Digital Intelligence Platform Empowers Teaching Reform and Practice of English Listening and Speaking Courses in Higher Vocational Colleges: A Case Study of iFLYTEK FIF Oral Training System and Intelligent Agent

Song Zhengfang
Shandong Vocational College of Information Technology, Shandong, China

Abstract: This study investigates the application effects and practical paths of integrating iFLYTEK FIF Oral Training System and Intelligent Agent in the teaching reform of English listening and speaking courses in higher vocational colleges. A quasi-experimental research design combined with questionnaire surveys and semi-structured interviews was implemented. Two parallel classes were selected for a 8-week comparative teaching intervention. The experimental class adopted a blended teaching model supported by the dual platforms, while the control class received traditional teacher-centered instruction. Data were collected through pre- and post-proficiency tests, learning behavior records from the platforms, and student feedback questionnaires. Statistical analysis reveals that the integrated application of dual platform improves students' English listening and speaking proficiency, learning engagement, and autonomous learning ability. This research provides an empirically validated teaching model for the digital transformation of vocational English education.

Keywords: FIF Oral Training System; Intelligent Agent; Higher Vocational English; Listening and Speaking Courses; Teaching Reform

1. Introduction

1.1 Research Background and Significance

Higher vocational education occupies a core position in China's modern vocational education system, focusing on cultivating high-quality technical and skilled talents for industrial development. English listening and speaking ability has become an essential professional competence for graduates in the context of economic globalization and international industrial cooperation. Traditional English listening and speaking teaching in higher vocational colleges faces multiple persistent challenges. Large class sizes make it difficult for teachers to provide individualized guidance to each student, resulting in uneven learning progress among students with different proficiency levels. Limited classroom time restricts students' opportunities for oral practice, and most speaking activities remain at the level of mechanical repetition rather than authentic communication. The single evaluation method centered on final exams fails to reflect students' real learning process and practical application ability.

The rapid development of digital intelligence technology has brought new opportunities for the reform of vocational education. Artificial intelligence, speech recognition, and natural language processing technologies have been gradually applied to language teaching, breaking the limitations of time and space in traditional teaching. iFLYTEK, as a leading enterprise in China's artificial intelligence industry, has developed the FIF Oral Training System and Intelligent Agent specifically for English language teaching. The FIF Oral Training System provides massive authentic language resources and intelligent speech evaluation functions, while the Intelligent Agent realizes real-time interactive dialogue, personalized learning recommendation, and intelligent tutoring. The integration of these two platforms can effectively make up for the shortcomings of traditional English listening and speaking teaching, improve teaching efficiency and quality, and better meet the personalized learning needs of vocational college students.

This research has important practical significance for promoting the digital

transformation of higher vocational English teaching. It explores how to deeply integrate digital intelligence platforms with curriculum teaching, reconstruct teaching processes, and innovate teaching models. The research results can provide reference for other higher vocational colleges to carry out English listening and speaking teaching reform, and help cultivate more technical and skilled talents with international communication capabilities to adapt to the needs of industrial upgrading and high-quality economic development.

1.2 Domestic and International Research Status

Domestic scholars have conducted extensive research on English listening and speaking teaching in higher vocational colleges. Wang (2011) explored the practical teaching of English listening and speaking courses and pointed out that practical teaching links are crucial for improving students' practical application ability. Zhang (2003) analyzed the classroom organization of English listening and speaking courses and proposed specific strategies for optimizing classroom teaching activities. Zhang (2013) studied the integrated teaching model of English listening and speaking courses, emphasizing the integration of listening and speaking skills training. Feng (2022) conducted research on blended teaching in English audio-visual and speaking courses in higher vocational colleges and found that blended teaching can effectively improve students' learning enthusiasm and teaching effects.

