International Journal for Housing Science and Its Applications
Publish August 6, 2025. Volume 46, Issue 3 Pages 7073-7082

https://doi.org/10.70517/ijhsa463607

Research on the Quality Improvement Path of Artificial Intelligence-Assisted Teaching in the Construction of Higher Vocational English Gold Classes under the Integration of Industry and Education

Xiaoya Zhang[1,*]
[1] Department of Public Instruction, Shandong Vocational College of Special Education, Jinan, Shandong, 250000, China
Corresponding authors: (e-mail: 13210559222@163.com).

Abstract Aiming at the pain points of the construction of higher vocational English gold classes in the context of the integration of production and education, this paper explores the path of teaching quality improvement empowered by artificial intelligence technology. Students' English pronunciation is captured using speech recognition technology, and learning style preferences are parsed based on SVM algorithm. The assessment method based on the standard speech reference template is proposed, and the pronunciation is corrected with the aid of the speech recognition results. The real-time pronunciation feedback and personalized resource recommendation mechanism are embedded into the smart teaching platform to promote the construction of English Golden Class. Setting up a 12-week teaching experiment, the mean difference between the scores of the two classes in the pre-test is 0.236, t=0.302, p=0.803>0.05, and there is no significant difference. In the posttest, the scores of the students in the experimental group were concentrated in the 35-50 point value range, while the scores of the control group were concentrated in the 20-35 point value range. The experiment proves that the proposed teaching model helps to promote the continuous deepening of the industry-teaching collaborative education model.

Index Terms higher vocational English teaching, speech recognition technology, SVM algorithm, English gold class construction

I. Introduction

The integration of industry and education can be understood as the establishment of links with industrial development in the process of promoting school education, forming a mutually beneficial education model. The scientific promotion of higher vocational English teaching in the context of industry-education integration is one of the important paths for the cultivation of higher vocational English talents [1]. One of the important points lies in the design of high-quality courses, which take the initiative to strengthen the connection with enterprises in education, so that English education is more in line with the actual requirements of industrial development [2]. The three distinctive features of a gold class are innovation, challenge and higher order, and the first task for English teaching is to keep in line with the actual needs of the country and society for talents [3]. So that the professional non-general English can be in the teaching process, according to the students themselves and the needs of future positions in the teaching design, not only to strengthen the cultivation of humanistic qualities of the students, to enhance cross-cultural communication skills, insight into cultural differences, and enhance the international outlook, but also to make the students through the English learning to master the knowledge of the English language and skills, at the same time, enhance the students' character cultivation and industrial literacy, for the future smooth to embark on the work position to lay a good foundation [4]-[7].

And with the arrival of the digital era, the digital society has put forward new requirements on teaching quality and teaching methods, which objectively activates the digital needs of students. Students in the digital era are more inclined to acquire knowledge and skills through digital technology to enhance the convenience and efficiency of learning [8]. From the analysis of potential effects, the English curriculum system after the digitalization of education can not only support the technological innovation requirements of high-end industries in the era of digital economy, but also meet the requirements of the digital upgrading of traditional industries in terms of industrial adaptability [9], [10]. On the supply side of resources can not only actively adapt to the digital educational environment, provide time and space conditions to guarantee students' personalized ubiquitous learning, but also optimize the allocation of teaching resources in an all-round way, solve the reality of insufficient supply of resources for students' independent learning, release the vitality of resources to the maximum extent possible, emancipate the productivity of teaching and ultimately realize the return of the value of the curriculum to be people-oriented, and pave the way for the construction of gold classes [11]-[14].

However, in the existing construction of higher vocational English gold course, the teaching content and market demand matching degree is low, the students' individual difference needs are not satisfied, resulting in the students' gap continues to widen, and the course construction you take the degree of aggravation, and the school-enterprise integration of the number of educational case resources is small, and the matching degree of the teachers' team is low, resulting in inefficiency of the integration of industry and education and thus affecting the quality of teaching [15], [16]. Therefore, how to use digital technology to better meet students' professional development needs and skill enhancement expectations, and enhance students' sense of access to digital learning and sense of achievement has become a key issue that needs to be urgently resolved in the construction of high-quality gold course system.

