International Journal of Continuing Engineering Education and Life-Long Learning

ISSN online: 1741-5055 - ISSN print: 1560-4624
https://www.inderscience.com/ijceell

Research on interactive audio-visual-oral teaching in college English

Haixia Yin, Liming Wang

DOI: 10.1504/IJCEELL.2025.10072638

Article History:

Received:	11 February 2025
Last revised:	07 May 2025
Accepted:	03 June 2025
Published online:	10 October 2025

Research on interactive audio-visual-oral teaching in college English

Haixia Yin*

School of English Language and Culture,
Xi'an Fanyi University,
Xi'an 710000, Shaanxi, China
Email: fiona_xxn@163.com
*Corresponding author

Liming Wang

Data Intelligence Business Group,
Lenovo Group,
Xi'an 710000, Shaanxi, China
Email: wanglm19@lenovo.com

Abstract: With the advancement of information and communication technologies and globalisation, English has become essential for accessing global information and promoting socio-economic development. This highlights the importance of effective English education in colleges. An interactive audio-visual-oral teaching mode (TM) is proposed to enhance students' language skills through active participation. Supported by virtual reality (VR) and artificial intelligence (AI), this approach creates immersive 3D learning environments, improving students' interest and self-directed learning. VR enriches audiovisual teaching with realistic sensory experiences, while AI personalises instruction. Together, they offer innovative tools for modern college English education, fostering both theoretical understanding and practical communication skills.

Keywords: English interactive teaching; audiovisual classroom; virtual reality; artificial intelligence.

Reference to this paper should be made as follows: Yin, H. and Wang, L. (2025) 'Research on interactive audio-visual-oral teaching in college English', *Int. J. Continuing Engineering Education and Life-Long Learning*, Vol. 35, No. 8, pp.129–144.

Biographical notes: Haixia Yin is an Associate Professor and holds a Master's degree from Xidian University. She is currently a full-time teacher at the School of English Language and Culture, Xi'an Fanyi University, where she teaches courses related to College English. Her work primarily involves teaching and research in public English, with specific interests in college English teaching. Up to now, she has published more than ten academic papers in relevant fields.

Liming Wang received his Bachelor's degree from Xi'an Technological University. He is currently employed at Lenovo (Beijing) Co., Ltd., where his work focuses on artificial intelligence. He has technical expertise and practical experience across multiple fields, including the implementation of enterprise digital applications, industrial AR quality inspection, and the application of enterprise large models.

1 Introduction

More and more people need to have cross-border exchanges in their daily lives, or some enterprises need more people who can use English skilfully when conducting cross-border trade and activities. However, the traditional college English teaching model still faces significant challenges: firstly, one-way knowledge imparting leads to weak listening and speaking abilities of students, making it difficult to adapt to cross-cultural communication needs; secondly, existing technological means have not fully simulated the real language environment, which hinders the cultivation of students' contextualised communication skills; thirdly, lagging teaching feedback and lack of personalised guidance hinder the improvement of learning efficiency. In this context, exploring interactive teaching models that integrate cutting-edge technologies has become a key breakthrough in breaking through the bottleneck of English education.

The VR-AI-TM applied in this article has achieved a dual breakthrough in teaching paradigms through technological innovation. On the one hand, the three-dimensional immersive scenes constructed by VR can simulate real contexts such as business negotiations and academic conferences, solving the problem of traditional classroom situations being monotonous; on the other hand, speech recognition and natural language processing systems empowered by AI can analyse students' pronunciation, grammar, and logical structure in real-time, providing personalised error correction and adaptive learning paths. This technological integration not only strengthens the multimodal input and output channels of 'sight listen speak', but also reconstructs the cognitive closed loop of language learning through the dual mechanisms of human-computer interaction and life collaboration.