In recent years, some scholars have begun to pay attention to the application of digital intelligence technology in vocational English teaching. Xie (2025) studied the teaching strategies of higher vocational mathematics courses under the background of the digital intelligence era, providing reference for the digital transformation of vocational education courses. Xu (2013) explored the construction of networked teaching platforms for higher vocational English courses and analyzed the functions and application effects of network platforms. However, most existing studies focus on the application of a single digital platform, and there is relatively little research on the integrated application of oral training systems and intelligent agents. The research on the specific application effects and practical paths of iFLYTEK's FIF series platforms in higher vocational English listening and speaking teaching is still insufficient.

International research on artificial intelligence in language teaching started earlier. Foreign scholars have conducted in-depth research on intelligent tutoring systems, speech recognition technology, and adaptive learning systems in language education. They have verified that AI technology can provide personalized learning experiences, improve students' language proficiency, and enhance learning motivation. However, due to differences in educational systems, teaching objectives, and student groups between China and foreign countries, foreign research results cannot be directly applied to China's higher vocational English teaching practice. It is necessary to conduct targeted research combined with the actual situation of China's higher vocational education.

1.3 Research Questions and Objectives

This research focuses on two core issues. The first is whether the integrated application of iFLYTEK FIF Oral Training System and Intelligent Agent can effectively improve the English listening and speaking proficiency of higher vocational college students. The second is what kind of teaching model and implementation strategies should be constructed to give full play to the advantages of the dual platforms in English listening and speaking teaching.

The main objectives of this research are as follows. To verify the effectiveness of the dual platforms in improving students' English listening and speaking ability, learning interest, and autonomous learning ability through empirical research. To construct a scientific and feasible teaching model of English listening and speaking courses supported by the FIF Oral Training System and Intelligent Agent. To summarize the practical experience and existing problems in the application process, and put forward targeted improvement suggestions to provide reference for the teaching reform of higher vocational English courses.

2. Theoretical Basis of Digital Intelligence Platform Empowering Higher Vocational English Listening and Speaking Teaching

2.1 Constructivist Learning Theory

Constructivist learning theory holds that knowledge is not passively received by learners but actively constructed by learners through interaction with the environment. Learners are

the main body of the learning process, and teachers play the role of guides and facilitators. Constructivism emphasizes learning in authentic situations and believes that learning occurs in specific social and cultural contexts.

The FIF Oral Training System and Intelligent Agent provide a good learning environment for constructivist learning. The platforms contain a large number of authentic language materials from real life and workplace scenarios, creating a realistic language communication environment for students. Students can conduct independent listening and speaking practice on the platforms, and construct their own language knowledge system through repeated practice and experience. The Intelligent Agent can conduct real-time interactive dialogues with students, simulate various communication scenarios, and guide students to use language knowledge to solve practical problems. Teachers can use the learning data generated by the platforms to understand students' learning situation, provide targeted guidance, and help students complete knowledge construction.

2.2 Blended Learning Theory

Blended learning theory combines the advantages of traditional face-to-face teaching and online learning. It emphasizes the organic integration of offline classroom teaching and online independent learning, giving full play to the leading role of teachers and the main role of students. Blended learning can break the limitations of time and space in traditional teaching, expand learning channels, and improve learning efficiency.

The integrated application of the FIF Oral Training System and Intelligent Agent provides strong support for blended learning. Students can complete pre-class preview, after-class review, and extended learning through the online platforms at any time and any place. The platforms can automatically record students' learning process and learning effects, providing data support for teachers to carry out targeted teaching. Offline classroom teaching focuses on solving the problems encountered by students in online learning, organizing interactive activities such as group discussions and role-playing, and deepening students' understanding and application of knowledge. The combination of online and offline teaching forms a complete learning closed loop, which can effectively improve the quality of English listening and speaking teaching.

2.3 Personalized Learning Theory

Personalized learning theory emphasizes that teaching should be carried out according to the individual differences of learners, meet the different learning needs of learners, and promote the all-round development of each learner. Each student has different cognitive styles, learning foundations, and learning interests, so a unified teaching model cannot meet the learning needs of all students.

