Journal of Theoretical and Applied Information Technology
31st March 2025. Vol.103. No.6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

# USING ARTIFICIAL INTELLIGENCE SOFTWARE FOR CORRECTING INCORRECT INTONATION IN SINGING

**XIAO HUIWEN**

Doctor of Arts, Senior Research Officer, Lecturer,

Sanming University, China

25 Jingdong Road, Sanming City, Fujian Province, China

E-mail: 091019591nb@gmail.com

## ABSTRACT

The purity of intonation has an impact on the expressiveness of vocal performance. The aim of this study is to determine the features of using artificial intelligence (AI) software for correcting incorrect intonation in singing. The research employed the following methods: analysis, comparison, calculations of two-factor analysis of variance (ANOVA), Cognitive Change Index (CCI). Training techniques included the use of the iZotope RX 10 application for the purpose of flexibility of performance based on the development of singing breathing. The Clip.audio application was used to work on ways of reproducing sounds; Adobe Audition – for creating a musical style. It was established that the advantages of digital applications are related to the possibility of automatic error detection and the development of individual vocal approaches that are important for students. The results showed that the learning process primarily contributed to the preservation of the correct intonation at different pitches, the development of correct pronunciation. The students developed a high level of vocal intonation during training, which affected the expressiveness of vocal performance. The practical significance of the article is aimed at the use of AI technologies for the development of vocal intonation. Artificial intelligence applications helped respondents achieve high vocal results, which also had a positive impact on the formation of vocal skills. Further research may focus on determining the ways of developing vocal intonation depending on the features of the musical genre.

**Keywords**: *Vocal Performance, Voice Mobility, Musical Articulation, Sound Image, Interactive Technologies*

## 1. INTRODUCTION

It is important for the vocalist to ensure a combination of technique and harmony when singing, which based on understanding the musical performance techniques. A vocalist must develop the vocal apparatus, performance aesthetics, and melodiousness. The development of correct intonation has an impact on these skills [1]. The use of AI makes it possible to achieve ease of sound during articulation. The foregoing gives grounds to state the relevance of the issue selected for the research.

The right vocal performance of a particular composition depends on the text and melody ratio, taking into account the expressiveness of individual sounds. Each sound should be expressive during singing, individual words should be pronounced correctly [2]. Correct pronunciation makes it possible for the audience to better perceive a piece of music and understand its meaning. The performance of sounds is characterized by length during singing, so it is worth considering the peculiarities of the combination of melody and lyrics. Chinese words are characterized by their own intonation, which must be harmoniously combined with the main melody. The Chinese phonetic system is characterized by a long distance between vowels and consonants, which must be taken into account when singing. Therefore, additional short words are often used to fill these gaps, which is related to the speed of performance [3]. It is important to maintain accuracy, purity and firmness of performance, which affects the creation of three-dimensional sound. Special attention should be paid to complex words, which is associated with the need to use mobile facial expressions, change the position of the oral cavity [4]. The performance of European compositions involves taking into account the absence of changes in stressed and unstressed syllables, which differs

from the peculiarities of the performance of Chinese music. Articulation is possible under the influence of adjusting the position of the lips, lower jaw, and tongue. A clear combination of words and melody during performance ensures the artistry of singing [5]. Achieving purity of intonation ensures the perception of musical images, increasing the relationship between hearing and voice.

The quality of correct pronunciation in compositions of different levels of complexity depends on the training level [6]. Incorrect pronunciation can be corrected during singing based on the use of AI software. AI is software that aims to develop students' professional and general abilities. Interactive learning affects the ability to perform complex tasks based on proposed algorithms [7; 8]. The advantages of AI in education are related to the use of structured approaches, which affects the improvement of the vocal abilities of each student. This is due to a decrease in the number of errors and an understanding of the correct approaches to vocalization. AI provides recognition of the quality of vocal performance of songs and helps to adjust the learning process [9]. AI affects the possibility of creating new educational materials, practical tasks and analysing the overall students' success.

After studying the theoretical material, a variety of approaches to the AI use in general for the production of a vocal voice was established. But the peculiarities of the development of vocal intonation with the help of AI have not been considered in detail separately. Attention is paid to the combination of vocal singing, articulation, breathing, emotionality, etc. The aim of the work involves taking into account the features of using AI software to correct incorrect intonation during singing.

The objectives of the study were to:

- Create educational techniques for the development of vocal intonation under the influence of the AI;

- Determine the advantages of AI obtained in the educational process (according to the students);

- Determine the skills that were formed in students during the educational process for the possibility of ensuring high-quality vocal intonation;

- Determine the level of students' vocal intonation, focusing on the general performance and the performance of a particular composition.

## 2. LITERATURE REVIEW

The use of automatic vocal pitch correction allows you to develop the skills of correct intonation. In this way, students can develop improvisation and harmonization skills based on Sumel's interactive system during their studies. The correction of incorrect intonation is connected with the transformation of musical layers, which ensures the adaptation of a piece of music to the capabilities of an individual voice. The neural network can increase the students' vocal efficiency by 93% [10]. Automated vocal intonation enhancement systems increase the accuracy and speed of intonation perception. Collecting data on the intonation of the performer's singing allows for the development of individualized learning algorithms. Correcting the intonation signal allows students to achieve emotionality in vocal performance [11]. The use of digital technologies affects the improvement of the teaching music, which is associated with the adjustment of educational methods. The analysis of audio and video materials affects the ability to track changes in the student's qualification level. The use of digital technologies affects the possibility of providing systematic and comprehensive approaches to learning, which affects a better perception by students. Digital technologies contribute to the transformation of abstract learning mechanisms into more understandable ones. This affects the accurate perception of one's own voice, which ensures the establishment of the relationship between different voices for the purpose of achieving musical harmony [12].