This article aims to construct a student-centred English audio-visual oral teaching model, and its core innovation lies in the systematic integration of VR scene rendering, AI semantic analysis, and multimodal interaction technology, breaking the temporal and spatial limitations of traditional classrooms. By proposing a collaborative filtering recommendation algorithm based on knowledge graph, combined with student behaviour data such as login duration and test scores, dynamic adaptation of learning content and networked knowledge association can be achieved. The study verifies the teaching effectiveness through a two-stage comparative experiment, and the results show that students' listening and speaking abilities are significantly improved under the new mode, and teaching efficiency is also significantly optimised. The main contribution of this article is to provide a scalable technology empowerment framework for the field of language education. Its innovation is reflected in the design of multimodal interaction mechanisms and adaptive recommendation algorithms, exploring new methodological paths for future educational technology integration.

2 Related work

In today's society, English is still the preferred language for cross-border economic activities and residents' cross-border exchanges, which also leads to the increasing importance of English TM. Ahmadi and Reza (2018) mainly explored the use of various emerging information technologies in English education and English learning. Through the analysis of the development background and current situation of these emerging technologies, he determined that various technologies can be better used in English

education. At the same time, he also made an in-depth analysis of some examples of the use of these technologies in English teaching. Khotimah et al. (2019) explored the feelings of English teachers and students in this learning mode of autonomous English learning. Through the analysis of some viewpoints of teachers and students in autonomous English learning, he explored how teachers can promote students' autonomous English learning and determined that middle school students can learn independently at the current stage. Due to the insufficient understanding of teachers, the effect of autonomous learning often fails to reach the expected goal. Villafuerte and Mosquera (2020) made an in-depth study of English teaching and the local education industry in a certain region and described some deficiencies in the current English TM in the region. He discussed how to reform the English TM in this region through some policies of local governments and schools and determined the gap between the optimised English TM in this region and the previous TM in terms of teaching effect and teaching efficiency. Ibrahim (2017) explored the differences of various teaching methods in English teaching in a certain region and investigated the teaching methods of several English teachers in the region. He determined that the TM of using English games for auxiliary education had a better teaching effect. Ghaemi and Golshan (2017) explored the role of new social platforms and media in English teaching. He determined the feasibility and reliability of this English TM combining the new social platform and media. Turan and Akdag-Cimen (2020) made an in-depth exploration on the use of flipped classroom (FC) in English TM. Through reviewing the literature on the use of the current FC TM in English teaching, he determined that FC in English teaching can not only improve teaching efficiency but also optimise the teaching architecture. Macaro et al. (2018) evaluated the TM used by English teachers in higher education. After a comprehensive description of the current importance of English teaching and the TM, he determined that the current English is more learned by students as a second language at the stage of higher education. This TM has a very negative impact on improving students' English level. Although many researchers in the field of education have studied the optimisation of English TM, there is still no English TM based on emerging technologies.

Researchers in the field of education have begun to pay attention to VR and AI technology and discussed the use of VR and AI technology in English teaching. Kim et al. (2019) studied and predicted the future English TM. Through the analysis of the development process of robot and AI technology at the current stage, he determined the feasibility of the use of AI and robot technology in English teaching. Fahimirad and Kotamjani (2018) explored a new TM, which was a new mode based on AI technology. Through in-depth analysis of various research achievements of current AI technology, he determined the feasibility of applying some algorithm models in AI technology to the construction of this TM. Shin (2021) studied various new educational applications combined with AI technology to develop a new English TM and determined the feasibility of this scheme. Tan (2020) studied an English teaching system combining AI technology and wireless network and verified the reliability of the system. Zhao (2021) proposed a new AI-based hybrid TM for college English and determined the superior performance of this TM. However, due to the limitation of the development progress of VR and AI technology, this kind of English TM based on VR or AI technology still cannot be well popularised.

With the help of in-depth analysis of VR and AI technology, this text determines the feasibility of VR and AI technology in college English education. At the same time, this

text sets up an interactive audio-visual-oral TM and classroom, which can effectively improve the teaching efficiency and effect of college English.