The Intelligent Agent has powerful adaptive learning functions, which can realize personalized learning for students. The agent can evaluate students' English listening and speaking level through initial tests, and formulate personalized learning plans and push targeted learning resources according to students' actual situation. For students with weak foundation, the agent will push more basic listening and speaking exercises and provide detailed guidance and correction. For students with good foundation, the agent will push more challenging learning content and expand their knowledge. The FIF Oral Training System can also provide personalized speech evaluation, point out the pronunciation, intonation, and fluency problems of students in oral expression, and help students improve their oral English level in a targeted manner.

3. Teaching Reform Design of Higher Vocational English Listening and Speaking Courses Empowered by Digital Intelligence Platform

3.1 Teaching Objective Reconstruction

The traditional teaching objectives of English listening and speaking courses in higher vocational colleges pay too much attention to the mastery of language knowledge, ignoring the cultivation of students' practical application ability and professional quality. The teaching reform supported by digital intelligence platforms needs to reconstruct teaching objectives around the training objectives of higher vocational education.

The reconstructed teaching objectives include three levels. The basic level objective is to enable students to master basic English listening and speaking skills, understand daily English conversations, and conduct simple oral communication. The improvement level

objective is to enable students to understand English conversations and speeches in professional-related scenarios, and express their opinions and ideas fluently in English. The professional level objective is to enable students to use English to complete professional-related communication tasks, such as job interviews, business negotiations, and technical exchanges, and have good cross-cultural communication awareness and ability.

The teaching objectives are formulated in combination with the professional characteristics of students. For example, for students majoring in international trade, the teaching objectives focus on cultivating their English communication ability in business scenarios. For students majoring in nursing, the teaching objectives focus on cultivating their English communication ability in medical scenarios. The FIF Oral Training System and Intelligent Agent provide a large number of professional-related language resources, which can effectively support the realization of professional level teaching objectives.

3.2 Teaching Content Optimization

The traditional teaching content of English listening and speaking courses is relatively single, mainly based on textbooks, and lacks connection with real life and workplace practice. The teaching reform needs to optimize the teaching content, integrate digital resources, and build a diversified teaching content system.

The optimized teaching content consists of three parts. The basic part is the content specified in the textbook, which is the foundation of students' English learning. The extended part is the authentic language resources provided by the FIF Oral Training System, including daily conversations, news reports, movie clips, and workplace dialogues. These resources are updated in real time, which can broaden students' horizons and improve their language perception. The personalized part is the learning content pushed by the Intelligent Agent according to students' learning situation and professional needs. The agent can generate targeted exercises and learning materials based on students' weak points, helping students make up for their deficiencies.

The teaching content is organized around themes and tasks. Each unit takes a specific theme as the core, and designs a series of listening and speaking tasks. Students complete the tasks through online independent learning and offline classroom interaction, and gradually improve their English listening and speaking ability. The teaching content pays attention to the integration of ideological and political education, incorporates Chinese cultural elements, and guides students to tell Chinese stories well in English.

3.3 Teaching Model Innovation

The traditional teacher-centered teaching model cannot meet the needs of modern vocational education. The teaching reform needs to innovate the teaching model, build a student-centered blended teaching model supported by the dual platforms, and realize the transformation from "teacher teaching" to "student learning".

The teaching model is divided into three stages: pre-class, in-class, and after-class. In the pre-class stage, teachers release learning tasks and related resources on the FIF Oral Training System. Students complete pre-class preview tasks independently, including listening exercises, vocabulary learning, and text reading. The Intelligent Agent provides real-time guidance and answers to students' questions encountered in the preview process. Teachers can view students' preview situation through the platform's background data, understand the problems existing in students' learning, and adjust the in-class teaching content in time.

In the in-class stage, teachers organize various interactive teaching activities based on students' pre-class learning situation. First, teachers explain the key and difficult points of the unit and answer students' common questions. Then, students carry out group discussions, role-playing, and situational dialogues using the functions of the Intelligent Agent. The agent can act as a dialogue partner, participate in students' communication activities, and provide real-time evaluation and feedback. Finally, teachers summarize the learning content of this class and assign after-class learning tasks.