It is necessary to focus on the peculiarities of the musical genre to reproduce the necessary intonation. It was established that the most popular genres of music for Chinese students are folk and popular music, which involves the use of improvisational technologies. The development of intonation is possible as a result of ensuring the expressive performance of melodies, preserving the smoothness and mobility of the sound. The work with the text should also be carried out during performance, which contributes to the regulation of vowel and consonant sounds [13]. Digital technologies influence the understanding of the principles of vocal intonation through the use of combined technical approaches. Such a learning process involves the use of built-in microphones and cameras that provide audio signal reception. The edTech software affects the tracking of the quality of vocal performance in real time, the

correctness of the display of rhythm and intonation. Students also develop the skills of dynamic performance in different ranges, clarity of performance. So, students contribute to achieving high results in owning their own voice [14]. Vocal improvisation promotes understanding of respiratory, technical, and aesthetic vocal skills. The combination of these skills affects the display of quality intonation. Interactive technologies are a high-quality tool for achieving the presented skills, as they influence the systematic assessment of knowledge, the involvement of different approaches to learning. This promotes the achievement of a high level of knowledge in 88% of students [15].

Control over the singing voice affects the achievement of quality singing. The learning process requires not only the reflection of vocal parts, but also their perception in the performance of other vocalists. Based on the analysis of the vocal singing of other musicians, it is possible to achieve greater purity of performance. This is an effective tool for finding valuable vocal approaches [16]. During the educational process, it is necessary to focus on the analysis of musical compositions in different performances. This affects the understanding of different types of intonation (adapted and emphatic). It was found that most performers prefer adapted intonation, which is associated with finding their own approaches to performance in order to achieve the most expressed emotions. This affects the creation of emotional experiences, reproducing dramatic intonations. Influencing factors for creating adapted intonation are movies, voice type, etc. [17]. Modern virtual reality technologies contribute to the transformation of abstract acoustic principles and more intuitive approaches. This ensures the search for wider opportunities for students, which affects the improvement of vocal singing. The process has an impact on the formation of emotional experience [18]. Intonation assessment is an effective tool for quality vocal performance. Self-controlled learning can be implemented on the basis of neural networks. The process of combining audio recordings affects the assessment of individual tones and pitch intervals. The accuracy of intonation assessment using automatic neural networks is up to 93.38%, which contributes to the achievement of visual assessment results [19].

The literature review contributed to the identification of the advantages of digital technologies for the development of vocal skills and intonation in general. But the intonation development was considered from the perspective of vocal performance, which excludes an independent study of this component. The research gaps are also related to the lack of creating separate educational approaches to the AI adaptation in order to achieve clarity of vocal intonation.

## 3. METHODOLOGY

### 3.1. Research design

The first stage of the research involved the development of teaching methods for the possibility of setting vocal intonation. AI-based applications iZotope RX 10 [20], Clip.audio [21], and Adobe Audition [22] were used to enable classes. The second stage of the study included an assessment of the benefits of AI for students. The Staple scale was used for the assessment aimed at measuring the values of a separate indicator. The results were aimed at determining the advantages of AI not in general, but as a result of the specific programmes used in the study. The third stage of the study included the dynamics of changes in acquired musical skills and the level of intonation development. The level of developed intonation was determined based on the results were evaluated in general and during the performance of individual compositions.

### 3.2. Sampling

The study involved 170 students to determine the features of the development of vocal intonation with the help of AI. The students studied at the School of Education and Music, Sanming University. The sample included 2nd-3rd-year students who could use the acquired professional knowledge, but did not understand the intricacies of vocal singing. Students with a medium and above levels of vocal knowledge were involved in the study. Fourth-year students were excluded from the study, as their learning process involved taking into account all stages of vocal singing, including intonation. Therefore, the results would not be correct, which involved a combination of academic training and training proposed in the research. All students who participated in the research were in equal conditions.

### 3.3. Methods

The development of teaching methods for setting vocal intonation was connected with the use of a general theoretical method of analysis. The analysis involved identifying advantages and disadvantages that affect the development of vocal intonation [23; 24; 25]. The choice of teaching methods involved the determination of the most favourable methods

for training of intonation and the possibility of their implementation in the research. The choice of AI applications was related to the comparison of efficiency and complexity of their use. A comparison of the specific features of 20 AI technologies identified three programmes that most clearly meet the specified requirements. The training period lasted 4 months.

Determining the advantages of AI was related to their influence on the development of musical intonation. Therefore, the authors presented a list of possible advantages for students to compare their own results. The use of the Staple scale involved assigning points to the identified advantages from 1 to 7, where 7 reflected the most influential value. The data were collected through WhatsApp service, which enabled collecting the necessary information within five hours. To assess the correctness of the presented results, students had to justify the most and least influential advantages. The total scores were calculated for each individual advantage, which made it possible to calculate two-factor ANOVA [26]. Analysis of variance determines the indicators that have the greatest impact. The sum of squares of indicators, the level of variance and mathematical expectation are taken into account during calculations. The most significant influence of a single indicator will be revealed if the experimental value is greater than the table value.

The dynamics of changes in the cognitive component of students, which is associated with the acquisition of vocal skills, were determined through an assessment of their development during 4 months of study. The skills were assessed by teachers based on students' observation, the level of their performance of various compositions. The quality of performance of the compositions depending on the complexity, the possibility of conveying musical images and decorations was taken into account. Results were obtained from students in the control and experimental groups — assessment of student performance involved monitoring the capabilities of students formed during the educational process. Comparison of the initial results of students with the final ones made it possible to determine the overall performance of students. The assessment of student performance was valid since it provided for the consideration of additional objective factors. This involved considering the meaningful perception of theoretical knowledge and the correctness of their application in practice. Students' level of independence and the ability to use logical approaches to vocal interpretation were also taken into account. Students' vocal development level also provided for the consideration of originality, which contributed to high-quality and individual performance. The emotionality of singing was also one of the objective factors for assessing student performance. The results obtained were valid since the process involved comparing the initial opportunities of students and those achieved during the study period. The Cognitive Change Index, which was developed by the authors of the article, was used for calculations:

$$c_c = \frac{\sum p + \sum t + \sum i}{m(m-1)}, \quad (1)$$

$p$ – the possibility of using a particular skill during the performance of a song of different difficulty levels;

$t$ – understanding of theoretical information regarding the practical application of a particular skill;

$i$ – a score for the possibility of using individual skills during the creation of improvisation;

$m$ – the highest score that could correspond to the development of a particular skill.

