vocal performance, shaping a distinctive artistic identity with intellectual property independence.

Against this developmental backdrop, pop singing has emerged as a vital expressive medium in Chinese musical theatre, owing to their flexibility and emotional versatility. Compared to classical or operatic singing, pop singing emphasizes techniques such as mix voice, falsetto, and diaphragmatic-chest coordination. These approaches enable nuanced timbral shifts, enhancing the emotional depth and individuality of theatrical characters (Lee et al., 2023). Such technical features also grant greater expressive freedom in character development. A study conducted in 2022 demonstrated that belt singing, as a hallmark of pop technique, outperforms traditional operatic methods in terms of glottal pressure and closure efficiency, offering increased vocal output and emotional intensity—especially suited for high-affect performance scenarios (Fleischer et al., 2022).

Simultaneously, advances in acoustic analysis have opened up quantitative pathways for studying the relationship between vocal technique and emotional expression. A study in 2023 reported that the Voice Wellness Index (VWI), based on multiparametric acoustic indicators, achieved high diagnostic accuracy (AUC = 0.99), offering a robust tool for distinguishing singing styles through objective measures (Uloza et al., 2023). This affirms the feasibility of acoustic-based voice training in musical theatre. From an affective perspective, research has shown that intense emotions such as anger and joy are often accompanied by increased sound

pressure and pitch variability, whereas tender or sad emotions tend to feature lower intensity and more stable frequencies (Scherer et al., 2017). These findings support the expressive advantages of pop singing in conveying complex emotional layers.

Moreover, the Frequency of Half Energy (FHE) metric was introduced as a way to quantify resonance energy distribution across formants, enabling structural comparisons of vocal timbre between voice types (Müller et al., 2022). This contributes to the understanding of how resonance patterns correlate with perceived emotion. Related studies have indicated that vocal modes such as falsetto and mixed voice differ in their acoustic and emotional perceptual cues. In particular, timbre has been identified as a primary factor in listeners' recognition of "pleasant" emotions (Nussbaum et al., 2022), further underscoring the potential of pop singing in delivering subtle emotional nuance.

Although previous studies have demonstrated the expressive potential of pop singings from technical, acoustic, and perceptual perspectives, there remains a lack of systematic investigation into their acoustic-emotional mechanisms within the context of Chinese original musicals. In particular, little is known about how different vocal modes manipulate formant frequency distribution, vocal tract resonance, and dynamic vibrato parameters to support multi-layered emotional expression in theatrical performance. These internal acoustic strategies have yet to be comprehensively theorized in a localized cultural framework.

To address this gap, the present study adopts an acoustic analysis approach to quantitatively compare pop and musical theater singing styles in terms of three key dimensions: formant frequency distribution, harmonic structure, and vibrato rate and extent. It further explores how these acoustic differences contribute to emotional expression and character development in the context of Chinese musical theatre. By integrating quantitative metrics with performance-based vocal data, this study aims to offer empirical support for technical training in musical theatre and provide theoretical grounding for understanding the cultural adaptability of emotion-driven acoustic strategies.

This research centers on the following questions in the context of Chinese original musicals:
How does the formant frequency of pop singing affect emotional expression through vocal tract adjustment?

How do vibrato rate and extent differ between pop and musical theater singing in reflecting emotional arousal?

How do acoustic differences between the two vocal styles influence their respective emotional expression strategies?

2. Research Methods

2.1 Research Design and Sample

This study employs acoustic analysis to quantitatively compare pop and musical theater singing styles in Chinese original musicals, focusing on differences in vocal production and emotional expression. The analysis centers on three key acoustic dimensions: vibrato rate and extent, harmonic distribution, and formant distribution. By establishing correlations between these acoustic parameters and expressive functions, the study aims to identify the underlying mechanisms of vocal timbre modulation and emotional delivery across styles, offering both theoretical grounding and practical guidance for vocal pedagogy and stage performance in musical theatre.

