Pathways for AI-Enabled Phonation Training of Vocal Music in Higher Education from the Perspective of the Dynamics of the Singing Voice

Fei Xu

School of Music and Drama, Sias University, Xinzheng, Henan Province, China 451100

Abstract: The dynamics of the singing voice is an applied research perspective on the dynamics system of vocal organs based on the complete theoretical framework of Voice Aerodynamics, by referring to the core viewpoints of Meribeth Bunch's "Dynamics of the Singing Voice". It focuses on explaining the mutual influence and synergistic effects among various factors such as breath control, vocal cords, resonance, and psychological control during the singing process. It is a crucial analytical model that analyzes the dynamics sources of the singing voice, and its balance & regulation, providing fundamental theoretical support for scientific Vocal Music training. AI has the immediate perception ability, enabling to quantify data and achieve dynamic feedback. This paper analyzes the profound reasons for integrating AI technology into Vocal Music dynamics training, explores multi-dimensional training operation paths, and identifies the important matters that need to be noted during the execution process, thereby constructing a new paradigm for human-AI collaborative phonation training of Vocal Music, expanding the boundaries of college Vocal Music teaching, promoting the shift of Vocal Music teaching from experiential perception to dynamic quantitative regulation, and providing new insights for the improvement and development of college Vocal Music teaching.

Keywords: Dynamics of the Singing Voice; Artificial Intelligence (AI); Vocal Music in Higher Education; Phonation Training; Human-AI Collaboration

DOI:10.12417/3029-2328.26.07.016

1.Introduction

Vocal art is a dynamic artistic form that uses vocal organs as its carrier. Its core lies in the synergistic operation of the vocal organs and the human bodies. The dynamics of the singing voice is precisely the core perspective for systematically analyzing the dynamic operation mechanisms. From the perspective of dynamics, it comprehensively explains the internal relations among the breath management, vocal-cord vibration, and resonance control, providing important theoretical basis for scientific phonation training of vocal music in higher education. The rapid development of AI technology has provided strong technical support for the visual presentation of phonation dynamics, the dynamic simulation in training processes, and the precision of teaching guidance. Leveraging the theory of the singing voice dynamics to achieve the deep integration of AI technology and phonation training in Vocal Music can break through the temporal-spatial limitations and the sensory restrictions of traditional Vocal Music teaching, restructure the college Vocal Music phonation training system, and promote the scientific and precision development of Vocal Music teaching. It is of significant practical significance and research value for improving the cultivation quality of vocal music talents in higher education in the new era.

2.The Concept of the Dynamics of the Singing Voice, and the Logical Framework of Phonation Training of Vocal Music in Higher Education

2.1 *Core Concept Definition and Academic Positioning*

The dynamics of the singing voice referred to in this paper is not a secondary discipline established separately in the field of music or voice science, but an applied interdisciplinary research perspective and theoretical analysis framework specifically designed for Vocal Music teaching practice. It originated from the ideas embodied in the translation of "Dynamics of the Singing Voice" created by Meribeth Bonch, and grew its theory based on Vocal Aerodynamics, a mature discipline. It is the process of transforming and implementing the theory of Voice Science into Vocal Music teaching scenes. Unlike the Voice Aerodynamics with a complete research system and clinical experimental support, the dynamics of singing voice abandons the purely physicochemical and clinical research

directions, anchors in the practical scenarios of vocal performance, focuses on analyzing the dynamic integration of various factors such as respiratory airflow, glottic function, vocal tract resonance, and psychological regulation during the singing process. It mainly pays attention to how the phonation dynamics are generated, how the various factors maintain balance with each other, and how to carry out dynamic adjustment. It is specifically used to explain the dynamics operation of vocal organs throughout the singing process, and is applicable to the teaching research of phonation training of Vocal Music in higher education. From the perspective of academic positioning, the dynamics of the singing voice has been incorporated into the analysis framework of Vocal Music teaching research in China's core music journals. It is mainly used for conducting attribution study on the phonation issues or study on the establishment of dynamics training systems. Among them, the core theory of dynamic balance among airflow, vocal cords, and the cavity is highly aligned with the data collection and quantitative analysis requirements of AI-based Vocal Music training. It is an appropriate theoretical tool for AI to assist in the scientific Vocal Music training [1].