In the after-class stage, students complete the listening and speaking exercises assigned by teachers on the FIF Oral Training System, and conduct extended learning according to their own situation. The Intelligent Agent tracks students' learning process, provides personalized learning suggestions, and urges students to complete learning tasks on time. Teachers check students' after-class learning situation through

the platform, communicate with students who have learning difficulties, and provide individual guidance.

3.4 Evaluation System Reform

The traditional evaluation system of English listening and speaking courses is mainly based on final exams, which cannot comprehensively and objectively evaluate students' learning process and practical ability. The teaching reform needs to reform the evaluation system, build a diversified process evaluation system, and realize the combination of process evaluation and summative evaluation.

The diversified evaluation system includes three parts: platform automatic evaluation, teacher evaluation, and student mutual evaluation. Platform automatic evaluation accounts for 40% of the total score, mainly evaluating students' online learning time, completion of learning tasks, and performance in listening and speaking exercises. The FIF Oral Training System can automatically evaluate students' pronunciation, intonation, and fluency, and generate detailed evaluation reports. Teacher evaluation accounts for 40% of the total score, mainly evaluating students' performance in classroom activities, task completion, and progress in learning. Student mutual evaluation accounts for 20% of the total score, mainly evaluating students' performance in group activities and cooperative learning.

The evaluation content focuses on students' practical application ability and comprehensive quality. In addition to evaluating students' listening and speaking proficiency, it also evaluates students' learning attitude, cooperative spirit, and cross-cultural communication ability. The evaluation results are fed back to students in time, helping students understand their own learning situation and adjust their learning strategies. Teachers can also use the evaluation results to reflect on their teaching work and continuously improve teaching quality.

4. Teaching Practice and Effect Analysis

4.1 Research Objects and Methods

The research objects are 86 freshmen from two parallel classes majoring in e-commerce. The two classes have similar English proficiency levels through pre-test, and there is no significant difference in their average scores. Class A with 43 students is the experimental class, and Class B with 43 students is the control class. The teaching experiment lasts for 8 weeks, with 2 class hours per week.

The research adopts a quasi-experimental research method. The experimental class adopts the blended teaching model supported by the FIF Oral Training System and Intelligent Agent, while the control class adopts the traditional teacher-centered teaching model. The same teacher teaches both classes, using the same textbook and teaching progress. In addition to the quasi-experimental method, the research also uses questionnaire surveys and semi-structured interviews to collect students' feedback on the teaching model and the dual platforms.

The research tools include pre- and post-English listening and speaking proficiency tests, student learning behavior records from the platforms, and student feedback questionnaires. The pre- and post-tests are designed according to the Higher Vocational College English Proficiency Test (Pretco) Level B listening and speaking standards, with a full score of 100. The questionnaire includes 20 items, covering students' learning interest, learning engagement, autonomous learning ability, and satisfaction with the platforms. The data collected are analyzed using SPSS 26.0 statistical software.

4.2 Teaching Implementation Process

The teaching implementation is carried out strictly according to the designed blended teaching model. In the first week of the experiment, teachers introduce the FIF Oral Training System and Intelligent Agent to the students of the experimental class, teach them how to use the platforms, and conduct an initial English proficiency test. The Intelligent Agent formulates personalized learning plans for each student based on the test results.

In the pre-class stage of each unit, teachers release learning tasks and related resources on the FIF Oral Training System 2 days in advance. The learning tasks include listening to the unit text, learning new words and expressions, and completing basic listening exercises. Students complete the tasks independently on their mobile phones or computers, and can ask the Intelligent Agent for help at any time when encountering problems. Teachers check students' learning situation through the platform background every day, and send reminders to students who have not completed the tasks.

In the in-class stage, teachers first spend about

10 minutes answering the common questions raised by students in the pre-class learning. Then, students are divided into groups of 4-5 people to carry out situational dialogue and role-playing activities. Each group uses the Intelligent Agent to simulate different communication scenarios, such as product introduction, customer consultation, and order negotiation. The agent participates in the dialogue as a customer or a colleague, and provides real-time evaluation and correction of students' oral expression. After the group activities, each group sends a representative to demonstrate the dialogue results, and teachers and other students make comments. Finally, teachers summarize the key points of the unit and assign after-class tasks.