The level of vocal intonation was determined when the skills acquired by the students were assessed. The assessment of the intonation of a particular composition involved the selection of a composition that was not considered during the learning process. Such a composition was the Chinese song The Sea, O My Hometown (Figure 1, Appendix 1), which was characterized by various stylistic forms of musical expressiveness.

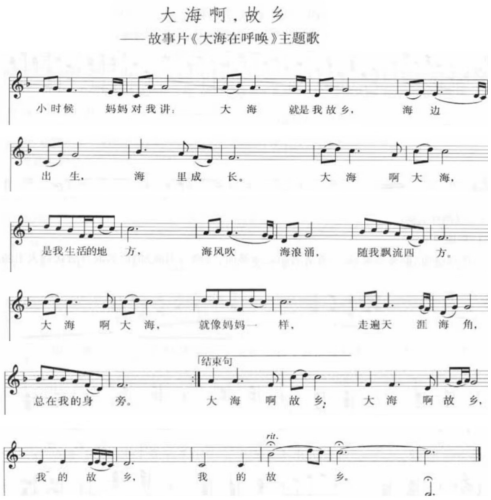

**Fig. 1.** Sheet music for the performance of the composition The Sea, O My Hometown (Music sonyselect, 2024 [27])

This made it possible to assess the level of development of vocal intonation on the basis of its adjustment to convey the artistic aspects of vocal expressiveness. This approach was aimed at assessing the level of knowledge acquired by students during training and as a result of performing an unfamiliar composition. The final results were obtained with the help of two-factor ANOVA calculations. The calculations were carried out by analogy with the previous stage of the study.

## 4. RESULTS

Singing should be based on achieving the correct pronunciation, which ensures the expressiveness of the performance. The authors developed appropriate training methods for the possibility of taking into account the positive or negative impact of AI on the process of setting vocal intonation. It was planned to use the following AI-based programmes during the educational process: iZotope RX 10 [28], Clip.audio [29], and Adobe Audition [30] (Figure 2).

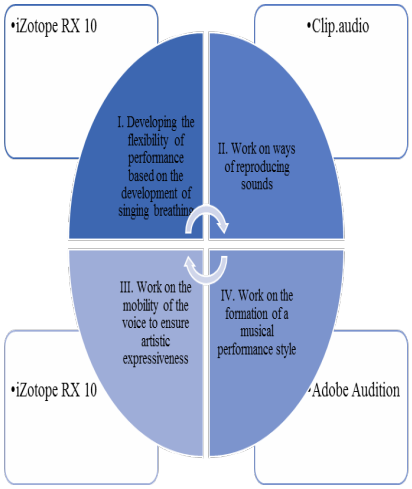


***Fig. 2.*** *Educational Techniques For The Development Of Vocal Intonation (Developed By The Author Based On The Analysis Of The Studies [31; 32; 33])*

Setting of the voice was primarily ensured through developing of the flexibility of performance based on the development of singing breathing. The process involved parsing pieces of music to ensure correct singing posture. Fragmentary singing was aimed at the possibility of regulating inhalations and exhalations, which ensure correct pronunciation. The development of the flexibility of performing is connected with the performance of musical fragments in different tonality, which forms an idea of the correct perception of words. Regulation of breathing allows you to ensure clear diction during the pronunciation of words. Processing of each sound was implemented with the help of vocal performance of words. This contributed to the understanding of ways of comparing the intonation of the word itself with the intonation of the melody and influenced the correct pronunciation of words. The AI-based programme iZotope RX 10 contributed to the implementation of the first learning method as a result of working with a musical text. iZotope RX 10 (Figure 3) facilitates the creation of text transcriptions on which voice changes (increasing and lowering sounds, musical accents) are displayed using spectrograms. Such segmentation enables vocalists to perceive appropriate prompts, which contributed to the personalized correction of vocal pronunciation.

***Fig. 3.*** *AI-Based Programme Izotope RX 10 For The Possibility Of Developing Vocal Intonation (Izotope, Inc., 2024 [34])*

The work on sound reproduction methods involved studying the peculiarities of pronunciation and intonation not of whole words or phrases, but of individual sounds. During the training period, students studied ways of setting the speech apparatus while pronouncing individual sounds. The process involved classifying sounds into vowels and consonants and correcting them sequentially based on interactive correction. The interactive platform Clip.audio (Figure 4) was used for the search for similar sounds in the same or other musical genres. The advantage of this interactive platform is related to the possibility of excluding additional noise, which would affect the quality of the perception of individual sounds. This ensures the exclusion of possible intonation errors and the search for new opportunities to improve vocal performance. This approach affects the development of musical hearing, which ensures a harmonious relationship between voice and pitch.

Journal of Theoretical and Applied Information Technology
31st March 2025. Vol.103. No.6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

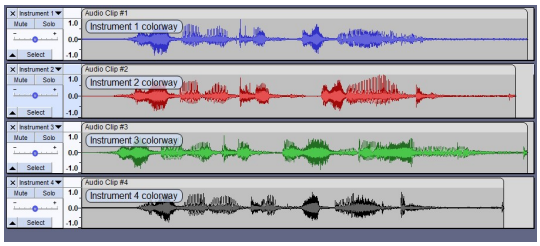


***Fig. 4.*** *Clip.Audio AI-Based Programme For The Development Of Vocal Intonation (Clip Audio, 2024 [35])*

Working on the voice mobility to ensure artistic expressiveness is connected with understanding the specifics of performing musical decorations. Musical decorations are a separate instrument of vocal performance and are aimed at not always generally accepted performance. Artistic expressiveness is interconnected with rendering of a certain intonation, which allows you to regulate the emotionality of the performance. Artistic expressiveness provides a clearer conveyance of musical images. The learning process involved the use of the iZotope RX 10 programme, which influenced the work with a particular expressive word. The students focused on understanding the techniques of the conversational and musical style of pronouncing artistic decorations during the training. This contributed to students' understanding of the need for the intonation used. The use of such exercises was aimed at the development of students' thinking based on understanding the approaches to maintaining intonation in pronunciation and singing. During the analysis, the students had to evaluate the singing attitude, evenness of sound, singing breath, sound production features, articulation setting.