2.2 *The Limitations in the Traditional Phonation Training of Vocal Music from the Perspective of the Dynamics of the Singing Voice*

It is possible to find that the traditional teaching methods suffer from several unavoidable structural limitations after introducing the theory of the singing voice dynamics into current phonation training of Vocal Music in higher education and conducting a review of this process. Currently, the phonation training of Vocal Music in higher education mainly relies on face-to-face demonstration teaching by teachers. Teachers leverage their years of teaching experience to assess students' phonation status through auditory judgment and observation of their external body movements. This evaluation method can only make judgment on superficial timbre issues and obvious phonation errors, struggling to accurately capture the imperceptible dynamics information such as the changes in student's internal breath pressure, the state of vocal cord closure, and the adjustment of the vocal tract tension [2]. For students, core concepts, such as breath support, throat broadening, and cavity resonance, are all abstract and experiential descriptions. Students can only try to fathom them based on their subjective sensation. It easily leads to problems such as compensatory muscle activation and imbalanced dynamics distribution. Over time, these problems will develop into rigid wrong phonation dynamics mode, making it difficult to fundamentally improve the phonation quality. Furthermore, traditional classroom teaching is enslaved to the conditions of large-scale teaching. Teachers are challenging to keep an eye on the whole training process of every student, and the training feedback is significantly delayed. The wrong phonation methods cannot be corrected in a timely manner.

3.The Adaptation Mechanism and Value Dimensions of AI Involvement in Phonation Training

3.1 *The Adaptation Mechanism of AI Technology to the Singing Voice Dynamics System*

The core research objective of the singing voice dynamics lies in accurately capturing, dynamically analyzing, and scientifically regulating the phonation dynamics process. The signal acquisition, feature extraction, dynamic modeling, and real-time feedback capabilities of AI technology precisely align with this objective, providing crucial technical support for the practical application of the theory of the singing voice dynamics. The operation process of phonation dynamics is characterized by crypticity, dynamicity and instantaneity. Traditional teaching methods are challenging to conduct quantitative detection of it. AI can leverage technologies, such as high-precision acoustic sensors, dynamic signal capture, and human movement feature recognition, to convert the invisible and imperceptible air flow, vocal cord vibration, and cavity deformation and other dynamics mechanisms into visual, quantifiable, and analyzable digital signals. These signals precisely correspond to the three core research categories of the dynamics of the singing voice. The Vocal Music dynamics models constructed based on deep learning algorithms can establish an accurate mapping relationship between acoustic features, movement features of vocal organs and phonation dynamics state, thereby transcending the limitations of traditional vocal music that can only identify timbre, and deeply analyzing the dynamics mechanism and deviations in the phonation process.

3.2 *The Intrinsic Value of AI-Enabled Phonation Training of Vocal Music in Higher Education*

The deep integration of AI technology and the phonation training of Vocal Music in higher education effectively breaks down the barriers between the theory of the singing voice dynamics and teaching practice. It holds significant value in multiple aspects such as teaching application, talent cultivation, and discipline development. Regarding talent cultivation, AI has broken through the temporal-spatial limitations of classroom teaching, enabling to provide students with timely and normalized guidance for self-training, and promptly put compensatory muscle activation and imbalanced dynamics distribution right, continuously optimize students' phonation dynamics habits, and make up for the deficiencies of limited classroom training time and insufficient personalized guidance. It is conducive to students establish a stable and scientific phonation dynamics system. Regarding discipline development, the empowerment of AI has advanced a paradigm shift in the Vocal Music teaching research, shifting it from a paradigm of experience summary and subjective judgment to a scientific research paradigm of quantitative analysis, process research, and precise control. It has vastly expanded the application scope of the theory of the singing voice dynamics, injected new energy into the optimization and upgrading of college Vocal Music teaching systems, and promoted the continuous development of the Vocal Music discipline towards digitalization, scientization, and precision.

4.Implementation Pathways for AI-Enabled Phonation Training of Vocal Music in Higher Education from the Perspective of the Dynamics of the Singing Voice