In the after-class stage, students complete the listening and speaking exercises assigned by teachers on the FIF Oral Training System, including imitation reading, dialogue practice, and topic retelling. The system automatically evaluates students' exercises and generates detailed error analysis reports. Students can view the reports and correct their mistakes in time. The Intelligent Agent pushes extended learning resources to students according to their exercise performance, such as related English videos and articles. Teachers communicate with students with poor performance individually through the platform, and provide targeted guidance and help.

4.3 Teaching Effect Data Analysis

After 8 weeks of teaching experiment, post-tests are conducted on both the experimental class and the control class. The independent samples t-test is used to compare the post-test scores of the two classes. The results showed that the experimental class scored significantly higher ($M=78.56$, $SD=6.23$) than the control class ($M=69.23$, $SD=7.15$), $t(84)=5.87$, $p<0.001$. This indicates that there is a significant difference in the post-test scores between the two classes. The English listening and speaking proficiency of the experimental class is significantly higher than that of the control class.

The paired samples t-test is used to compare the pre-test and post-test scores of each class. The results show that the average pre-test score of the experimental class is 62.35 ($SD=5.82$), and the average post-test score is 78.56 ($SD=6.23$). The t-value is 12.45, and the p-value is less than 0.001, indicating that the experimental class has made significant progress in English listening and speaking proficiency. The average pre-test score of the control class is 61.89 ($SD=5.76$), and the average post-test score is 69.23 ($SD=7.15$). The t-value is 7.23, and the p-value is less than 0.001, indicating that the control class has also made some progress, but the progress is much smaller than that of the experimental class.

The analysis of students' learning behavior records from the platforms shows that the students of the experimental class have high learning enthusiasm. The average weekly online learning time is 3.5 hours, and the completion rate of learning tasks is 95.3%. The Intelligent Agent has conducted more than 2000 interactive dialogues with students, and the average number of dialogues per student is 46.5. These data indicate that the dual platforms have effectively stimulated students' learning interest and improved their learning engagement.

4.4 Student Feedback and Teacher Reflection

After the experiment, a questionnaire survey is conducted on the students of the experimental class. A total of 43 valid questionnaires are collected, with a recovery rate of 100%. The survey results show that 88.4% of the students are satisfied with the blended teaching model supported by the dual platforms. 90.7% of the students believe that the FIF Oral Training System has helped them improve their listening and pronunciation skills. 86.0% of the students believe that the Intelligent Agent has provided them with more opportunities for oral practice and improved their oral expression fluency. 83.7% of the students believe that the personalized learning function of the platforms has helped them make up for their learning deficiencies.

Semi-structured interviews are conducted with 10 randomly selected students from the experimental class. Most students say that they used to be afraid of speaking English in class because they were worried about making mistakes. Now they can practice oral English freely with the Intelligent Agent without feeling nervous, and their oral English level has improved significantly. Some students mention that the real-time evaluation function of the platforms allows them to know their own problems in time and correct them quickly. A few students say that sometimes the intelligent agent cannot understand their expressions accurately, and the response is not natural enough.

Teachers reflect that the application of the dual platforms has greatly reduced their workload of correcting homework and evaluating students' oral performance. The learning data generated by the platforms provide a scientific basis for teachers to understand students' learning situation and carry out targeted teaching. However, teachers also find that some students have poor autonomous learning ability and cannot complete online learning tasks on time without supervision. In the future teaching, it is necessary to strengthen the guidance and supervision of students' autonomous learning, and further optimize the teaching design to improve the effectiveness of online learning.

5. Conclusion

This research constructs a blended teaching model of higher vocational English listening and speaking courses supported by iFLYTEK FIF Oral Training System and Intelligent Agent, and verifies the effectiveness of the model through empirical research. The results show that the integrated application of the dual platforms can significantly improve students' English listening and speaking proficiency, learning interest, and autonomous learning ability. The model effectively solves the problems existing in traditional English listening and speaking teaching, such as large class sizes, insufficient individual guidance, and limited oral practice opportunities.