The vocal material selected for analysis is a representative excerpt titled *Dangshi* from the Chinese original musical *Jinsha*. This segment features rich emotional dynamics, contextual shifts, and diverse vocal techniques—including mixed voice and falsetto—making it an ideal sample for examining how different vocal modes perform under varying emotional conditions. To ensure data validity and stylistic comparability, two experienced vocalists were recruited for recording. One is a professional musical theater actor with extensive stage experience in large-scale productions; the other is an active pop vocalist engaged in both commercial performances and university-level instruction in contemporary vocal techniques. All recordings were conducted in a professional studio with controlled acoustic conditions. The microphone model, sampling rate, and recording distance were standardized across all sessions to minimize extraneous variables and ensure the reliability and consistency of the acoustic data.

2.2 Data Collection

During the data collection phase, Adobe Audition was employed to ensure audio clarity and consistency across samples. All recordings were conducted in a professionally soundproofed studio using high-quality condenser microphones. The sampling rate was standardized at 44.1 kHz (i.e., 44,100 Hz), a widely accepted resolution in acoustic research that enables the precise capture of subtle sonic variations and supports high-resolution analysis.

To avoid the distortive effects of compression on acoustic properties, all audio files were saved in uncompressed WAV format. Throughout the recording process, Adobe Audition was used for real-time waveform monitoring to regulate input levels and prevent peak distortion. Basic

signal conditioning was applied, including the removal of background noise and electrical interference, in order to improve the acoustic purity and scientific validity of the recordings.

Following data capture, Adobe Audition was further utilized for initial audio preprocessing. Procedures included targeted noise reduction and spectral equalization, ensuring that the recordings were consistent in both amplitude and frequency distribution. These steps established a stable foundation for subsequent acoustic analysis using Praat.

2.3 Data Analysis Methods

2.3.1 Vibrato Analysis

Vibrato characteristics in both pop and musical theater singing were analyzed using Praat, with particular focus on two parameters: vibrato rate and vibrato extent. The original recordings were imported at a 44.1 kHz sampling rate to ensure high temporal and frequency resolution. Pitch contours were first extracted from each vocal segment using Praat, capturing the time-series variation of fundamental frequency (F0). These time-domain pitch trajectories were then subjected to frequency-domain analysis using fast Fourier transform (FFT) in MATLAB, enabling the identification of periodic oscillation patterns.

Vibrato rate was defined as the number of pitch modulations per second, while vibrato extent referred to the amplitude of the pitch fluctuations, i.e., the frequency deviation range. To improve analytical precision, the pitch series were smoothed and high-pass filtered prior to processing, reducing background noise interference. This method allowed for the quantification of pitch modulation strategies across vocal styles, serving as a foundation for subsequent modeling of acoustic-emotional relationships.

2.3.2 Harmonic Structure Analysis

To further explore spectral distinctions between singing styles, harmonic structure analysis was conducted on all audio samples. Using FFT-based spectral analysis on the raw waveforms extracted in Praat, frequency spectra ranging from 0 to 15 kHz were generated. The analysis focused on the number of harmonic components, their frequency positions, and amplitude distribution. Prior to spectrum extraction, all samples underwent band-pass filtering to eliminate environmental noise and highlight primary frequency content.

The resulting spectrograms were evaluated in terms of timbral richness, spectral brightness, and energy concentration—key acoustic parameters associated with stylistic timbre differences

and emotional rendering capacity. This analysis provided insight into how harmonic structure contributes to the expressive affordances of each vocal mode.

2.3.3 Formant Analysis

Formant extraction was performed to analyze vocal tract strategies and vowel shaping mechanisms in each singing style. Using Praat, the first three formants (F1, F2, and F3) were identified and extracted from each vocal sample. The results were then imported into Origin software for visualization and comparative analysis. Frequency positions and amplitude peaks of F1–F3 were plotted using curve-fitting and peak annotation functions.

Formant frequencies served as critical indicators of articulatory configuration. F1 is generally associated with glottal aperture, while F2 reflects tongue positioning and front oral cavity shaping. Together, these parameters revealed structural patterns of vowel formation and timbre control in different vocal techniques. This analysis provided empirical data for interpreting the relationship between vocal style and resonant tract modulation.