In this paper, firstly, the speech signal processing mechanism of speech recognition technology is systematically elaborated, and the pronunciation scoring method based on standard speech reference template is proposed. The SVM algorithm is used to train and classify the extracted speech features to determine students' learning styles. Design a technology-enabled golden lesson construction strategy to transform the technological capability into teaching effectiveness. Set up three groups with different amounts of vocabulary to explore the effect of different amounts of vocabulary on the speech recognition rate. Utilize waveform fundamental frequency maps for speech visualization comparison to examine the effectiveness of pronunciation scoring and correction methods. Combine the results of speaking pre-test and post-test to evaluate the effectiveness of language recognition technology to assist English teaching.

II. Application of speech recognition technology in English language teaching

Under the background of deepening the integration of industry and education, higher vocational English teaching is facing the urgent need to transform from knowledge transmission to ability cultivation, and the introduction of artificial intelligence technology provides technical support for the construction of intelligent and dynamic English gold classes. This paper focuses on the practical path of AI-assisted teaching in the construction of senior vocational English gold classes, and explores the formulation of personalized teaching strategies by analyzing the core principle of speech recognition technology and its application in teaching, aiming to form a paradigm of construction of English gold classes that is collaborated between industry and education and empowered by technology.

II. A. Speech Recognition Overview

II. A. 1) Speech Recognition Fundamentals

Speech recognition technology, broadly speaking, refers to semantic recognition and voiceprint recognition; narrowly speaking, it refers to the understanding and recognition of speech semantics, also known as automatic speech recognition (ASR). As the most natural way of human-computer information interaction, the basic idea of speech recognition technology takes speech as the research object, and converts the input speech signals into corresponding text commands through the machine's recognition and comprehension process, so as to realize the control of machines by speech.

In the development of speech recognition technology, although different researchers have proposed many different solutions, the basic principles are similar. The principle of speech recognition is shown in Figure 1, in the processing of speech signals, any speech recognition system can use Figure 1 to represent its approximate recognition principle, the most important module of speech recognition system is speech feature extraction and speech pattern matching.

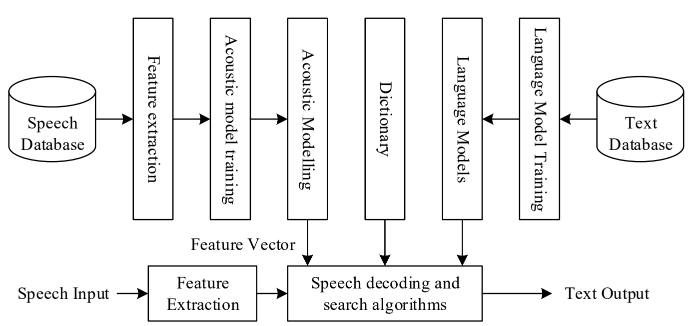

Figure 1: The principle of speech recognition

II. A. 2) Speech signal preprocessing

The first step of speech recognition is speech signal preprocessing, speech signal preprocessing is the premise and foundation of speech recognition, and it is also a very critical step for feature extraction of speech signal. Only in the stage of speech signal preprocessing to extract the characteristic parameters that can represent the nature of the speech, only to be able to compare the speech with the standard speech comparison to get the best similarity effect language. The preprocessing module of the speech signal generally includes two steps: digitization of the speech signal and pre-emphasis.

(1) Digitization of speech signal

Speech signal is a waveform that changes with time, is an analog signal, but because the computer only recognizes digital signals, so in order for the computer to be able to process the speech signal, it is necessary to digitize the speech signal. The process of digitizing speech signal includes sampling and quantization, after the processing of sampling and quantization, the speech signal becomes a discrete digital signal.

In this paper, the acquisition of voice signals is recorded by the microphone that comes with the Android phone, and in the background, the API android.media.AudioRecord provided by Android is used to configure the sampling rate and the number of quantization bits in the digitization process, so as to get the digital audio signals to be used. On the Web side is used in the Silverlight plug-in, the same can be recorded on the audio.