4.1 *Building Dynamics-Tiered AI-Enabled Phonation Training Modules*

Based on the hierarchical logic of the singing voice dynamics from basic dynamics to system synergy, integrating with the progressive objective teaching law of phonation training of Vocal Music in higher education, tiered, progressive and synergistic AI-based specific phonation training modules can be constructed to form a systematic and gradient digital training framework. The training modules are strictly stratified according to the law of dynamics construction. The foundational tier focuses on core respiratory dynamics training, and leverages AI's signal collection capability to capture key dynamics parameters such as breath flow velocity, airflow pressure, and breath stability, and monitor the activation status of the thoracic diaphragm and intercostal muscle groups, aiming to guide students to form a uniform, continuous, and controllable airflow supply method, thereby solidify the dynamics foundations for the phonation in Vocal Music. The intermediate tier focuses on the vocal cord vibration dynamics training, and leverages AI's acoustic waveform analysis capability to accurately identify common vocal cord dynamics problems such as incomplete vocal cord closure, excessive vocal cord compression, and unbalanced vocal cord vibration, aiming to help students adjust the tension of the vocal cords and establish a correct and stable vibration-based phonation mode [3]. The advanced tier focuses on the resonance coupling collaborative training, and relies on AI's resonance peak monitoring and vocal tract morphology recognition capabilities to coordinate the ratio relationship among breath dynamics, vocal cord vibration and cavity resonance, aiming to optimize the vocal tract shaping ability, and realize the synergistic operation of the multi-organ dynamics system.

4.2 *Establishing a Human-AI Collaborative Dynamic Feedback Training Mechanism*

Based on the essential characteristics of vocal art that integrates humanities and science, a teacher-led and AI-assisted bidirectional collaborative feedback training mechanism should be established to clearly define the functional boundaries of humans and AI, and achieve the organic integration of the quantitative advantages of technology and the expertise of art teaching. In this mechanism, AI undertakes the fundamental, quantitative and real-time technical tasks. It continuously collects the dynamics signals, acoustic features and body movement data during students' phonation process, instantaneously generates visualized dynamics models to clearly present the deviations between current phonation dynamics state of students and the standard dynamics models, precisely mark the issues of force application in breath, vocal cords, resonance and other dimensions, and achieve visual and dynamic monitoring of the whole training process. Vocal Music teachers assume the leading guidance responsibility

and artistic gatekeeping responsibility based on their own professional aesthetic judgment and teaching experience. They provide professional explanations for the quantitative data and dynamic deviation results generated by AI, and integrate student's individual differences such as voice features, physiological conditions, and learning foundation to deeply analyze the underlying reasons for the imbalance of dynamics in order to avoid falling into the trap of mechanical data-based judgment, thereby forming a complete closed-loop training ecosystem: Students undergo normalized phonation training, AI captures real-time feedback, and teachers provide precise diagnosis and guidance.

4.3 *Establishing a Personalized Vocal Music Dynamics Training Archives System*

Relying on the data storage, feature tracking and intelligent analysis capabilities of AI, integrating the core evaluation indicators of the dynamics of the singing voice, a personalized Vocal Music dynamics training archive system can be established for students majoring in Vocal Music in higher education to realize precise and differentiated training. This training archive takes the phonation dynamics system as its core architecture to comprehensively record key data such as breath dynamics, vocal cord vibration, resonance coupling of students at different training phases, in different vocal ranges, for different pieces of music, and under different phonation state, completely preserve the development trajectory and evolutionary history of students' phonation dynamics habits, and intuitively present the training advantages and long-standing dynamics deficiencies of students. Through the comparison and intelligent analysis of big data, AI can accurately identify the repetitive problems, such as compensatory muscle activation, imbalanced dynamics distribution, and deviations in cavity regulation and then forecast potential difficulties and improvement directions during training.

5.The Practical Adjustment and Development Trend of AI-Enabled Phonation Training of Vocal Music in Higher Education

5.1 *The Control of Artistic Expression in the Human-AI Integrated Training*

AI technology provides precision support for the scientific training of phonation dynamics of Vocal Music. However, Vocal Music itself integrates humanities and science. The specifications of dynamics are the foundation, while artistic expression is the ultimate goal. Neither should be overemphasized. In the practical operations of human-AI collaborative training, it is essential to strike the right balance between technical quantification and artistic expression so as to avoid falling into the mechanized training pitfalls caused by technical tools. AI can only conduct objective evaluations of the phonation dynamics and acoustic parameters. It is difficult for it to perceive the emotional connotation, artistic conception, and style characteristics embodied in vocal music works, nor can it make aesthetic judgment on the artistic appeal and emotional expression of singing. Therefore, in the training practice, it is a must to adhere to the humanistic core of vocal art and refrain from using AI-generated dynamics data as the sole evaluation criterion, prevent students from stiff phonation and empty expression led by mechanical imitation of standardized dynamics models.