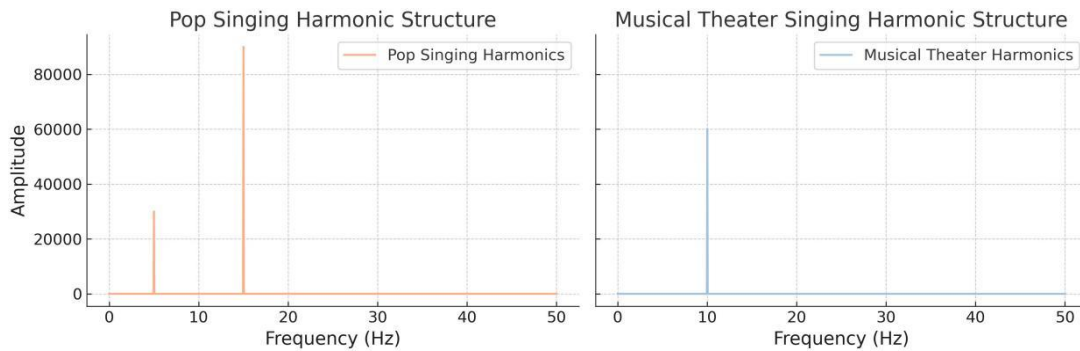
3. Research Results

3.1 Differences in Harmonic and Formant Distribution

To further examine the essential acoustic differences in timbre construction between pop and musical theater singing, this section presents a comparative analysis of harmonic structures and formant distributions across the two styles. Key parameters under investigation include spectral bandwidth, energy density, resonance frequency positions, and their variation patterns. These acoustic features not only reflect underlying vocal tract adjustment mechanisms but also directly influence perceptual qualities such as timbral fullness, projection, and speech intelligibility—critical indicators for distinguishing stylistic vocal identities.

In the harmonic analysis, spectral data revealed that pop singing exhibits a richer harmonic structure, with energy concentrated primarily in the 0–10 kHz range. The dominant spectral peaks are generally located around 4–5 kHz, and energy decays gradually at higher frequencies, forming a relatively even harmonic attenuation slope (*Figure 1*).

Figure 1. Harmonic Structures in Pop vs. Musical Theater Singing



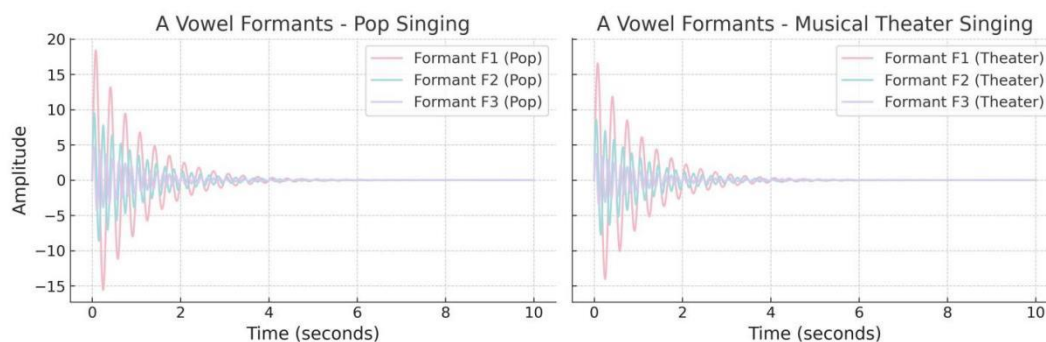
Note: This spectrogram shows that pop singing has stronger mid-range harmonics (~4–5 kHz), resulting in a brighter timbre. Musical theater emphasizes lower fundamentals, yielding a rounder tone. These spectral patterns reflect stylistic differences in resonance and harmonic emphasis.

Source: Developed for this research

This spectral pattern contributes to a warmer and rounder timbral quality in pop singing, creating a sense of acoustic envelopment that is particularly effective for expressing intimate emotional states. The harmonic structure is also more uniform and continuous, endowing the voice with greater emotional elasticity and dynamic layering—traits well-suited for conveying subtle and fluid internal emotions.

As shown in **Figure 2**, the formant trajectories in musical theater singing are more centralized and display less fluctuation. In contrast, pop singing exhibits lower resonance frequency concentrations, suggesting a tendency toward “softened” vocal production. For example, in the vowel /a/, the first three formant frequencies approximate 800 Hz (F1), 1200 Hz (F2), and 2500 Hz (F3); for /e/, they are approximately 500 Hz, 1600 Hz, and 2700 Hz, respectively. This lower-frequency formant layout enhances vocal warmth and smoothness while also indicating a more relaxed laryngeal and pharyngeal posture during phonation. Such adjustments likely facilitate a more natural and less constricted sound pathway, supporting the timbral qualities associated with emotional subtlety and authenticity.