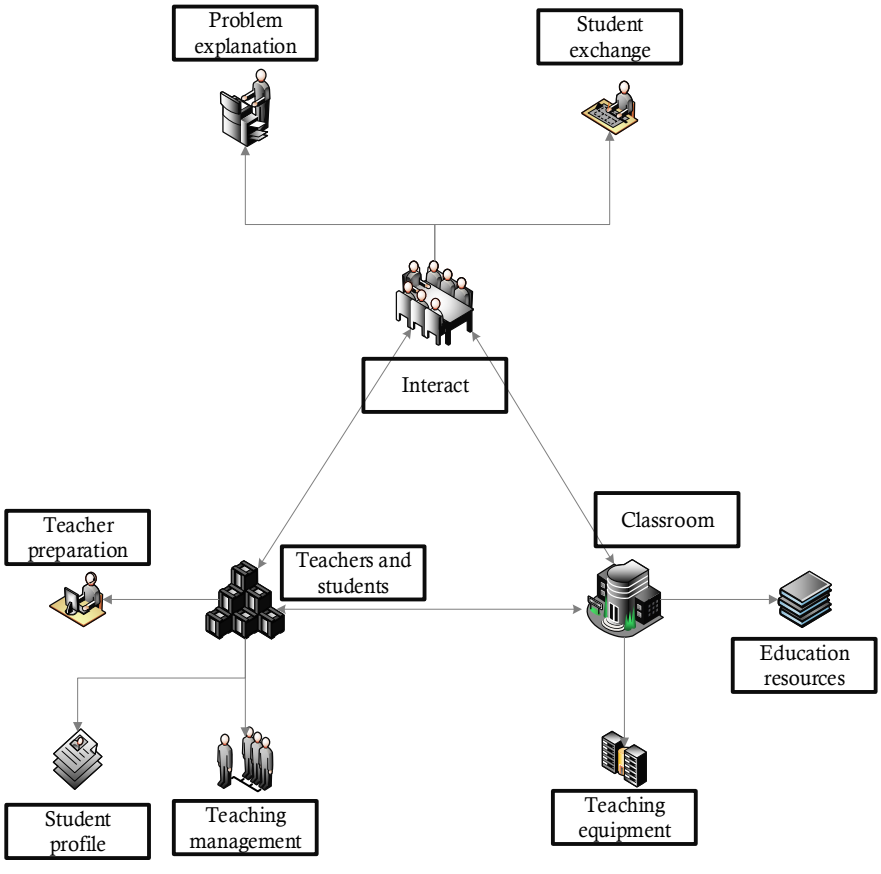
3 Interactive audio-visual-oral English TM

English is one of the most widely used and important languages at present. It plays a central communication role in international trade and commerce. Contract signing, negotiation, mail, etc., are inseparable from English. It is also the main term of standardised documents to ensure consistent global understanding. English also promotes cross-cultural cooperation and global marketing. From the perspective of language choices in different regions, a large number of people have received English education and are proficient in using English for communication, and English is also the preferred language in many regions. Therefore, people's proficiency in English is the basic requirement for cross-cultural or cross-border communication. If a person can communicate in English skilfully, the upper limit of his development can also be improved. At the same time, the increase in the number of people mastering English skills is also a necessary condition for the economic development of a region. Through a large number of talents who have mastered English skills to play their due role in various jobs and absorb excellent management or other technical concepts from other regions, they can better build the local economic system. Although this new TM also pays attention to grammar teaching, it also pays attention to the flexibility of teaching, making great efforts to improve the learning effect of students. On the other hand, English viewing, listening and speaking are also important in English teaching. The mastery of English viewing, listening, and speaking not only determines whether students can skilfully use English for oral communication, but also has a very important impact on English teaching efficiency. Therefore, the new interactive English TM has also been constructed in the audio-visual-oral classroom, which is mainly reflected by the communication between students and students in the English teaching classroom. Teachers are essential in interactive teaching, not only to impart knowledge but also to guide and promote students' learning. Teachers need to design enlightening activities to stimulate students' interest and initiative, answer questions, and provide necessary guidance. The structure of the interactive English TM proposed in this text is shown in Figure 1.