(2) Pre-emphasis

The purpose of pre-emphasis is to enhance the high-frequency signals of the voice, remove the low-frequency signals in the voice signal, so that the signal spectrum becomes flatter. For the spectrum of the speech signal, usually the higher the frequency the smaller the amplitude, in the frequency of the speech signal increases twice, the amplitude of the power spectrum decreases by 6 dB, in order to make the spectrum of the signal become flat, in order to facilitate the analysis of the spectrum and other characteristics of the parameter, the need for pre-emphasis of the speech signal. High-frequency speech signals and low-frequency speech signals are acquired with different difficulties, and pre-emphasis is designed to resolve this contradiction. The common practice is to pass the speech signal through a digital filter with 6Db/oct of boosted high frequency characteristics, which is a first order digital filter:

$$H(z) = 1 - \mu * z^{-1} \tag{1}$$

If expressed in the time domain, the pre-emphasized signal $S_2(n)$ is:

$$S_2(n) = S(n) - \mu * S(n-1) \tag{2}$$

where μ takes the value 0.9375.

II. B.Pronunciation Scoring Mechanisms and Methods

(1) Introduction to the Pronunciation Scoring Mechanism

By evaluating students' pronunciation accurately and scientifically, and giving certain corrective suggestions to students' pronunciation status, students can improve their English proficiency on the basis of continuous improvement of their pronunciation level.

The method of evaluating students' pronunciation level is to compare the matching degree between students' pronunciation and the reference template of standardized pronunciation, and a high matching degree means that the students' pronunciation is close to the standardized pronunciation and is of high level, and vice versa, the level is low and needs to be improved.

By analyzing the needs of the voice evaluation function of English learning for higher vocational students, it is clear that the requirements for the voice evaluation algorithm are to be able to accurately and reliably give voice evaluation scores, and to be able to give the calculation results in a relatively short period of time after the students have finished following the readings, i.e., the real-time nature of the evaluation algorithm has to meet the requirements.

Based on the limitations of hardware and real-time requirements, it is necessary to find a recognition method that can have a high recognition rate for isolated words and small vocabulary, and does not require other training. Therefore, the evaluation method based on the standard speech reference template becomes a good choice.

The scoring system flow of the standard speech reference template is shown in Fig. 2, which is the system flowchart of the standard speech reference template based assessment method. The first step is feature extraction; the second step is pattern matching using the extracted feature values; the third step is to obtain the similarity between two feature values by pattern matching, followed by mapping to the percentage scoring mechanism to derive a specific score.

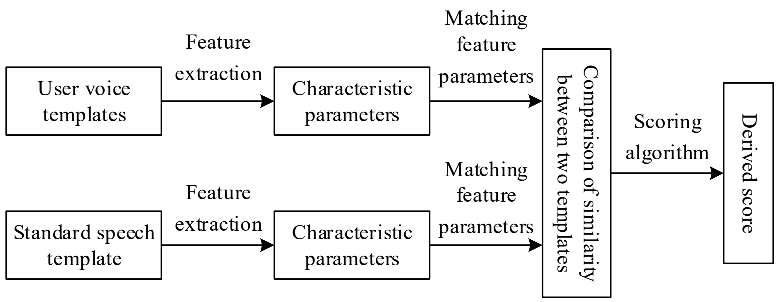

Figure 2: The process of the voice scoring mechanism

The Euclidean distance $D_{\min}(N,M)$ distance is a numerical measure of the similarity between the to-be-evaluated template and the standard reference template, and the similarity of the two speech signals can be visualized through this scale.

The scoring mechanism of pronunciation assessment is to obtain a mapping relationship between the matching distance between the speech signal to be assessed and the signal of the labeled parameter template to the final pronunciation score displayed on the user interface, and to give a method to obtain the pronunciation score through the matching distance. The previous matching distance between the to-be-assessed template and the reference template is $D_{\min}(N,M)$, and since the frame length of each speech signal is not the same, and the longer each frame is, the longer the total matching length is obtained, so in order to make each matching distance have the same metric, the notion of average matching distance is introduced, and average matching distance is The quotient of the sum of the matching distances and the frame length of the template to be tested, i.e:

$$\bar{d} = \frac{D(N,M)}{N} \tag{3}$$

By using the average of the distances of each frame as a measure, the effect of the unequal length of the speech signals can be eliminated, and therefore, this average is used to evaluate the user's pronunciation level.

(2) Scoring method based on a single reference template

The principle of single-reference template-based scoring is to calculate the frame-averaged distance between the user's speech to be evaluated m_1 and the standard reference template m_2, and then to obtain the final pronunciation scores presented on the interface by function mapping.