Figure 2. Formant Trajectories of Three Vowels in Pop and Musical Theater Singing



Note: The graph compares F1–F3 trajectories for /a/, /e/, and /i/. Pop singing shows wider formant variation, especially in F1 and F2, indicating flexible vocal tract tuning. Musical theater displays greater formant stability, supporting articulation clarity.

Source: Developed for this research

In contrast, the harmonic structure of musical theater singing exhibits greater concentration and high-frequency emphasis. The spectral extension spans up to 0–15 kHz, with dominant frequency peaks commonly located above 6–7 kHz. These spectral maxima reflect a strong frequency-focusing effect, enhancing vocal brightness and projection—an intentional functional adjustment suited to the acoustic diffusion demands of theatrical spaces.

Moreover, the spectrogram reveals that energy distribution is more concentrated near the primary resonance zones, particularly around F2 and F3. This reflects a “focused” resonance strategy that facilitates clearer articulation of lyrics and maintains high levels of speech intelligibility.

With regard to formant distribution, musical theater singing displays higher vowel formant frequencies overall, characterized by greater vocal tension and elevated vocal tract positioning. For example, the /a/ vowel registers at approximately 900 Hz (F1), 1500 Hz (F2), and 3000 Hz (F3), while the /e/ vowel reaches 600 Hz, 2000 Hz, and 3200 Hz, respectively. This high-frequency resonance configuration supports linear acoustic projection, enabling the voice to effectively penetrate orchestral textures and reverberant theatre environments, thereby preserving clarity at long distances. It also reflects a vocal production approach that prioritizes speech intelligibility and semantic delivery in performance.

In summary, pop singing emphasizes a broader harmonic spread and lower-frequency formant reinforcement, yielding a softer, more intimate, and emotionally nuanced timbral profile.

Conversely, musical theater singing achieves a brighter, more focused acoustic output through enhanced high-frequency harmonics and elevated formant placement—characteristics that align with the demands of stage performance and long-range audience communication.

3.2 Comparative Analysis of Vibrato Rate and Extent

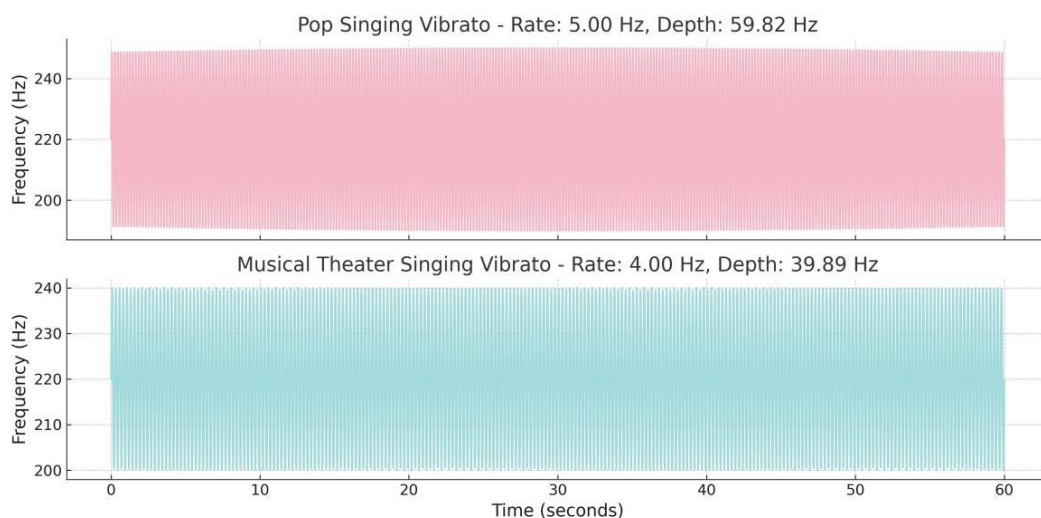
To investigate stylistic differences in emotional expression, this section focuses on the acoustic analysis of vibrato rate and vibrato extent in pop and musical theater singing. As a key indicator of vocal modulation, vibrato rate reflects the number of frequency oscillations per second, while vibrato extent represents the amplitude of pitch fluctuations. Together, these parameters influence the dynamic layering and emotional tension of vocal timbre.