The work on the formation of the musical performance style involved the analysis of recordings of one's own singing. Adobe Audition (Figure 5) includes elements of AI-based sound analysis, which facilitates working with your own recordings. Editing of the audio file involved correction of incorrect sounds, noise reduction. So, students improved their vocal performance and focused on achieving a professional sound. Processing vocal recordings provided an understanding the principles of using musical intonation during the performance of a particular text and the possibility of making additional changes. This educational technique contributed to the development of vocal skills, ensuring the achievement of correct technique, emotional, and intonation expressiveness. Consequently, students eliminated difficulties while singing and developed their own individual characteristics.

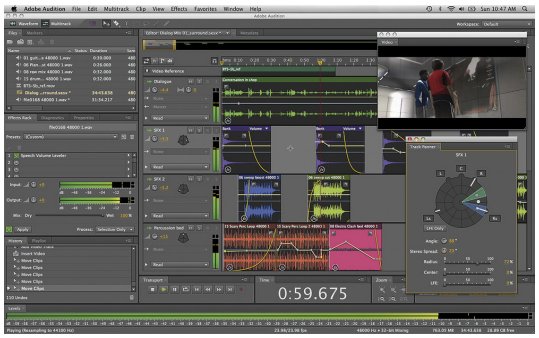


***Fig. 5.*** *Adobe Audition AI-Based Programme For The Development Of Vocal Intonation (Adobe Audition, 2024 [36])*

The next stage of the research was to determine the advantages of AI during training. Preferences were determined after four months of training using two-factor ANOVA calculations. The results are presented in Table 1.

***Table 1.*** *Advantages Of The AI-Based Programmes Used By The Students For Training Of Vocal Intonation (Calculated By The Author On The Basis Of Two-Factor ANOVA)*

| Possible advantages | SS | df | MS | $F_{exp.}$ | p | $F_{cr.}$ |
|---|---|---|---|---|---|---|
| Ability to automatically identify errors | 0.934 | 1.0 | 0.932 | 4.894 | 0.01 | 4.196 |
| Ability to adjust vocal singing | 0.856 | 1.0 | 0.855 | 4.540 | 0.02 | 4.196 |
| The possibility of studying individual vocal approaches, namely intonation | 0.901 | 1.0 | 0.900 | 4.826 | 0.01 | 4.196 |
| The possibility of using a non-standard approach to learning | 0.857 | 1.0 | 0.858 | 4.541 | 0.02 | 4.196 |
| Understanding the principles of vocal and intonation coordination of one's | 0.839 | 1.0 | 0.837 | 4.042 | 0.03 | 4.196 |

| | | | | | | |
|---|---|---|---|---|---|---|
| own singing | | | | | | |
| The possibility of developing a musical ear | 0.848 | 1.0 | 0.847 | 4.212 | 0.02 | 4.196 |
| Ease of sound formation | 0.835 | 1.0 | 0.834 | 4.037 | 0.03 | 4.196 |
| There are no advantages | - | - | - | - | - | - |

AI had advantages for students in the educational process, primarily *the possibility of automatic error detection*. As a result, students could form the correct performance style and focus on maintaining the correct intonation. During training, this approach made it possible to eliminate voice tension and maintain the evenness of singing. Identifying errors in performance allowed students to perform compositions at different tempos, ensuring expressiveness and purity of intonation. *Elaboration of individual vocal approaches, including intonation*, developed professional skills of students. The approach made it possible to exclude low-quality performance and ensure the possibility of performing more complex compositions. *The use of a non-standard approach to teaching* motivated students to develop vocal skills. It also promoted better memorization of information and the ability to freely use theoretical knowledge in practice. *The ability to adjust vocal singing* made it possible to improve vocal skills based on the constant development of new ones. Automatic obtaining of possible tasks for vocal correction affected the quality of information perception. AI also influenced the *possibility of developing musical ear* based on the performance of various tasks. The processing of each sound made it possible to perceive characteristic features by ear. The acquired skills contributed to the achievement of high results by students in vocal performance. It also made it *easier to create sounds* as a result of working on individual musical details.

The dynamics of changes in the cognitive component under the influence of AI was determined among students in the course of the research. The results were obtained with the purpose to understand the students' skills that ensured the quality of vocal singing. The data were displayed in accordance with the peculiarities of vocal intonation (Figure 6).

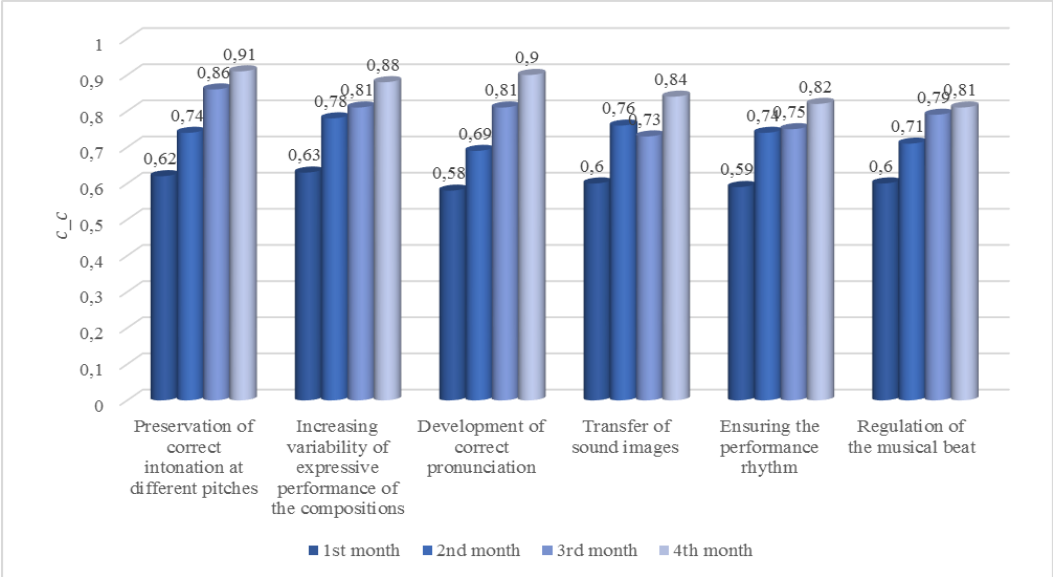


*Fig. 6. The Dynamics Of The Development Of Students' Vocal Skills Under The Influence Of The Development Of Vocal Intonation (Calculated By The Author Based On The CCI).*