At present, the initial goal of English teaching is to focus on cultivating students' ability to use English in daily life, which does not mean that students can only accept the education of English grammatical structure or other relevant theoretical knowledge. It means that students need to accept more professional English teaching. At this time, the proposed interactive teaching mode (ITM) also provides new ideas for researchers in related fields of English education. In the ITM, language teaching is compared to a real-time interactive process, which requires efficient communication between teachers and students. The students can consult the teachers who are responsible for all kinds of problems arising in the interaction process. Teachers should teach students in accordance with their aptitude and answer their questions. Compared with the traditional TM, this TM can enhance students' learning interest and effect. For example, interactive teaching enables students to participate more actively in learning and increase interest through group discussions and role-playing. This model allows students to become passive to active, deepening their understanding and mastery of knowledge. Therefore, interactive

teaching significantly enhances students' interest in learning. The ITM of college English also has higher requirements for teaching content. In interactive teaching (IT), the teaching content not only needs to integrate teachers' cognition and theoretical knowledge of English at a deep level, but also needs to unify the oral expression form and grammar of English. Secondly, IT does not attach importance to the sequence of programs in teaching. As long as the expected teaching goals can be achieved, the sequence of programs in teaching can be more flexible. The procedures in teaching need to be formulated according to the provisions of English teaching in schools. At the same time, IT also pays more attention to the effective combination of teaching methods and the team building among students. Students can flexibly form various types of learning groups and can better find some problems in English communication by simulating and practicing a certain oral English communication situation in the group. First, through the communication within the group, the problem is preliminarily discussed, and finally, the teacher is asked to solve the problem. In a word, the ITM of college English provides a relatively good learning atmosphere and communication environment for teachers and students. At the same time, it also improves the teaching efficiency and teaching effect, so as to cultivate more high-quality talents in a shorter time.

Figure 1 Schematic structure of the new interactive English TM (see online version for colours)



4 VR and AI technology analysis

VR is a newly emerging application technology. Through the use of 3D image technology, multimedia technology, display technology, and other information technologies, it can create a highly realistic 3D perceptual virtual environment with computing equipment such as computers. It is currently widely used in film and television entertainment, education, design, and military related fields (Kavanagh et al., 2017). There are three main characteristics of current VR technology, namely immersion, interaction, and imagination (El Beheiry et al., 2019). Immersion mainly means that the three-dimensional perceptual environment constructed by the simulation system theory in the computer is more realistic. This allows users to devote themselves to the use of the three-dimensional perceptual environment, so that they cannot distinguish the virtual world from the real world. VR technology is a double-edged sword. Although immersing in VR technology can enhance people's use experience in the three-dimensional perceptual environment, it also causes users to be unable to distinguish where the virtual world is, thus causing certain harm to people's physical and mental health. The second is interaction, which generally refers to that users can interact with various virtual objects in the virtual world and get realistic feedback in a three-dimensional perception virtual world built by computers. For example, users can watch and move objects in the virtual world through their own vision, touch, and hearing. The establishment of the interaction function has greatly increased the scope of activities that people can do, which is also the foundation of VR technology. It helps people to visit and travel in the virtual world in various free time in places that cannot be visited in the real world. The last is imagination, which mainly means that the virtual world through three-dimensional perception has greatly expanded people's imagination. It can not only build various real objects and scenes in the virtual world but also build some objects and scenes that do not exist in the real world. This also makes VR technology have greater development space. The application fields of VR technology in life are shown in Figure 2.

On the other hand, AI technology is a technology that makes devices similar to human intelligence through in-depth analysis of human intelligence and attempts to find the essence of human intelligence (Davenport and Ronanki, 2018). AI technology refers to various methods and means that enable computer systems to simulate human intelligent activities, including learning, reasoning, perception, understanding, and autonomous decision-making capabilities. It covers multiple fields such as machine learning, deep learning, natural language processing, computer vision, and expert systems. It is widely used in scenarios such as autonomous driving, intelligent customer service, medical diagnosis, and financial analysis. It is an important driving force for the intelligent transformation and innovative development of various industries in society. AI technology has been widely used in current daily life. At present, AI technology can be divided into special AI and general AI according to its research direction in the application field. Among them, dedicated AI is a kind of AI technology with rapid research progress, and even has obtained some research results and applied in various fields. General AI is the initial R&D goal of AI technology, which is to build a general model with human like intelligence, so it involves many scientific and technological fields. The data contained therein is also more complex, and data modelling is also more difficult, so the development of general AI is still in its infancy at present. This text constructs a new ITM of audio-visual-oral teaching through various new algorithms and theories in VR and AI technology. Among them, VR mainly makes use of the three