The analysis revealed that the vibrato rate for pop singing is approximately 5.00 Hz, indicating five pitch modulations per second. This relatively fast vibrato contributes to a more expressive and flexible timbral quality, enhancing dynamic variation and emotional coloration. Moreover, the vibrato extent in pop singing reached approximately 59.82 Hz, demonstrating a broader pitch modulation range. This deeper vibrato adds richness and emotional layering to the vocal output, making it particularly effective in heightened emotional scenarios—such as character portrayals involving tension, intimacy, or passion in pop-infused musical theater productions.

In contrast, musical theater singing exhibited a slightly slower vibrato rate of approximately 4.00 Hz, with more rhythmically even and consistent modulation patterns. This slower vibrato supports vocal stability and clarity, which are essential for ensuring that lyrics remain intelligible to audiences in large theatrical spaces. The vibrato extent was approximately 39.89 Hz, indicating a shallower frequency fluctuation. This restrained vibrato helps maintain pitch stability and prevents excessive tonal variation, thereby enhancing text delivery and tonal precision. Such a controlled pitch modulation strategy aligns with the technical demands of musical theater performance, where clarity of diction and melodic accuracy must be preserved for long-distance acoustic transmission.

As shown in *Figure 3*, the vibrato trajectories clearly illustrate the differences between the two styles. The blue curve, representing pop singing, displays dense peaks and wider oscillations, reflecting a high level of dynamic modulation. In contrast, the red curve, representing musical theater singing, shows fewer oscillations and a smoother contour, indicative of stable pitch control and speech-directed vocal technique.

Figure 3. Vibrato Rate and Depth in Pop and Musical Theater Singing



Note: Pop singing shows faster and deeper vibrato (~5 Hz, ~60 Hz), supporting expressive dynamics. Musical theater shows slower, shallower vibrato (~4 Hz, ~40 Hz), favoring pitch stability and textual clarity.

Source: Developed for this research

In summary, pop singing—characterized by faster vibrato rate and greater pitch fluctuation—offers broader expressive potential in both vocal layering and emotional articulation. It is particularly effective for conveying complex, subtle, or emotionally heightened content. In contrast, musical theater singing emphasizes pitch stability and intelligibility, making it more suitable for theatrical spaces and ensemble coordination. These differences in vibrato strategies reflect distinct acoustic adaptation mechanisms aligned with the performative objectives, audience reception modes, and spatial demands of each vocal style.

3.3 Application of Acoustic Features in Emotional Expression

The preceding acoustic parameter analysis revealed significant differences between pop and musical theater singing in terms of formant distribution, harmonic structure, and vibrato characteristics. To further clarify how these parameters translate into concrete emotional expression, this section draws upon the comparative acoustic indicators presented in Table 1 to examine their emotional strategies and functional implications within the context of Chinese original musical theater.

Table 1 summarizes quantitative differences across key dimensions—vibrato rate and extent, mean and stable fundamental frequency (F0), and the first formant (F1) of the vowel /a/. These core metrics provide a technical foundation for understanding how vocal styles relate to emotional function on stage.

Table 1. Summary of Acoustic Parameters in Pop and Musical Theater Singing Styles

Parameter	Pop Singing	Musical Theater Singing
Vibrato rate (Hz)	6.10	5.40
Vibrato extent (Hz)	7.53	16.04
Mean fundamental frequency (F0, Hz)	312.5	289.3
F0 stability (SD, Hz)	18.7	11.2
First formant F1 (/a/ vowel, Hz)	710.4	648.9

Note: Acoustic parameters were extracted from two performances of the same phrase. F0 stability refers to the standard deviation of pitch. Only F1 of the vowel /a/ was analyzed; F2, F3, and other vowels are reserved for future studies.

Source: Developed for this research

Specifically, with respect to formant distribution, the F1 value in pop singing is comparatively higher (710.4 Hz vs. 648.9 Hz), indicating a resonance focus in the mid-to-low frequency range. This configuration contributes to a warm, soft, and intimate vocal timbre, which is particularly effective in expressing subtle or restrained emotional content. Such a low-frequency resonance layout enhances emotional resonance in passages involving inner monologues or reflective transitions (Hakanpää et al., 2023). In contrast, musical theater singing emphasizes a more centralized and focused resonance strategy to enhance clarity and projection, aligning with the spatial and communicative demands of theatrical performance.