AI contributed to the acquisition of the necessary vocal performance skills, correcting the possibility of intonation pronunciation. An important skill was maintaining the correct intonation at different

pitches, which affected the expressive performance. Students achieved technical clarity of performance, ensuring correct spacing. During the training, students developed the skills of correct pronunciation while singing. The development of articulation is connected with the use of interactive programmes that contributed to the work on each sound. Increased variability of the expressive performance of compositions made it possible to perform not only compositions of different complexity, but also to focus on different sounds of the same song. This is due to the clarity of application of the acquired skills to obtain the required result. When the students were reproducing the necessary intonation, they achieved the conveyance of sound images. The use of intonation promoted the conveyance of various emotions that reflect sound images. Regulating the musical tempo and rhythmicity of performance are related to the ability to maintain the necessary tempo during performance, focusing on the appropriate musical accompaniment.

The level of achieved correctness of intonation under the influence of the AI by students is measured below. The results were presented in general after training and during the performance of a particular composition for the possibility of comparing the obtained data. Calculations were obtained using two-way ANOVA (Table 2).

*Table 2. The Level Of Acquired Vocal Intonation Regulation Skills By Students (At $F_{cr.}$=4.196) (Calculated By The Author On The Basis Of Two-Factor ANOVA)*

| Possible level | General development of intonation | | | | | Performance of a particular composition | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Percentage of students | Obtained score | Two-factor ANOVA | | | Percentage of students | Obtained score | Two-factor ANOVA | | |
| | | | SS | df | $F_{exp.}$ | | | SS | df | $F_{exp.}$ |
| High | 78 | 9.8 | 0.947 | 1 | 4.373 | 91 | 9.7 | 0.980 | 1 | 4.597 |
| Medium | 21 | 7.5 | 0.634 | 1 | 4.041 | 9 | 7.3 | 0.562 | 1 | 4.027 |
| Low | 1 | 1.4 | 0.523 | 1 | 1.215 | - | - | - | - | - |

It was established that the students were able to develop high vocal intonation skills, both during the performance of a particular composition and during the study in general. The performance of the compositions involved focusing on a harmonious combination of sounds, which is connected with the conveyance of the most expressive elements. The students also achieved new skills in reproducing musical associations using intonation. The development of intonation skills also promoted loudness of performance because of the increased expressiveness and emotionality.

A high level of performance was also maintained during the performance of the composition The Sea, My Hometown. The students preserved the artistry of the musical piece, which was adjusted by intonation. This ensured the preservation of the lyricism and mood of the musical work. The students' preserved artistic logic, dynamic performance in their singing, which ensured an accurate combination of music and text.

## 5. DISCUSSION

High-quality vocal performance depends on the pre-processing of the composition, which is related to the principles of automatic alignment of the text, transcription of melodies. The processing of the composition affects the possibility of dividing the voice depending on the type of performance and the processing of the vocal melody. This approach makes it possible to create classical and modern approaches to vocal interpretation [37]. The development of different singing styles depends on the intensity of the acoustic sound and their configurations. Digital applications Twang and Belting provide the possibility of using similar approaches to the performance of the composition, which affects the search for the most pronounced vocal contrast. Avoiding noise interference allows for better intonation. Comparing the real singing of vocalists with experimentally possible ones allows to achieve better vocal performance [38]. The formation of types of voicing and correct intonation depends on the sequence of analysis of their singing. Speed parameters, pitch errors can be used for voice analysis. Consistency between vocal intervals depends on the level of training of vocalists. Also, the acoustic environment has an effect on maintaining quality intonation [39].

During the development of musical intonation, it is necessary to take into account that music and text have a direct connection. Therefore, it is necessary to develop sound hearing, phonological processing, and pronunciation in the educational process. During the performance of compositions from different countries, it is necessary to ensure mutual communication between the performer, the music teacher, and the foreign language teacher. This will contribute to the awareness of the performance process and the conscious perception of the musical text, which affects its consistency with the melody [40]. The published works aim to develop musical intonation based on the melody and text ratio and the development of aesthetic performance. However, there is no detailed description of the correct pronunciation of words associated with reproducing the necessary intonation. In our article, attention was paid to the preservation of musical intonation, primarily under the influence of the development of musical articulation.

Cognitive processes during singing take place during the study of choral scores. When singing, the use of sheet music is related to the search for visual cues to coordinate performance. It was established that the density of notes affects the achievement of fixation on one's own cues, but not on the cues of other voices. Ensuring the note density provides the possibility of combining vocal parts with a musical structure [41]. The self-regulation processes promote the achievement of educational goals and the development of student motivation. The use of Plectrus software develops the ability to provide quality intonation in real time. An interactive learning approach affects the possibility of using the most effective learning strategies that motivate students to achieve high results [42]. In the process of vocal performance, the relationship between auditory images and tonal and temporal accuracy of singing is important. Digital technologies enable auditory feedback. Receiving auditory feedback allows you to coordinate musical intonation and create the necessary tonal speech. Auditory pattern recognition is not a formal learning tool, but facilitates more efficient and expressive performance [43]. In the presented works, special attention is paid to the formation of vocal intonation under the influence of the development of students' independence. However, the presented analysis does not reflect the advantages and disadvantages of such training and excludes detailed ways to develop students' independence to achieve a high level of vocal performance. Our article was focused on the possibility of developing vocal intonation under the influence of developed educational processes using AI-based applications. Vocal performance promotes performers' self-expression, which improves the level of intonation development and artistic expression. Correct and beautiful performance depends on the development of technical and artistic skills. The trained intonation enable students to reproduce a musical image and develop their own potential [44]. The individual style of performance depends on the features of the voice timbre, which is connected with the achievement of technical correctness of performance. The SVQTD programme facilitates the downloading of music compositions from YouTube for their analysis. Challenge and INTERSPEECH computer algorithms are used to develop vocal skills and achieve clearer recognition of vocal intonation [45].