In order to make the differentiation more obvious, the evaluation score of this system adopts the form of percentage, and the frame average distance and the evaluation score have the following mapping relationship:

$$score = \frac{100}{1+a(d)^b} \tag{4}$$

d is the average frame matching distance, which is converted into a percentage by equation (4). a and b are scoring parameters, which are derived from the relationship between expert scores and matching distances. The formulas for multiple features and scoring parameters are as follows:

$$score = w_1 \cdot \frac{100}{1+a_1(d_1)^{b_1}} + w_2 \cdot \frac{100}{1+a_2(d_2)^{b_2}} + \cdots + w_n \cdot \frac{100}{1+a_n(d_n)^{b_n}} \tag{5}$$

where $w_1 + w_2 + \cdots + w_n = 1$, w is the weight of each feature parameter, and a, b are the corresponding transformation parameters.

In the process of pronunciation evaluation calculation, the Meier cepstrum coefficient plays a larger role, so the system chooses the Meier cepstrum coefficient as the feature parameter of evaluation.

(3) Scoring method based on dual reference templates

The scoring method based on a single template will be affected by the chance of the standard parameter template in the evaluation, so a scoring method based on a double reference template is proposed.

The scoring method of dual reference templates is to match the speech to be evaluated m_3 with two standard reference templates m_1 and m_2 respectively. Three frame average distances d_{12}, d_{13} and d_{23} can be obtained by matching between the two, and thus the frame distances between the to-be-evaluated template and the reference templates are the average of d_{13} and d_{23}, which is denoted as $d=(d_{13}+d_{23})/2$. This scoring method improves the stability of the scoring method by calculating the average distance between the speech to be evaluated and the Meier cepstrum coefficients frames of the two reference templates, and finally obtaining a final result from the two averages, which is determined by two parameters, which will keep the final result from being too arbitrary. d_{12} is the distance between the speech frames of the two reference templates, so the distance is set to be full for d_{12}. By comparing the size of d and d_{12}, it can be concluded that the closer d is to d_{12}, the higher the pronunciation level is.

The relationship between the three speech signal frames is shown in Fig. 3, and m_1 and m_2 are used as two reference templates to compare with the speech m_3 to be evaluated. When $(d_{13}+d_{23})>2d_{12}$, the frame matching distance is taken as:

$$d=(d_{13}+d_{23})/2 \tag{6}$$

When $(d_{13}+d_{23})\leq 2d_{12}$, it is reasonable to assume that the levels of the three pronunciations are close together, so the minimum value of $d_{13}+d_{23}$ is taken as the matching distance:

$$d=\min(d_{13}+d_{23}) \tag{7}$$

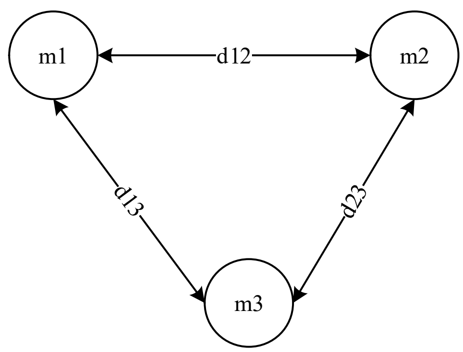

Figure 3: Relationship between three speech signal frames

II. C.Speech Recognition and Learning Style Classification

After completing the feature extraction of students' speech signals, machine learning algorithms are used to classify the learning styles. The article uses Support Vector Machine (SVM) algorithm to train and classify the extracted speech features.SVM is a commonly used supervised learning algorithm, which is especially suitable for the classification problem in high-dimensional feature space.SVM separates the data points of different categories by constructing hyperplanes and improves the classification accuracy by maximizing the interval. Its mathematical model can be expressed as

$$f(k)=sign(w^T x+b) \tag{8}$$

where, w denotes the normal vector of the hyperplane; x denotes the input feature; and b denotes the bias term; $f(k)$ denotes the classification result. When $w^T x+b>0$, $f(k)=+1$, indicating that the input feature x is classified as a positive class; When $w^T x+b<0$, $f(k)=-1$, indicating that the input feature x is categorized as a negative class; When $w^T x+b=0$, $f(k)=0$, indicating that the input feature points lie on the classification boundary, which rarely happens in real models. In the training process of SVM, the optimal w and b are solved by the optimization algorithm to make the classification accuracy optimal. In this paper, SVM is used to classify learning styles, including different types of students such as visual, auditory and kinesthetic. By learning from the training data, the SVM classification model can accurately determine the learning styles of students based on their speech characteristics, thus providing a basis for personalized teaching.

