A Study on the Application Effects of Intelligent Practice Companion Systems in Vocal Music Teaching

Gong Qing

Xiamen Huaxia University, Xiamen, Fujian 361000, China.

DOI: 10.64649/yh.eg.issn3078-4794.20260516

Abstract: The current problems of post-class vocal training in higher education institutions is that feedback is delayed and specific vocal errors cannot be identified. The research subject of this paper is the application value of an intelligent coaching system based on pure fifths and dotted-eighth notes. We have built a training pipeline to collect audio, identify any deviations in the sound, provide prompt correction, make targeted corrections, and have also developed an pre- and post-test assessment system. A comprehensive index system has been developed that includes pitch deviation, rhythm errors, the overall assessment of teachers, how accurately one can correct oneself, and how long the training takes. The above framework provides a general basis for determining when to repeat the performance and can also be used to adjust the difficulty of the training tasks; that is, they can be increased or decreased.

Keywords: intelligent accompaniment system; vocal music teaching; acoustic recognition; targeted error correction; teaching effectiveness

0.Introduction

Vocal Pedagogy in higher education institutions has two main purposes: skill training and cultivation of artistic expression. Although students are taught in class, they still need to practice after class to improve their pitch accuracy, rhythm and vocal control. The old way of self-study and self-practice relies on auditory memory and repeated singing; it cannot detect particular errors or the extent of performance fluctuations in real time, and thus is difficult to achieve specific training goals accurately. Now that acoustic analysis and real-time feedback technology have been combined, we can see any deviations in performance in real time; this data can help students correct themselves promptly and allow teachers to conduct professional assessments of vocal performance and training progress.

1. Application Requirements for Intelligent Practice Companion Systems in Vocal Music Teaching

1.1 Delayed feedback on post-class singing exercises

Training for improving the voice is mainly based on hearing and controlling one´s own voice. Although the teacher finds the problem with singing during class, students are unable to correct it on their own in the days that follow. When students are given scales, intervals and musical phrases to practise alone, they often fail to notice any errors in their singing; for instance, if the target pitch is too low, they may not be aware of it promptly and repeatedly make the same mistake. In the following classes, the teacher will have to go over these problems and thus reduce the time for teaching skills. Late feedback after class does not allow us to correct errors in practice in time, and thus it will have an adverse effect on the overall results of the training [1].

1.2 Variations in Guidance for Individual Vocal Training

Some students are unable to consistently reach the high note of ascending pure fifths; thus, they need to strengthen their interval recognition skills and the smooth register transitions. Although some students can perform the melodic sections correctly, there are still problems with the timing of dotted notes and syncopated rhythms; therefore, regular training in metric subdivision

needs to be carried out. The range of notes, speed and repetition of practice are all the same in the post-class exercises; strong-foundation students will perform repeated low-intensity drills, and those with weaker foundations will be continually exposed to music beyond their current ability. Teachers develop training content according to the short time available in the classroom; however, this method often fails to address all kinds of long-term vocal errors and thus does not match students' actual proficiency levels.

1.3 Insufficient support for students' self-correction

When students find that their performance is abnormal, they will identify the error, determine the cause, modify their study methods, check the results, and thus correct the mistake. In actual training, students may feel that their high notes are unstable and are unable to determine why; thus, they will often repeat an entire section of music without learning. As there are no original recordings and no playback data, we cannot objectively compare the changes in pitch during the various practice sessions to determine whether a single correct performance was due to chance or the result of sustained skill. As the number of training sessions has increased, the same problems keep happening; without clear standards for identifying these errors, improvement is unlikely.

2. Practical Application of Intelligent Practice Partner Systems in Vocal Music Teaching

2.1 Vocal Deviation Diagnosis Based on Acoustic Recognition

A four-bar vocal training exercise from the foundational vocal course for music education majors at higher education institutions is taken as the case study in this paper. The instructor sets the tempo to 72 beats per minute in 4/4 time, adds an ascending pure fifth leap in the third bar, and arranges a rhythmic combination of a dotted quarter note and an eighth note in the fourth bar; then, according to the students' vocal parts, the instructor changes the starting pitch and enters the standard melody and rhythm information into an intelligent coaching system that has been equipped with a pitch-tracking and rhythm-localization module. For the recording stage, a quiet place far from noise and reverberation is chosen, and the microphone is set about 20cm away; to avoid external sound interference, headphones are used to play back the accompanying music. After a student completes their first performance, the system will capture the median fundamental frequency of stable vocal segments, compare it with the target pitch, convert the pitch deviation into a pitch score, and use the onset time of each note to calculate how far away it is from the standard beat pattern for rhythm detection. The teacher focuses on the landing note of the leap in the third bar and the entry point of the short note in the fourth bar, points out problems such as consistently low pitch or premature rhythm onset directly in the relevant score, and then has the student perform vibrato, glissando, and sections with abnormal pitch recognition again based on the original recording for acoustic misjudgment review by the teacher. Finally, the system will keep the bad notes, the misplaced items and the recording file for later review. Research on intelligent vocal assessment to date has used tools such as Praat, Sing&See and Vocal Pitch Monitor for acoustic analysis; although it has shown that acoustic analysis can help with vocal training, problems still exist in the evaluation of timbre, recognition and assessment of emotions, etc [2].