The discussion of existing published articles gives grounds to identify the peculiarities of vocal style formation. It was suggested that AI be used for intonation development, but it was not considered for specific educational programmes. In our research, the choice of AI technologies was associated with specific educational techniques. The used approaches made it possible to determine the archived level of the students' vocal intonation skills. The level of acquired knowledge was also assessed. Our research results are related to American, European, and Asian studies. The results were implemented based on the involvement of Chinese students, which contributed to the study of vocal intonation production approaches. The authors obtained favourable results in acquiring practical musical skills due to the use of existing AI technologies, which will contribute to the student's professional success in the future. Our research results confirm that setting the correct intonation contributes to a more emotional and technically accurate performance of songs of various musical genres.

### 5.1. Limitations

The limitations of the study are the lack of ways to develop vocal intonation depending on the musical genre. This approach will allow us to determine in more detail the benefits of using artificial intelligence to train vocalists. It will also contribute to vocalists' understanding of ways to correct the wrong intonation. But the study does not have a negative impact on the obtained results. The authors determined the advantages of AI in the educational process, focusing on the skills acquired by students and the level of knowledge obtained in the methods of vocal intonations.

### 5.2. Recommendations

Vocal singing is characterized by a combination of technical and emotional components. Therefore, in order to achieve high-quality vocal results, it is necessary to engage in setting vocal intonation. The authors recommend using modern educational methods, namely AI-based programmes, which contribute to achieving high results in the shortest possible time. Setting the right intonation will enable performing musical works of various levels of complexity.

## 6. CONCLUSIONS

The authors achieved the aim and confirmed the relevance of the issue under research in the course of the study. Our research was designed to develop educational methods aimed at developing vocal intonation skills. The AI-based applications iZotope RX 10, Clip.audio, Adobe Audition were used for the educational process. During training, students paid attention to the techniques for training flexibility of performing activities, ways of reproducing sounds, voice mobility, formation of musical style. The learning process ensured the identification of the advantages of AI by students. It was established that the very first advantage was related to the possibility of automatic error detection, which provided an understanding of the principles of vocal singing ($F_{exp.}$=4.894). The advantages also included the possibility of working out individual vocal approaches, namely intonation ($F_{exp.}$=4.826), using a non-standard approach to learning ($F_{exp.}$=4.541), adjusting vocal singing ($F_{exp.}$=4.540).

It was found that the educational process helped to preserve the correct intonation at different pitches ($C_c$=0,91), develop correct pronunciation ($C_c$=0,90). The students were able to ensure the expressiveness of a musical composition ($C_c$=0,88), provide rhythmicity ($C_c$=0,82) and musical beat ($C_c$=0,81). These skills contributed to the achievement of clear and emotional singing during the learning process. The positive value of the educational process is associated with students' achievement of a high level of musical intonation (among 78% with a score of 9.8). High results were also confirmed based on the performance of random compositions, where the correctness of vocal intonation was preserved.

Based on the results, the authors believe that the applications iZotope RX 10, Clip. Audio and Adobe Audition contribute to students' vocal development, which is reflected in the work with incorrect intonation. The final results confirm the need to transform the traditional training system, which facilitates the perception of educational materials and expands practical skills in vocal performance.

The practical significance of the study is the possibility of applying the proposed teaching methods for the development of vocal intonation among students of different years of study. The obtained results can be useful for specialized educational institutions or be used in vocal groups. Therefore, the work's strengths are related to the selection of effective interactive applications, focusing on specific mechanisms, such as iZotope RX 10, Clip.audio, and Adobe Audition, for the implementation of student vocal training. Research prospects may be related to the search for characteristic elements for the development of vocal intonation in various musical genres. The lack of study of various musical genres is a weak element of the study, the elimination of which will determine the effectiveness of interactive applications due to the study of multiple compositions.

**Acknowledgements**

Sanming Social Science Planning Project "Research on the Method of Intonation Correction of Artificial Intelligence Technology in College Vocal Music Teaching", Project No. 24042.

About the author: Xiao Huiwen is mainly engaged in the research of vocal music singing and artificial intelligence assisted music teaching.

## REFERENCES

[1] S.P. Hinkley, "Vocal Modeling: A Review of Literature", *Update: Applications of Research in Music Education,* Vol. 42, No. 2, 2024, pp. 23-30. https://doi.org/10.1177/87551233221146285

[2] B. Liu, and F. Ye, "The Problem-Based Approach in Online Music Education: How to Teach Students to Control Singing with Piano Accompaniment?", *Interactive Learning Environments,* Vol. 32, No. 6, 2023, pp. 2862–2873. https://doi.org/10.1080/10494820.2022.2160471

[3] F.M.B. Lã, and M.B. Fiuza, "Real-time Visual Feedback in Singing Pedagogy: Current Trends and Future Directions", *Applied*

*Sciences,* Vol. 12, No. 21, 2022, paper 10781. https://doi.org/10.3390/app122110781

[4] A. Wang, and L. Zhang, "Design and Implementation of Computer Visualization Assisted Vocal Music Teaching System", *Computer-Aided Design and Applications*, Vol. 19, No. S7, 2022, pp. 44-55. https://doi.org/10.14733/cadaps.2022.S7.44-55

[5] K.O. Islamovich, and Q.G. Durdimurotovna, "Formation of Vocal-Choir Skills in Music Lessons", *Miasto Przyszłości,* Vol. 33, 2023, pp. 57-62. https://uniwork.buxdu.uz/resurs/13099_1_2BD684295614F0589A10E1660220374EEC3ACD7F.pdf