III. Strategies for the construction of senior English gold classes in the context of intelligent teaching and learning

On the basis of clarifying the supporting capability of speech recognition technology, the construction of gold classes should focus on the systematic integration of technical elements and teaching scenarios. Based on the speech feature extraction and learning style classification model proposed in this paper, this paper will explore how to transform the technical capability into teaching effectiveness from the two dimensions of teaching concept reconstruction and platform architecture design. Specifically, by embedding real-time pronunciation feedback and personalized resource recommendation mechanism, the intelligent teaching platform becomes the carrier of technology, so as to achieve the goal of “assessment for learning, learning for teaching”.

(1) Establishing the concept of building the English Golden Class

The golden lesson is not just a course construction concept, and its construction process requires a scientific and systematic layout. Especially in the context of intelligent teaching, the reform and innovation of higher vocational English teaching should be taken as the goal to strengthen the ideology of the construction of the golden class. Specifically, the following measures should be taken: on the one hand, based on the concept, mode and platform of wisdom teaching, the traditional concept of test-taking education in higher vocational English should be transformed, and the guided teaching of students should be strengthened, so as to realize that the concept of wisdom teaching is rooted in the construction of the golden class. On the other hand, change the curriculum concept of higher vocational English teachers. Because of the long-term senior vocational English teachers stick to the stereotypes to carry out teaching work, a single solidified mode of thinking has been formed, usually passive acceptance of the curriculum teaching arrangements, coping with the completion of the English teaching task, the limitations of these ideas and thinking, the wisdom of teaching reform and the construction of the Golden Class has a negative impact on the construction of the Golden Class. Therefore, we should start from the dimension of ideology and conception, guide higher vocational English teachers to be deeply aware of the development trend of wisdom teaching, make clear the value and goal of higher vocational English gold class construction, and always pursue the construction concept of gold class of excellence and quality teaching, so as to promote the construction and perfect development of higher vocational English gold class.

(2) Construct and improve the intelligent teaching platform

Under the background of wisdom teaching, the construction of higher vocational English gold class needs to take the wisdom teaching platform as a carrier, and steadily promote the platform construction around the actual needs of English gold class. First, widely integrate the relevant materials and information of higher vocational English teaching, such as teaching materials and lesson plans, to promote the construction of intelligent teaching materials, summarize and refine the English knowledge, highlight the epochal nature of the teaching content and the novelty of the way of presenting the teaching content, in order to adapt to the development requirements of intelligent teaching. Secondly, strengthen the construction of information system for intelligent teaching, focus on the syllabus and practice requirements of higher vocational English courses, focus on building the teaching system, resource sharing system and data application and analysis system of the English Golden Class, and lay a solid foundation for the construction of the English Golden Class by giving full play to the functions of different systems. Thirdly, to carry out online education of higher vocational English based on the intelligent teaching platform, it is necessary to build a perfect buddy and group system and classroom interaction system. These two types of systems are not only the key to the construction of the intelligent teaching platform, but also play a cornerstone role in the construction of the English Golden Class. On the basis of building a perfect intelligent teaching platform, it promotes the transformation of informatization, digitization and intelligence in the teaching of higher vocational English courses, extensively collects the data information of students' learning in the system, and summarizes and refines it to understand the students' learning progress, learning status and learning problems, so as to facilitate the teachers of higher vocational English to combine the actual situation with the adjustment of the content of the English teaching, the teaching methodology and the teaching strategy.

IV. An empirical study of artificial intelligence-assisted construction of higher vocational English gold classes

This study adopts a quasi-experimental design to verify the effectiveness of speech recognition technology in assisting higher vocational English teaching, integrating quantitative analysis and qualitative assessment through a mixed research method. The experimental subjects are 100 students majoring in Business English in the class of 2024 in a higher vocational college, randomly divided into the experimental group (n=50) and the control group (n=50). The experimental group adopts a teaching mode assisted by speech recognition technology, which uses speech recognition technology to collect students' English pronunciation, parse learning style preferences, and then generate personalized teaching strategies. At the same time, the results of speech recognition were combined to assist in the correction of pronunciation. In the control group, the traditional multimedia teaching mode was used, with only PPT lectures and audio exercises accompanying the textbook. The experimental period was 12 weeks, including a 2-week technology adaptation period.