2.2 Targeted error-correction training leveraging real-time feedback

To address the problem of the target note being too low in the pure fifth leap of the third subsection, the teacher divided the training content into four stages: auditory discrimination of the target note, single-note singing, interval linking and complete vocalisation. First, the students heard the auditory signals for the beginning and end notes, finished the single-note singing exercise, and then started slow leap training at a speed of 60 beats per minute. At the same time, show the target note pitch trajectory and have the instructor identify any pitch deviations in the stable note-holding section. A reference threshold of ±0.50 yuan was set in this paper for the target note deviation of classroom training. If a student meets the target conditions in all

three rounds of independent vocalisation and can speak naturally, their training speed will be increased to 72 beats per minute. If the pitch deviation is repeatedly above the upper limit, we will narrow the range of the interval and reassess both the quality of auditory discrimination for the target note and the smoothness of transitions among vocal registers to avoid an excessive increase in the number of vocalisation attempts. In the fourth sub-section of the dotted rhythm, at a training speed of 72 beats per minute, the instructor pointed out that the theoretical duration of a dotted quarter note is 1.250 seconds and that of a quarter note is 0.417 seconds. First, have the students break down the rhythmic unit by clapping together according to a metronome; then, add a singing task for the melody. The actual start time of each note was regularly compared with the predetermined target time points, and the first training period had a tolerance of ±100ms. After successfully completing a set number of accurate vocalisations, students were able to perform whole four-bar phrases and confirmed that the local corrective improvements could be maintained cohesively in the music. Before the start of the official training, the teachers will hold a trial test to establish a baseline for students´ vocal foundations and will not directly use the same numerical standard for evaluating all students´ singing at that time. Instructors have been observing the students´ breath support, tone quality and laryngeal effort during training, and if there are any signs of vocal tension, have reduced the amount of training promptly. Research on the "dual-track integrated" vocal pedagogy in higher education institutions has shown that the application of digital and intelligent technologies should be based on the professional characteristics of vocal courses, improve cooperation between instructors and intelligent technologies, and be cautious of pedagogical evaluation risks of technology application [3].

2.3 Personalized training adjustments based on training records

Each week, the teacher will gather data on how well the system is being used by observing playback, effective time, target pitch deviation, rhythm errors, compliance with repetitions, and so on, to keep a running record of training. For students who repeatedly show a pitch deviation in the leap-to-landing notes of the third measure, the teacher will continue to give them the target-pitch auditory discrimination task and slow-paced interval connection exercises but will not increase the range of their pitch at this time. If the student can maintain the same pitch but consistently plays the rhythm of the fourth measure too early, the teacher will stop having the student repeat scales and will add exercises that use dotted rhythms and changes in tempo. Adjustments to the training will be made based on the performance evaluation results. After a student can accurately reproduce a single repeated-note section at the same speed, hits all the notes and rhythms properly, and has received praise for good singing habits from the teacher, they can gradually add a high or low note. If the student does not meet the above standards for two consecutive weeks, the instructor will adjust the training difficulty and determine the reasons for the errors; then, either slow down the speed or divide up the practice into smaller parts. For each adjustment cycle, the instructor records the original problems, the training content and the results of re-assessment, and then plans the next phase of training to address specific vocal deviations; thus, the number of practice sessions is not used as the sole basis for increasing the difficulty of the training.

3. Evaluation of the Application Effects of the Intelligent Practice Companion System in Vocal Music Teaching

3.1 Teaching Effectiveness Evaluation Indicators and Control Design

For the purpose of this teaching effectiveness study, students enrolled in the basic vocal music course of the university music education program were selected as the subjects. Based on the results of the pre-test, divide the participants into groups according to their vocal sections, and then randomly allocate some of these groups to the experimental group and others to the control group. The length of the training period was 8 weeks, and in the experimental group, both instructor guidance and intelligent assistance were used in the joint training model; the

control group received the same number of instructional hours from instructors and participated in traditional self-study. The total training time, repertoire difficulty and instructional needs in both groups were the same. Based on the pre- and post-training evaluations, the main metrics of assessment included the mean absolute pitch deviation, rhythmic duration error and general vocal performance score; auxiliary indicators were the accuracy rate of self-correction and the time taken to achieve the training goals [4]. Pitch- and rhythm-related data were obtained from the acoustic information of the performance, and three vocal teachers independently judged the general quality of the voice without knowing which group the participant was in. The four divisions of the scores were breath support, register change, articulation and expression. Participants had to learn some new musical phrases at the same level as the original ones to see if they could apply what they had been taught.