In terms of vibrato control, Table 1 shows that pop singing exhibits a higher vibrato rate (6.10 Hz) and a wider vibrato extent (7.53 Hz), resulting in more dynamic pitch modulation and heightened expressive intensity, particularly in high-arousal emotional scenes (Scherer et al., 2017). Musical theater singing, by contrast, maintains a more stable vibrato profile (5.4 Hz rate, 16.04 Hz extent), supporting vocal balance and speech intelligibility, especially suitable for ensemble sections or narrative-driven performance contexts (Becker & Peter J. Watson, 2025).

Additionally, the analysis of fundamental frequency (F0) reflects stylistic preferences: pop singing features a higher average F0 with greater variability (F0 SD = 18.7 Hz), which supports expressive pitch fluctuation and emotional engagement. In contrast, musical theater singing shows a more centralized F0 distribution (SD = 11.2 Hz), emphasizing rhythmic control and clarity in delivery.

Taken together, the widespread use of pop singing in Chinese original musicals is not only a result of its technical adaptability and expressive flexibility, but also of its strong acoustic alignment with multi-dimensional emotional expression. Through the nuanced manipulation of formant positioning and vibrato parameters, pop singing establishes a controllable and expressive vocal framework that enhances the emotional intensity and psychological depth of stage characters. Accordingly, acoustic features function not only as technical markers of stylistic difference but also as mechanisms that support artistic communication and stylistic selection.

It should be noted that this study extracted formant data only for the vowel /a/, excluding other vowel types and higher formants such as F2 and F3. Future studies are encouraged to expand the phonetic dataset to build a more comprehensive acoustic model of stylistic differentiation.

4. Discussion

By quantitatively analyzing key acoustic features of pop and musical theater singing, the present study directly responds to the research gap highlighted in the introduction—namely, the lack of systematic analysis concerning how vocal acoustic mechanisms shape emotional expression. Through the integrated examination of vibrato parameters, formant distribution, and harmonic structures, distinct functional divergences were identified in terms of emotional arousal and vocal regulation. These findings not only offer empirical evidence for understanding the expressive utility of pop singing in Chinese original musicals but also reaffirm earlier assumptions that such techniques enhance character portrayal. Furthermore, theoretical implications for vocal pedagogy, phrase composition, and stylistic integration in musical theater practice have been substantiated.

From an acoustic standpoint, the amplification of lower-frequency formants (F1 and F2) in pop singing has been shown to yield a timbral profile that is warmer and softer, facilitating more natural resonance pathways for conveying subtle emotional content. This observation supports prior findings indicating a positive correlation between low-frequency resonance and perceptions of tenderness or introspection (Hakanpää et al., 2023). Additional evidence has suggested that musically sensitive individuals demonstrate heightened performance in vocal

emotion recognition, largely attributable to greater responsiveness to timbral and frequency-based cues (Nussbaum et al., 2022).

In contrast, musical theater singing exhibited a higher overall formant configuration and a more focused spectral energy distribution, a structure that enhances intelligibility and projection within large-scale performance venues. This pattern aligns with the notion that high-frequency resonance plays a crucial role in maintaining clarity on stage (Müller et al., 2022). Empirical research has also shown that high-frequency energy serves as a salient cue in the neural processing of emotional categories (Paquette, 2018). Further evidence has demonstrated that pitch perception mediates the relationship between musical sophistication and emotional decoding accuracy, suggesting that structured vocal control—typical of musical theater techniques—may enhance emotional precision in performance (Vigl et al., 2024).

In the domain of vibrato, the analysis revealed that pop singing tends to employ faster pitch oscillations (~5 Hz) and wider frequency deviation (~60 Hz), whereas musical theater singing displays greater temporal consistency. These contrasting vibrato profiles reflect divergent affective strategies: the former is associated with dynamic, high-arousal expressivity; the latter with intelligibility and structural coherence. Prior studies have indicated that high-arousal emotions are commonly accompanied by elevated sound pressure levels and rapid frequency variation, while low-arousal emotions favor stable pitch regulation (Scherer et al., 2017). Additionally, recent findings have suggested that trained musicians are more adept at decoding emotions via pitch contour rather than timbre, reinforcing the relevance of musical theater singing in melody-driven emotional communication (Nussbaum et al., 2024).