[6] S. McCarther, and C. Arneson, "Reorganizing the Pillars of Voice Pedagogy: A Tiered Approach to Teaching Voice", *Journal of Singing,* Vol. 80, No. 2, 2023, pp. 135-143. https://muse.jhu.edu/article/910308

[7] D.S. Miller, "The Effect of Register, Direction, and Magnitude on Musicians' Evaluations of Chamber Ensemble Intonation: A Within-Study Comparison for Analysis of Repeated Measures", *Journal of Research in Music Education,* Vol. 70, No. 3, 2022, pp. 339-360. https://doi.org/10.1177/00224294211060457

[8] J. Tejada, and M.Á. Fernández-Villar, "Design and Validation of Software for the Training and Automatic Evaluation of Music Intonation on Non-Fixed Pitch Instruments for Novice Students", *Education Sciences,* Vol. 13, No. 9, 2023, paper 860. https://doi.org/10.3390/educsci13090860

[9] H. Lu, W. H. Tan, K. W. Cheong, and Y. Cong, "Unlocking Sound: New Trends and Innovations in Intonation Education Over the Past Decade", *International Journal of Music Education,* 2024. https://doi.org/10.1177/02557614241237528

[10] R. Shashidhar, M.P. Shashank, B. Sahana, "Enhancing Visual Speech Recognition for Deaf Individuals: A Hybrid LSTM and CNN 3D Model for Improved Accuracy", *Arabian Journal for Science and Engineering,* Vol. 45, 2023, pp. 11925-11941. https://doi.org/10.1007/s13369-023-08385-5

[11] H. Tang, "An Automatic Correction System of Singing Intonation Based on Deep Learning", *International Journal of Information and Communication Technology,* Vol. 22, No. 4, 2023, pp. 422-437. https://doi.org/10.1504/IJICT.2023.131190

[12] D. Zheng, and Y. Wang, "The Application of Computer-Aided System in the Digital Teaching of Music Skills", *Computer-Aided Design & Applications,* Vol. 19, No. S7, 2022, pp. 154-164. https://doi.org/10.14733/cadaps.2022.S7.154-164

[13] L. Zhang, and J. Hu, "Pop Music Singing in Education with Modern Innovative Technologies: How the Chinese Language Shapes the Creation of Popular Singing", *Journal of Psycholinguistic Research,* Vol. 52, 2023, pp. 2677-2691. https://doi.org/10.1007/s10936-023-10014-z

[14] A. Acquilino, and G. Scavone, "Current State and Future Directions of Technologies for Music Instrument Pedagogy", *Frontiers in Psychology,* Vol. 13, 2022, paper 835609. https://doi.org/10.3389/fpsyg.2022.835609

[15] S. Wang, "Vocal Improvisation Using Interactive Music Technology", *Interactive Learning Environments,* Vol. 32, No. 1, 2022, pp. 245-256. https://doi.org/10.1080/10494820.2022.2084755

[16] G.A. Martínez, and H. Daffern, "The Role of Vibrato in Group Singing: A Systematic Review", *Journal of Voice*, Vol. 5, 2022, pp. 1-19. https://doi.org/10.1016/j.jvoice.2022.11.004

[17] A. Jankowska, J. Pilarczyk, K. Wołoszyn, and M. Kuniecki, "Enough is Enough: How Much Intonation Is Needed in the Vocal Delivery of Audio Description?", *Perspectives,* Vol. 31, No. 4, 2022, pp. 705-723. https://doi.org/10.1080/0907676X.2022.2026423

[18] L. Duan, "The Application of Modern Virtual Reality Technology in the Teaching of Vocal Music", *Curriculum and Teaching Methodology,* Vol. 6, No. 19, 2023, pp. 77-81. https://doi.org/10.23977/curtm.2023.061912

[19] W. Wang, N. Zhang, W. Peng, and Z. Liu, "A New Intonation Quality Evaluation Method Based on Self-Supervised Learning", *Journal of Intelligent & Fuzzy Systems,* Vol. 45, No. 1, 2023, pp. 989-1000. https://doi.org/10.3233/JIFS-230165

[20] iZotope. Inc, https://www.izotope.com/en/learn/why-upgrade-to-rx-10.html?srsltid=AfmBOorlo0Zp_76E5OKRo