IV. A. Analysis of the effectiveness of speech recognition technology-assisted pronunciation correction

IV. A. 1) Speech recognition

In order to investigate the effect of different number of vocabularies on the speech recognition rate, three groups with different number of vocabularies were set up, which were 10, 50, and 100 isolated words, and 200 random tests were conducted for each group respectively. The test speech from the experimental group and the control group of students, and the three groups of data were collected in the laboratory, the small number of people outdoor scene 1, the crowded and noisy outdoor scene 2, and the collection of the test template library composed of this time. The recognition rate test results are shown in Table 1. Although the influence of outdoor noise leads to the lowest recognition rate in outdoor environment 2, the recognition results of this time still meet the needs of real-world applications, and the recognition rates are 95%, 93.5%, and 90.5% when the number of isolated words is 10, 50, and 100, respectively. In addition, according to the test results, the larger the test vocabulary is, the lower the recognition rate is. However, the recognition rate can still reach more than 90%, and in general, the recognition rate of the model can meet the requirements of practical use.

Table 1: Test Results of Recognition Rate

Environment	Vocabulary count	Total number of tests	The correctly identified number	Recognition rate
Laboratory	10	200	196	98%
	50	200	194	97%
	100	200	190	95%
Outdoor Scene 1	10	200	194	97%
	50	200	190	95%
	100	200	185	92.5%
Outdoor Scene 2	10	200	190	95%
	50	200	187	93.5%
	100	200	181	90.5%

IV. A. 2) Pronunciation Scoring and Correction

The results of the speech recognition were scored using the scoring method of the standard speech reference template. The logarithmic base was taken to be 10, and out of nearly 500 speech sentences used in the scoring experiment, a total of 17,535 valid phoneme scores were obtained, with a minimum of -3351 and a maximum of -2.5. However, since only 83 were less than -1000, they were not considered. The distribution of logarithmic scores in the experiment is shown in Figure 4. The horizontal axis of the figure represents the log likelihood and the vertical axis represents the number of distributions of the likelihood. The length of the window for finding the quantities is 5. The graph clearly shows that there are high points and low points at both ends. Scores between -110 and -93 are the most concentrated, while those below -500 are rare.

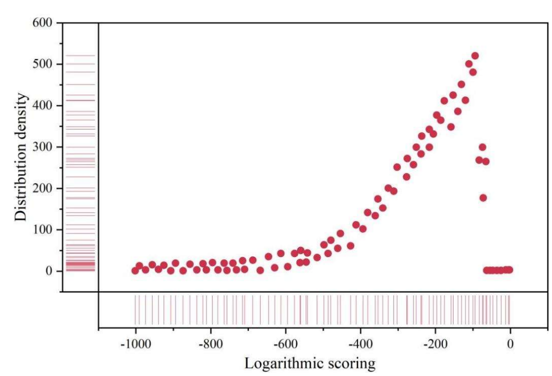

Figure 4: Distribution Density of Likehood Scoring

A student's voice was selected for analysis when reading a certain word, and the waveform fundamental frequencies of the reference voice, the student's voice and the rhyme-corrected voice of the word are shown in

Figure 5. It can be seen that the student's pronunciation appears to be continuous in fundamental frequency but discontinuous in vocal tract spectrum, and the transition of pronunciation is not natural. After feedback and correction, the student's pronunciation is close to the reference speech and the pronunciation tends to be standardized.