From a perceptual perspective, prior work has shown that listeners often rely on vocal analogs of laughter, tremor, and roughness as key heuristics for identifying emotional content in sung expressions (Bedoya et al., 2021). This supports the interpretation that pop singing—through nuanced control of timbre and source parameters—enables broader emotional latitude. Moreover, stage-based audience studies have revealed that emotional intensity is more strongly judged through visual than auditory channels in theatrical settings, highlighting the need to integrate visual-aural coordination into vocal training curricula (Lange et al., 2022). Further supporting this, it has been observed that timbre plays a dominant role in affective recognition and that variations in timbral coloration—arising from differences in vocal technique—can directly alter listeners' interpretation of emotional intent (Nussbaum et al., 2022). These findings converge with the current acoustic data, suggesting that flexible resonance shaping and harmonic richness in pop singing create the technical conditions necessary for wide-spectrum emotional delivery and character construction in musical theater.

Despite its contributions, this study is not without limitations. First, the sample was drawn from a single excerpt of one musical, which may not sufficiently capture the full stylistic diversity

of the genre and may underrepresent variability between singing techniques (Paquette, 2018). Second, although high-resolution tools such as Praat and Origin were employed, emotional processing in real-world contexts involves complex psychological and situational dynamics, and acoustic markers alone may not capture the full range of perceptual outcomes (Bedoya et al., 2021). Future research should integrate subjective evaluations or neuroimaging techniques to verify the causal relationship between acoustic dimensions and emotional perception (Vigl et al., 2024). Moreover, this study did not control for gender, voice type, or age, all of which may influence acoustic outputs. Subsequent studies should adopt more diverse performer samples to enhance the generalizability and external validity of the findings.

In summary, the widespread adoption of pop singing in Chinese original musicals reflects not only its aesthetic appeal but also its technical adaptability. The flexible control over formant placement, vibrato modulation, and harmonic design enables multilevel emotional expression. By aligning experimental data with theoretical frameworks, this study contributes a data-informed foundation for vocal pedagogy, while offering new perspectives for localizing and globalizing musical theater within the Chinese cultural context.

5. Conclusion

Situated within the context of Chinese original musicals, this study examined the acoustic and emotional expression mechanisms of pop singing, using quantitative acoustic analysis to systematically compare its core features—formant distribution, harmonic structure, and vibrato parameters—with those of traditional musical theater singing. Beyond identifying physical-acoustic divergences in vocal production strategies, the findings also integrated theoretical models and perceptual psychology to demonstrate the influence of these differences on audience emotional reception and character portrayal.

At a macro level, this study responds to the growing demand for scientific acoustic evaluation and theoretical support amid the stylistic diversification of singing in Chinese musical theater. As a highly integrated performative system, musical theater relies not only on performers' training and experience, but also on the structural modulation of sound as an expressive medium. This research provides a conceptual model for understanding the fit between singing style and emotional function, offering a framework for differentiated vocal pedagogy and guiding early-stage decisions in musical composition and staging.

At a micro level, the study highlights the expressive advantages of pop singing in specific dramatic contexts. Through flexible formant shaping and dynamic vibrato features, pop singing constructs a timbral architecture rich in emotional tension—especially effective for portraying

emotional transitions, psychological struggles, and internal monologues. In contrast, the stability and high-frequency clarity of musical theater singing is better suited for passages that prioritize semantic clarity and narrative coherence. Rather than being oppositional, the two vocal styles represent complementary expressive strategies that together form the pluralistic vocal aesthetics of contemporary Chinese musical theater.

While the study yields valuable preliminary insights, limitations remain in terms of sample scope and methodological depth. Future research should include a broader range of productions and vocalists, and incorporate audience-based perception studies and theater-specific acoustic measurements to further validate and expand upon these findings. Ultimately, the use of pop singing in Chinese original musicals reflects not only a technical evolution, but an ongoing negotiation between artistic language and cultural context. This study aims to interpret this phenomenon from an acoustic perspective, contributing quantifiable and verifiable methods and frameworks to the future creation, pedagogy, and scholarly analysis of musical theater.