The research has certain limitations. The research objects are only from one major of one higher vocational college, and the sample size is relatively small. The research results may not be fully applicable to other majors and other regions. The teaching experiment only lasts for 8 weeks, and the long-term effects of the teaching model need to be further verified.

Future research can expand the sample size and research scope, and conduct in-depth research on the application of the dual platforms in different majors and different grades of higher vocational colleges. It can also further optimize the functions of the intelligent agent, improve the accuracy of speech recognition and the naturalness of dialogue, and provide better support for students' personalized learning. The research results can provide more comprehensive reference for the digital transformation of higher vocational English teaching, and promote the high-quality development of vocational education.

Acknowledgements

Shandong Provincial Society of Vocational and Technical Education, Digital Intelligence Platform Empowering Teaching Reform and Practice of English Listening and Speaking Courses in Higher Vocational Colleges , WYKT2024Y019.

References

[1] Ministry of Education of the People's Republic of China. Action Plan for Digital Transformation of Vocational Education (2023-2025)[Z]. 2023.

[2] Zhang L. Practical Path of Artificial Intelligence Empowering Blended Teaching of Higher Vocational English[J]. Vocational Education, 2025, 24(03): 45-49.

[3] Chen M. Research on Application of Intelligent Speech Technology in Higher Vocational English Listening and Speaking Teaching[J]. Journal of Foreign Language Educational Technology, 2024(02): 89-94.

[4] Li N. Construction of Process Evaluation System of Higher Vocational English Supported by Digital Intelligence Platform[J]. Chinese Vocational and Technical Education, 2026(05): 78-82.

[5] Wang Y. Research on Personalized Teaching Mode of Higher Vocational English under the Background of AI + Education[J]. Higher Vocational Education (Journal of Tianjin Vocational Institute), 2025, 34(01): 67-71.

[6] Wang, W. P. Application of task-based teaching method in English audio-visual and speaking courses in higher vocational colleges: Taking English audio-visual and speaking courses in teaching ability competition as an example[J]. Knowledge Window (Teacher Edition), 2022(10): 126-128.

[7] Feng, D. Research on blended teaching in English audio-visual and speaking courses in higher vocational colleges[J]. Industrial & Science Tribune, 2022, 21(9): 192-193.

[8] Xie, D. Research on teaching strategies of higher vocational mathematics courses under the background of digital intelligence era[J]. Digital Communication World, 2025(10): 250-252.

[9] Liu, T., Ouyang, C. S. Teaching exploration of telling Hunan stories well in higher vocational business English listening and

speaking courses[J]. Folk Story Selection, 2024(20): 262-264.
[10] Yang, H. J. On the teaching of business English listening and speaking courses in higher vocational colleges. Campus English, 2014(32): 68.
[11] Yang, S. G. Research on innovation of English listening and speaking course teaching mode in higher vocational colleges [J]. Think Tank Era, 2020(31): 213.
[13] Wu, Y. Analysis of project-based teaching in higher vocational English courses: Taking nursing English listening and speaking as an example[J]. Journal of Changchun Education Institute, 2014, 30(13): 118-119.
[14] Li, N. Research on the application of authentic assessment in English audio-visual and speaking course teaching in higher vocational colleges[J]. Hunan University of Science and Technology, 2011.
[16] Xu, P. P. Exploration on the construction of networked teaching platform for higher vocational English courses[J]. Time Education, 2013(7): 181.
[17] Yan, D. P. Discussion on basic theory of English audio-visual and speaking courses in higher vocational colleges[J]. Journal of Hubei Radio & Television University, 2010, 30(6): 28-29.
[18] Anonymous. iFLYTEK appears at China Unicom Partner Conference to jointly expand the new ecology of digital intelligence integration[J]. Computer Products & Circulation, 2025(9): 13-15.
[19] Lei, Z. C. iFLYTEK Speech Cloud Platform[J]. New Computer, 2010(12): 89.
[20] Lin, L. Representative of capability platforms: iFLYTEK[J]. Entrepreneur, 2012(6): 76.