characteristics of its technology to play a positive role in the interaction of English teaching, such as improving the interest in the teaching process and cultivating students' ability to learn independently. AI technology is responsible for the analysis and management of all kinds of English teaching materials, and can also accurately evaluate students' learning progress and learning effect. Artificial intelligence can collect and analyse learning data such as progress, grades, interaction records, etc., and construct portraits of students to comprehensively evaluate the learning effect. Its advantage lies in multi-dimensional objective evaluation, reducing human interference and improving accuracy, and providing personalised feedback and suggestions to help students understand their learning status and formulate efficient plans. The structure of AI technology in college English ITM is shown in Figure 3.

Figure 2 Schematic diagram of the application areas of VR technology in life (see online version for colours)

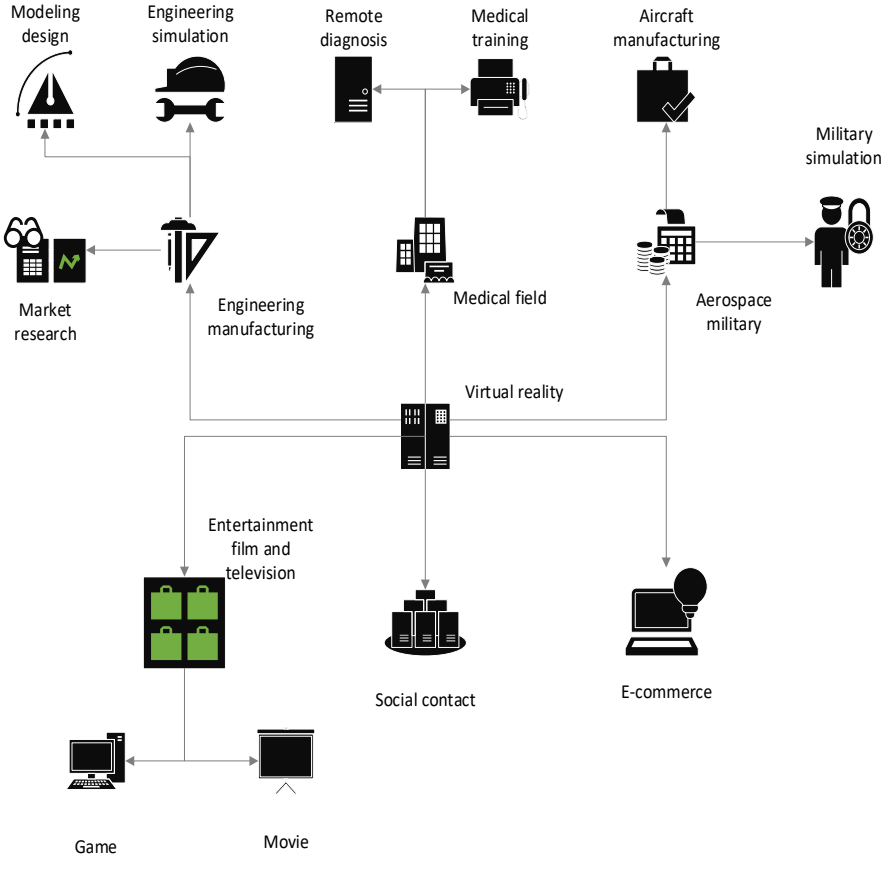
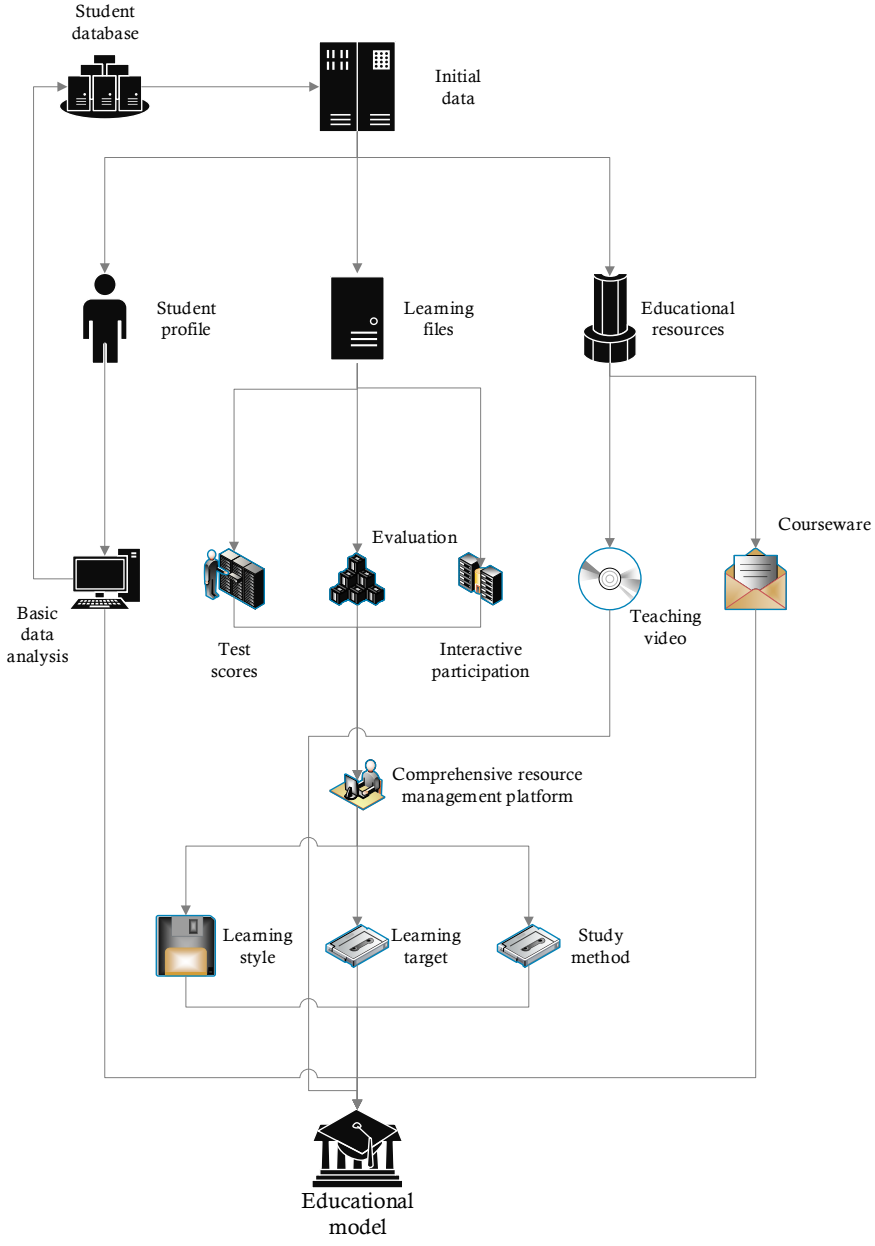


Figure 3 Schematic diagram of the structure of AI technology in the ITM of college English (see online version for colours)



In the interactive teaching of English listening, speaking, and watching in colleges, VR and AI technology can create an immersive and intelligent learning environment, break the time and space limitations of traditional classrooms, and enhance students' language practice opportunities and interactive experience. VR can stimulate students' learning interest and oral expression desire through real-life scene simulation, and enhance the naturalness of language input and output; AI can realise personalised learning

recommendations, intelligent evaluation, and instant feedback, and help students discover and correct problems in language use in a timely manner.

5 Algorithm analysis of AI technology

At present, the social and economic integration and the development of diversity and diversification of the world have greatly enhanced the status of English in the global multi-languages. Many regions have even adopted English as their preferred language of communication for education and use. Therefore, in the current social environment, people who can master English skills more skilfully also have an important means to obtain high-value resources and information in various fields. Therefore, the current emphasis on English education is increasing, and more and more new technologies are being used in English teaching. This text optimises college English TM to a certain extent through AI and VR technology. Through the use of various types of media resources in VR technology, a suitable classroom for English audio-visual-oral teaching can be constructed.