3.2 Comparison of pitch accuracy, rhythm, and overall singing ability

To determine the pitch, one finds the fundamental frequency of a stable part in the voice that corresponds to the desired note, and then calculates the mean absolute difference between this measured value and the standard pitch. A small value indicates that the pitch is close to the target. For rhythm evaluation, the absolute error of the onset time of each note from the target onset time is recorded, and special attention should be paid to verifying rhythmic continuity after pure fifth leaps and controlling the duration of dotted notes accurately. The total score for the comprehensive vocal performance evaluation is 100 points, and the assessing teachers will independently grade aspects such as breath support and voice control, pronunciation, accuracy of pitch and rhythm, and artistic expression; professional judgment will be given precedence over system scoring. The following data will show how much improvement has been achieved in practice; that is, the pre- and post-test means for pitch deviation, rhythm error and instructor ratings of both groups of participants, as well as the amount of improvement and differences between the two groups, will be shown with appropriate confidence intervals and effect sizes. If there are considerable improvements in the metrics of pitch and rhythm, but only slight changes in artistic expression, then system feedback should be added that explains the reasons based on measurable acoustic features; however, progress in individual skills should not be considered an improvement in overall vocal performance ability.

3.3 Analysis of Autonomous Error-Correction Ability and Practice Efficiency

In the autonomous error correction evaluation stage, short musical phrases that were not covered in the regular training were selected. Students had to, without the help of system prompts, find the location of the error, determine the type of deviation, and correct it within a limited number of rehearsal attempts. The total number of errors found was used to calculate the localisation accuracy rate, and the pitch and rhythm deviations after the rehearsals were used to determine the effect of error correction. The purpose of this protocol is to see how well students can self-correct without system prompts. To improve the efficiency of training, only one indicator was used: the actual time it took to train the students until they met the predetermined requirements for pitch and rhythm was recorded, and samples of students who did not meet this standard within a specific period were also collected separately. System logs were collected from the experimental group, and at the same time, audio materials and training logs with the same specifications were also obtained for the control group. Therefore, it was decided not to introduce bias in the results through different collection methods. In this paper, the autonomous error-correction accuracy rates and the time to meet the training conditions of the two groups are compared to determine how well students can learn unfamiliar short musical phrases. If the students had fewer recurring mistakes, met the requirements in a shorter period, and their overall scores did not change significantly, these results should be separated to distinguish between development in independent learning and improvement in special singing skills.

4. epilogue

Another way to teach with an intelligent coaching system is to obtain the results of acoustic diagnosis and generate a customised training plan; for example, according to the pitch contour, onset time and choral recording markers of a performance, particular exercises such as pure fifth-degree leaps and dotted-rhythm recognition can be set. Based on the above data, teachers can design many exercises to train all the above areas, such as differentiating target pitches, dividing rhythms and changing tempos, etc., and adjust the difficulty level dynamically according to the continuous compliance of students. The five areas of the performance assessment in this paper are pitch accuracy, rhythm precision, overall vocal performance score, self-correction accuracy rate and the time taken to reach proficiency goals. The several indicators will be given different weights so that no single one dominates the final score, and the teacher will also review the system feedback after it has been released. Since we have not obtained all of the original control data, we have not been able to verify whether this intelligent coaching system can effectively improve teaching and what conditions are suitable for its application.

About the Author

Gong Qing (born July 1990), female, Han ethnicity, from Xiamen, Fujian Province, holds a master´s degree and is a lecturer. Her research interests include musicology, musical performance, and music education.

References

[1] Li Chen. Technological Empowerment: A New Paradigm of Hybrid Teaching for the "Vocal Music" Course in Higher Education Institutions [J]. Drama Home, 2024(36):125-127.

[2] Cui Fengbin, Liu Xiaoyuan, Zheng Guodong. Research on the Application of Intelligent Voice Assessment in Vocal Performance Training [J]. Drama Home, 2026(24):132-135.

[3] Sun Na, Yuan Jie, Chen Dayang. "Dual-Track Integrated" Vocal Music Teaching in Higher Education Institutions in the Digital and Intelligent Era: Advantages, Challenges, and Countermeasures [J]. Journal of Beihua University (Social Sciences Edition), 2025,26(01):132-138+156.

[4] Zheng Yuanyuan. Exploration on Constructing a Hybrid Teaching Model for Vocal Music in Higher Education Institutions under the Context of Smart Education [J]. Journal of Heze University, 2025,47(06):128-135.