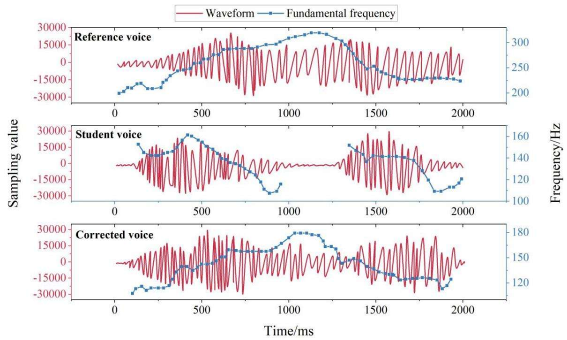

Figure 5: Waveform fundamental frequency comparison

IV. B. Classification of learning styles

After carrying out the formal experiment, SVM was used to categorize the learning styles of the students in the experimental group, and the results of the learning style classification are shown in Table 2. The learning styles of the students in the experimental group can be categorized into six types: visual, auditory, kinesthetic, experiential, group, and individual, among which the auditory type of students accounted for the highest proportion of students, reaching 30%. The number of experiential students is the lowest, only 3.

Table 2: Classification Results of Learning Styles

Style type	Number of people	Proportion
Visual type	9	18%
Auditory type	15	30%
Hands-on type	7	14%
Experiential type	3	6%
Group type	12	24%
Personal type	4	8%

IV. C. Analysis of application effects

Before and after the teaching practice, this paper conducted the speaking pre-test and post-test respectively, and the full score of the test was 50 points. EXCEL and spss software were used to collect, organize and analyze the data, and the results of the speaking pre-test and post-test are analyzed and discussed below.

IV. C. 1) Pre-testing results and analysis

An independent samples t-test was conducted for the pre-test data of the two classes and the results are shown in Table 3. From the results in Table 3, the difference between the means of the two classes is 0.236 and the standard error value is 0.973, which indicates that the difference is not large. The results of the independent samples t-test by spss software show that t=0.302 and the p-value is 0.803, which is greater than 0.05. It can be concluded that before the experiment, there is no significant difference between the performance of the experimental group and the control group in the speaking pre-test, and that the overall speaking performance of the students in the two classes is at the same level.

Table 3: Results of independent sample t-tests

		Levene test		Mean variance t-test						
		F	Sig.	t	df	Sig. (Bilateral)	Mean difference	Standard error value	95% confidence interval of the difference	
									Lower limit	Upper limit
Grade	Equal	.633	.501	.302	62.263	.803	.236	.973	-1.973	2.535
	Not equal			.302	60.186	.803	.236	.973	-1.987	2.582

IV. C. 2) Post-test results and analysis

After 12 weeks of teaching practice, this paper conducted an oral post-test for students in the experimental and control groups, and the frequency distribution of the post-test scores of the experimental and control groups is shown in Figure 6. The scores of the students in the experimental group are concentrated in the 35-50 score range, while the scores of the control group are concentrated in the 20-35 score range. It shows that there is a difference in the speaking scores of the two groups of students after teaching practice, which verifies the effectiveness of speech recognition technology-assisted teaching.

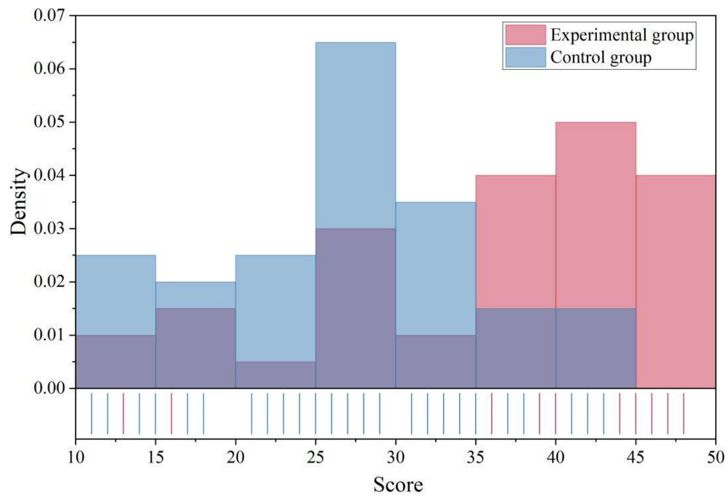

Figure 6: Frequency distribution of posttest scores

V. Conclusion

This paper proposes an artificial intelligence-assisted English teaching strategy based on speech recognition technology, and sets up experiments to explore the effect of its practical application.

The speech recognition rate in all the experiments exceeds 90%, and the recognition rate in the outdoor environment2 is the lowest when the isolated words are 10, 50 and 100, but still reaches 95%, 93.5% and 90.5%. The voice of a student was selected to be analyzed when reading a certain word, the student's pronunciation transition was unnatural, after feedback and correction, the student's pronunciation became standardized.