First, this text uses the linear regression algorithm model to classify the teaching resources in the new college English ITM. The operation is shown in formula (1).

$$H = q_0 + q_1 x_1 + q_2 x_1 \tag{1}$$

In the formula, H represents the error term of sample data in linear regression, while q_0 represents the offset term of data in the process of sample data calculation. q_1 and q_2 represent the weight of each data during the operation, while x_1 and x_2 represent different sample data. Then, the error of the specified data is calculated, as shown in formula (2).

$$H(x_m) = \sum_{i=0}^{n} q_i x_i \tag{2}$$

In the formula, m represents the number of sample data. Next, the loss function is expressed, and its operation is shown in formula (3).

$$y = q^T x_i + \varepsilon_i \tag{3}$$

ε_i mainly represents the error between sample data and output value. Then, the decision tree is used to control the background programs in the process of VR technology building a 3D perception world. First, it needs to calculate the information entropy, as shown in formula (4).

$$D = -\sum_{i=1}^{n} p_i \log_2 p_i \tag{4}$$

Among them, D represents the value of information entropy and is also an indicator to evaluate the purity of sample datasets. p_i represents the proportion of sample data of the i^{th} category in the dataset. Then, the information entropy increase rate is calculated, as shown in formula (5).

$$V = -\sum_{i=1}^{m} \frac{D_m}{D} \log_2 \frac{D_m}{D} \tag{5}$$

The next step is to calculate the probability of inconsistent sample categories. The lower the value, the higher the value of the sample. The calculation formula is shown in formula (6).

$$G = 1 - \sum_{k=1}^{n} p_i^2 \tag{6}$$

Finally, the optimal value is calculated by the above probability value, as shown in formula (7).

$$G(D) = \sum_{v=1}^{v} \frac{D_v}{D} G(D_v) \tag{7}$$

AI technology collects learning data such as login time, course duration, test scores, etc., and analyses it with machine learning algorithms to identify learning efficiency and patterns. Natural language processing technology analyses the speeches of students in the discussion forum to understand and master knowledge points. There are cases that show that AI can monitor and warn students' learning progress in real-time to improve evaluation efficiency.

These are the introduction of some algorithm models used in the construction of a new interactive college English audio-visual-oral TM based on VR and AI technology. These algorithm models not only greatly simplify the operation efficiency of the TM, but also provide students with better teaching services, so that students can learn better English in the same time.

Interactive teaching effectively improves students' interest in learning and the depth of knowledge mastery by stimulating students' active participation, strengthening immediate feedback and collaborative communication. In college English listening and speaking teaching, the interactive model can improve the actual language application ability and significantly increase the classroom activity. Similarly, in medical education, through interactive training of simulated surgery, students can master clinical skills faster; in the field of engineering, interactive experimental teaching has also been proven to effectively promote students' problem-solving ability and innovative thinking development. These successful cases show that interactive teaching has shown wide applicability and superior teaching effects in different disciplines, providing strong practical support for language teaching.

6 System application

The interactive audio visual oral teaching system (IAVOTS) based on VR and AI technology is an immersive interactive learning platform tailored for college English teaching. The system adopts a layered architecture design, mainly including four core modules.

- Module 1 is VR virtual scene generation, using Unity 3D engine and Unreal Engine to build a multi-context virtual environment, supporting students to enter the 3D interactive space through head-mounted devices for real-time dialogue practice.
- Module 2 is AI intelligent evaluation, integrating natural language processing technology. Based on BERT and transformer models, it provides real-time

evaluation of students' speech input in terms of pronunciation accuracy, grammatical logic, and fluency, and offers personalised error correction feedback.

- Module 3 is the personalised learning path recommendation module, which analyses students' learning data through machine learning algorithms, dynamically adjusts the difficulty of learning content, and pushes customised listening materials and oral tasks.
- Module 4 is an interactive teaching management platform, where teachers can create virtual classrooms through web or mobile devices, monitor student progress, initiate group collaboration tasks, and use data visualisation tools to track overall class performance.