yD3CnL2QLRdc3_OZIq7iXEYen8N-M8nVMrn.
[21] Clip Audio, https://apps.apple.com/ru/app/clip-audio/id1618985158.
[22] Adobe Audition, https://www.adobe.com/pl/products/premiere/campaign/pricing.html?gclid=CjwKCAjwlbu2BhA3EiwA3yXyu0QObxBrHjNBY4kz2T1E8rRdZBQf3FC8TcbBzT2gd2W2xJowe61M.
[23] Y. Xu, W. Wang, H. Cui, M. Xu, and M. Li, "Paralinguistic Singing Attribute Recognition Using Supervised Machine Learning for Describing the Classical Tenor Solo Singing Voice in Vocal Pedagogy", *EURASIP Journal on Audio, Speech, and Music Processing,* Vol. 8, 2022, pp. 1-16. https://doi.org/10.1186/s13636-022-00240-z
[24] H. Tang, "An Automatic Correction System of Singing Intonation Based on Deep Learning", *International Journal of Information and Communication Technology,* Vol. 22, No. 4, 2023, pp. 422-437. https://doi.org/10.1504/IJICT.2023.131190
[25] A. Vurma, E. Meister, L. Meister, J. Ross, M. Raju, V. Kala, and T. Dede, "The Intensities of Vowels and Plosive Bursts and Their Impact on Text Intelligibility in Singing", *The Journal of the Acoustical Society of America,* Vol. 154, No. 4, 2023, pp. 2653-2664. https://doi.org/10.1121/10.0021968
[26] A. Fernández-Barros, D. Duran, and L. Viladot, "Peer Tutoring as a Tool for Developing the Intonation of Violin and Viola Students in Elementary Music Education", *Music Education Research,* Vol. 25, No. 2, 2023, pp. 176-189. https://doi.org/10.1080/14613808.2023.2193210
[27] Music Sonyselect, https://music.sonyselect.net/
[28] iZotope. Inc, https://www.izotope.com/en/learn/why-upgrade-to-rx-10.html?srsltid=AfmBOorlo0Zp_76E5OKRoyD3CnL2QLRdc3_OZIq7iXEYen8N-M8nVMrn
[29] Clip Audio, https://apps.apple.com/ru/app/clip-audio/id1618985158
[30] Adobe Audition, https://www.adobe.com/pl/products/premiere/campaign/pricing.html?gclid=CjwKCAjwlbu2BhA3EiwA3yXyu0QObxBrHjNBY4kz2T1E8rRdZBQf3FC8TcbBzT2gd2W2xJowe61M
[31] Y. Xu, W. Wang, H. Cui, M. Xu, and M. Li, "Paralinguistic Singing Attribute Recognition Using Supervised Machine Learning for Describing the Classical Tenor Solo Singing Voice in Vocal Pedagogy", *EURASIP Journal on Audio, Speech, and Music Processing*, Vol. 8, 2022, pp. 1-16. https://doi.org/10.1186/s13636-022-00240-z
[32] H. Tang, "An Automatic Correction System of Singing Intonation Based on Deep Learning", *International Journal of Information and Communication Technology,* Vol. 22, No. 4, 2023, pp. 422-437. https://doi.org/10.1504/IJICT.2023.131190
[33] A. Vurma, E. Meister, L. Meister, J. Ross, M. Raju, V. Kala, and T. Dede, "The Intensities of Vowels and Plosive Bursts and Their Impact on Text Intelligibility in Singing", *The Journal of the Acoustical Society of America,* Vol. 154, No. 4, 2023, pp. 2653-2664. https://doi.org/10.1121/10.0021968
[34] iZotope. Inc, https://www.izotope.com/en/learn/why-upgrade-to-rx-10.html?srsltid=AfmBOorlo0Zp_76E5OKRoyD3CnL2QLRdc3_OZIq7iXEYen8N-M8nVMrn
[35] Clip Audio, https://apps.apple.com/ru/app/clip-audio/id1618985158
[36] Adobe Audition, https://www.adobe.com/pl/products/premiere/campaign/pricing.html?gclid=CjwKCAjwlbu2BhA3EiwA3yXyu0QObxBrHjNBY4kz2T1E8rRdZBQf3FC8TcbBzT2gd2W2xJowe61M
[37] R. Monir, D. Kostrzewa, and D. Mrozek, "Singing Voice Detection: A Survey", *Entropy,* Vol. 24, 2022, p. 114. https://doi.org/10.3390/e24010114
[38] M. Fleischer, S. Rummel, F. Stritt, J. Fischer, M. Bock, M. Echternach, B. Richter, and L. Traser, "Voice Efficiency for Different Voice Qualities Combining Experimentally Derived Sound Signals and Numerical Modelling of the Vocal Tract", *Frontiers in Physiology,* Vol. 13, 2022, paper 1081622. https://doi.org/10.3389/fphys.2022.1081622
[39] P. Bottalico, N. Łastowiecka, J.D. Glasner, and Y.G. Redman, "Singing in Different Performance Spaces: The Effect of Room Acoustics on Vibrato and Pitch Inaccuracy", *The Journal of the Acoustical Society of America*, Vol. 151, No. 6, 2022, pp. 4131-4139. https://doi.org/10.1121/10.0011675

[40] J. Lehtinen-Schnabel, and S. Levänen, "Singing Language – Integrating Second Language Learning into Choir Practice", *Music & Science,* Vol. 7, 2024, pp. 1-25. https://doi.org/10.1177/20592043241244542

[41] M. Timoshenko-Nilsson, M. Nyström, and E. Huovinen, "Sight-Singing in a Group Context: An Eye-Tracking Study with Experienced Choral Singers", *Journal of New Music Research,* 2024, pp. 1-15. https://doi.org/10.1080/09298215.2024.2375557

[42] F. López-Calatayud, and J. Tejada, "Self-Regulation Strategies and Behaviours in the Initial Learning of the Viola and Violin with the Support of Software for Real-Time Instrumental Intonation Assessment", *Research Studies in Music Education,* Vol. 46, No. 1, 2024, pp. 48-65. https://doi.org/10.1177/1321103X221128733

[43] C.N. Reed, M. Pearce, and A. McPherson, "Auditory Imagery Ability Influences Accuracy when Singing with Altered Auditory Feedback", *Musicae Scientiae,* Vol. 28, No. 3, 2024, pp. 478-501. https://doi.org/10.1177/10298649231223077

[44] T. Ekici, "An Evaluation on the Human Voice and the Act of Singing", *Turkish Online Journal of Educational Technology*, Vol. 21, No. 3, 2022, pp. 1-14.

[45] Y. Xu, W. Wang, H. Cui, M. Xu, and M. Li, "Paralinguistic Singing Attribute Recognition Using Supervised Machine Learning for Describing the Classical Tenor Solo Singing Voice in Vocal Pedagogy", *EURASIP Journal on Audio, Speech, and Music Processing,* Vol. 8, 2022, pp. 1-16. https://doi.org/10.1186/s13636-022-00240-z

**Appendix A**

The text of the composition The Sea, My Hometown
小时候妈妈对我讲
大海就是我故乡
海边出生 海里成长
大海 啊大海 是我生活的地方
海风吹 海浪涌 随我飘流四方
大海 啊大海 就像妈妈一样
走遍天涯海角 总在我的身旁

(repeat)
大海啊故乡 大海啊故乡
我的故乡 我的故乡

When I was little, my mother would tell me
The sea is my hometown
Born by the sea, grown up in the sea
Sea oh sea, it's where I live
The sea wind blows, the waves surge, following me wherever I go
Sea oh sea, just like my mother said
Traveling to the ends of the earth, she is always by my side
(Repeat)
Sea oh sea, the sea, my hometown
My hometown, my hometown