In the teaching experiment, the mean difference between the scores of the two classes in the pre-test was 0.236, $t=0.302$, $p=0.803>0.05$, and there was no significant difference. In the post-test the scores of the students in the experimental group were concentrated in the 35-50 point value range, while the scores of the control group were concentrated in the 20-35 point value range. It shows that there is a difference in the speaking scores of the two groups of students after the teaching practice, which verifies the effectiveness of speech recognition technology-assisted teaching.

References

[1] Danlu, L. (2023). A Study on the Talents Cultivation Model of Vocational Tourism English Students Based on the Integration of Industry and Education from the perspective of Stakeholders. Frontiers in Educational Research, 6(16).

[2] Qian, W. (2019, December). Research on the Construction of Public English Gold Course in Higher Vocational Education Against the Background of New Era. In 6th International Conference on Education, Language, Art and Inter-cultural Communication (ICELAIC 2019) (pp. 140-143). Atlantis Press.

[3] Zhang, M. (2021, June). of English Mixed Gold Course Learning Mode in Colleges and Universities Based on Mixed. In Data and Information in Online Environments: Second EAI International Conference, DIONE 2021, Virtual Event, March 10–12, 2021, Proceedings (Vol. 378, p. 431). Springer Nature.

[4] Yu, X. (2020). On the Construction of Integrated English Course with the Purpose of Eliminating" Water Course" and Creating" Gold Course". Theory and Practice in Language Studies, 10(11), 1447-1452.

[5] Zeng, Z. (2019). Outcome-Based Approach to Teaching Students Comprehensive English in China: From" Golden Course" to" Golden Lessons". English Language Teaching, 12(12), 112-118.

[6] Lv, N. (2023). Thoughts on the Construction of" Golden Course" in Business English under the Mixed Online and Offline Teaching Mode. The Educational Review, USA, 7(4), 483-486.

[7] Zhaodan, T. A. I. (2021). “Advanced English” Curriculum Design based on “Gold Course” Standard and Cross-cultural Perspective. The Theory and Practice of Innovation and Enntrepreneurship, 4(3), 27.

[8] Dobricki, M., Evi-Colombo, A., & Cattaneo, A. (2020). Situating vocational learning and teaching using digital technologies-a mapping review of current research literature. International journal for research in vocational education and training, 7(3), 344-360.

[9] Rohimajaya, N. A., & Hamer, W. (2023). Merdeka curriculum for high school english learning in the digital era. KLAUSA (Kajian Linguistik, Pembelajaran Bahasa, dan Sastra), 7(1), 1-8.

[10] Nasihin, A. (2022). The Impact of Using English Curriculum Design Based on Industry Needs in English Teaching on Vocational School to Improve Students English Skill for Industry Standard Working Communication. Education Quarterly Reviews, 5(1), 138-144.

[11] Song, S. J., Tan, K. H., & Awang, M. M. (2021). Generic digital equity model in education: Mobile-assisted personalized learning (MAPL) through e-modules. Sustainability, 13(19), 11115.

[12] Zhao, M., & Wang, L. (2022). Research on the Optimization and Allocation Management of Teaching Resources for English Teaching. Wireless Communications and Mobile Computing, 2022(1), 1182197.

[13] Karademir, T., Alper, A., Soğuksu, A. F., & Karababa, Z. C. (2021). The development and evaluation of self-directed digital learning material development platform for foreign language education. Interactive Learning Environments, 29(4), 600-617.

[14] Costa, C., Bhatia, P., Murphy, M., & Pereira, A. L. (2023). Digital education colonized by design: Curriculum reimagined. Education Sciences, 13(9), 895.

[15] Huang, Y. (2020, January). Research on the “Gold Course” Model of English Film and Television Appreciation Course Based on the Production-Oriented Approach. In 5th International Conference on Economics, Management, Law and Education (EMLE 2019) (pp. 1160-1163). Atlantis Press.

[16] Zhang, X. (2025). Optimization of Teaching Quality Improvement Path Based on Recursive Algorithm in the Construction of Higher Vocational English Gold Classes under the Background of Industry-Teaching Integration. J. COMBIN. MATH. COMBIN. COMPUT, 127, 6457-6472.