5.2 *Iterative Optimization Pathways for the AI-Based Vocal Music Dynamics Model*

Current AI dynamics models aligned with Vocal Music training still suffer from certain application limitations, mainly manifested in insufficient model adaptability, limited precision in identifying subtle dynamics, and difficulty in fully adapting to individual phonation characteristics of learners. Human body vocal organs exhibit significant individual differences. The optimal phonation dynamics ratios vary among learners with different voice parts, different physiological structure and different singing styles. General models are challenging to achieve refined personalized training. In response to this issue, it is necessary to establish a normalized model iteration and optimization mechanism, continuously collect diverse phonation dynamics samples from professional vocalists and outstanding Vocal Music learners, expand the types of voice parts, styles and vocal ranges in sample libraries, and construct a comprehensive and finely divided personalized dynamics benchmark library. Concurrently, it is a must to use the core principles of the dynamics of singing voice as the underlying constraints to optimize the algorithm logic, strengthen the correlation analysis between acoustic signals and the dynamics state of vocal organs, reduce the

weight of pure reliance on timbre recognition, and enhance models' acuity to identify subtle dynamics deviations and potential force application issues [4].

5.3 *The Cultivation of Digital Teaching Capabilities of College Vocal Music Teachers*

After the deep integration of AI technology and the phonation training in Vocal Music, the teaching models of Vocal Music in higher education has been reconstructed, Vocal Music teachers are also faced with new requirements regarding their comprehensive teaching capabilities. It has been difficult for the traditional monolithic professional vocal music capabilities to align with the needs of digital teaching scenarios. In the new era, college Vocal Music teachers not only need to deeply study the Vocal Music theories and singing practice, and master the core principles and training logic of the dynamics of the singing voice expertly, but also need to possess compound capabilities such as the application of digital teaching tools, dynamics data analysis, and human-AI collaborative teaching planning. Higher education institutions should establish a systematic faculty digital capability cultivation system, regularly conduct special teaching and research training, and organize symposiums and skills training on the perspective of the dynamics of the singing voice, practical operations of AI-based vocal music training tools, dynamics data analysis and interpretation, digital classroom design, and other content to help teachers update their teaching philosophy and break through traditional teaching mindset [5]. Concurrently, they should build inter-disciplinary communication platforms, facilitate in-depth cooperation between Vocal Music teachers and researchers in AI, acoustic engineering, digital education and other sectors, explore new pathways for digital Vocal Music teaching, promote teachers to actively apply AI technology in daily phonation training, continuously enhance digital teaching literacy and human-AI collaborative teaching capabilities, and provide solid talent safeguards for the implementation and long-term development of AI-enabled Vocal Music teaching.

6.Conclusion

Guided by the theory of the singing voice dynamics, integrating AI technology into the phonation training of Vocal Music in higher education serves as an important path for the scientific and digital innovative development of Vocal Music teaching in the new era. It effectively solves the problems of strong subjectivity, obvious lag, and prominent homogenization in the traditional phonation training of Vocal Music. With its advantages of precise dynamics perception and quantitative feedback, AI technology enables the hidden phonation dynamics process to become visible and controllable. It complements with the subjective teaching guidance of teachers to build a new training system featuring tiered training, dynamic feedback, and precision education. In the future, human-AI collaboration will become the main development direction of the phonation training of Vocal Music in higher education. With the continuous deepening of the theory of the singing voice dynamics and the continuous iteration of AI technology, college Vocal Music teaching is expected to achieve standardized dynamics training, precise teaching guidance, and personalized talent cultivation, thereby continuously improving the quality of college Vocal Music education systems in the new era.

References:

[1] *Yuanyuan Zheng.(2025)Construction of the Blended Teaching Model for Vocal Music in Higher Education under the Background of Smart Education[J].Journal of Heze University,47(6),128-135.*

[2] *Shupeng Zheng.(2025)Optimization Strategies for Vocal Music Teaching under the Context of Digital Transformation[J].China National Exhibition,18,163-165.*

[3] *Ping Xiao.(2025)Curriculum Design and Implementation of Vocal Music Talent Training in Higher Education Based on Artificial Intelligence[J].Modern Music,8,172-174.*

[4] *Xiaomin Chen.(2025)Reform of College Vocal Music Teaching in the New Era[J].Music Life,8,90-93.*

[5] *Sudu Langteng.(2025)The Integration of Traditional and Modern Teaching Methods in College Vocal Music Teaching Reform[J].Art Panorama,13,157-159.*