References

- Becker, A. S. & Peter J. Watson. (2025). The Use of Vibrato in Belt and Legit Styles of Singing in Professional Female Musical-Theater Performers. *Journal of Voice*, 39(1), 280.e7-280.e14. <https://doi.org/10.1016/j.jvoice.2022.07.018>
- Bedoya, D., Arias, P., Rachman, L., Liuni, M., Canonne, C., Goupil, L., & Aucouturier, J.-J. (2021). Even violins can cry: Specifically vocal emotional behaviours also drive the perception of emotions in non-vocal music. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1840). <https://doi.org/10.1098/rstb.2020.0396>
- Feng, Z. (2024). Analysis of the Business Model of Chinese Musical Theater Based on Community Economy. *Journal of Applied Economics and Policy Studies*, 12(1), 77–82. <https://doi.org/10.54254/2977-5701/12/2024114>
- Fleischer, M., Rummel, S., Stritt, F., Fischer, J., Bock, M., Echternach, M., Richter, B., & Traser, L. (2022). Voice efficiency for different voice qualities combining experimentally derived sound signals and numerical modeling of the vocal tract. *Frontiers in Physiology*, 13, 1081622. <https://doi.org/10.3389/fphys.2022.1081622>
- Hakanpää, T., Waaramaa, T., & Laukkanen, A.-M. (2023). Training the Vocal Expression of Emotions in Singing: Effects of Including Acoustic Research-Based Elements in the Regular Singing Training of Acting Students. *Journal of Voice*, 37(2), 293.e7-293.e23. <https://doi.org/10.1016/j.jvoice.2020.12.032>
- Lange, E. B., Fänderich, J., & Grimm, H. (2022). Multisensory integration of musical emotion perception in singing. *Psychological Research*, 86(7), 2099–2114. <https://doi.org/10.1007/s00426-021-01637-9>
- Lee, Y., Oya, M., Kaburagi, T., Hidaka, S., & Nakagawa, T. (2023). Differences Among Mixed, Chest, and Falsetto Registers: A Multiparametric Study. *Journal of Voice*, 37(2), 298.e11-298.e29. <https://doi.org/10.1016/j.jvoice.2020.12.028>
- Müller, M., Wang, Z., Caffier, F., & Caffier, P. P. (2022). New objective timbre parameters for classification of voice type and fach in professional opera singers. *Scientific Reports*, 12(1), 17921. <https://doi.org/10.1038/s41598-022-22821-w>
- Nussbaum, C., Schirmer, A., & Schweinberger, S. R. (2022). Contributions of fundamental frequency and timbre to vocal emotion perception and their electrophysiological correlates. *Social Cognitive and Affective Neuroscience*, 17(12), 1145–1154. <https://doi.org/10.1093/scan/nsac033>
- Nussbaum, C., Schirmer, A., & Schweinberger, S. R. (2024). Musicality – Tuned to the melody of vocal emotions. *British Journal of Psychology*, 115(2), 206–225. <https://doi.org/10.1111/bjop.12684>
- Paquette, S. (2018). Decoding musical and vocal emotions. *The Journal of the Acoustical Society of America*, 144(3_Supplement), 1841–1841. <https://doi.org/10.1121/1.5068112>
- Scherer, K. R., Sundberg, J., Fantini, B., Trznadel, S., & Eyben, F. (2017). The expression of emotion in the singing voice: Acoustic patterns in vocal performance. *The Journal of the Acoustical Society of America*, 142(4), 1805–1815. <https://doi.org/10.1121/1.5002886>
- Uloza, V., Pribušis, K., Ulozaite-Staniene, N., Petrauskas, T., Damaševičius, R., & Maskeliūnas, R. (2023).

Accuracy Analysis of the Multiparametric Acoustic Voice Indices, the VWI, AVQI, ABI, and DSI Measures, in Differentiating between Normal and Dysphonic Voices. *Journal of Clinical Medicine*, *13*(1), 99. <https://doi.org/10.3390/jcm13010099>

Vigl, J., Talamini, F., Strauss, H., & Zentner, M. (2024). Prosodic discrimination skills mediate the association between musical aptitude and vocal emotion recognition ability. *Scientific Reports*, *14*(1). <https://doi.org/10.1038/s41598-024-66889-y>