The implementation details of the system functions are as follows:

1 VR virtual scene interaction

Scene construction: the 3D modelling technology is used to recreate real communication scenes, such as airport check-in counters, multinational company conference rooms, etc. Students can manipulate virtual characters to engage in situational dialogues.

Gesture and speech recognition: the leap motion gesture controller and Google speech to text API are used to select props through gestures and synchronise voice communication in a virtual environment.

2 AI intelligent evaluation system

Pronunciation score: based on the Kaldi speech recognition framework, compared with the standard pronunciation library (BBC Learning English), the score and label weak phonemes in pronunciation are quantified.

Semantic understanding: the dialogue flow technology is used to parse dialogue intentions, detect logical coherence, and prompt answers that deviate from the topic.

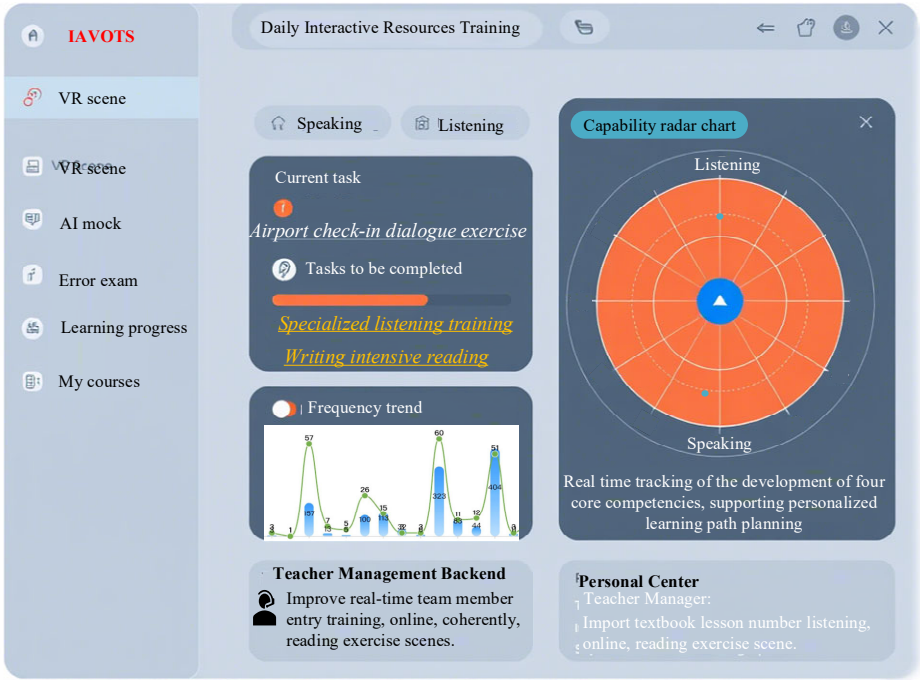
3 Personalised learning engine

Knowledge graph driven: a knowledge graph covering over 2,000 high-frequency English topics, linking vocabulary, sentence structures, and corresponding scenarios is built to achieve networked access to knowledge points.

Adaptive recommendation algorithm: the collaborative filtering algorithm is used to dynamically match learning resources based on similar student performance clusters (such as weak speaking group/excellent listening group).

The system interface is shown in Figure 4. In Figure 4, the left navigation bar of the student side main interface includes VR scenes, AI simulations, incorrect exams, learning progress, and the entrance to the course functions. The current task in the centre of the main interface is highlighted, and the remaining tasks are marked to be completed. The ability radar chart on the right features an orange fan-shaped real-time tracking of the development trajectory of two core abilities, listening and speaking, emphasising support for personalised learning path planning. The bottom teaching management area prompts to import textbook chapters to create listening/reading scenarios, presenting an intelligent and visual language training system architecture as a whole.

Figure 4 Screenshot of interactive audio visual oral teaching system (see online version for colours)



7 Experimental study on the interactive audio-visual-oral English teaching model based on VR and AI

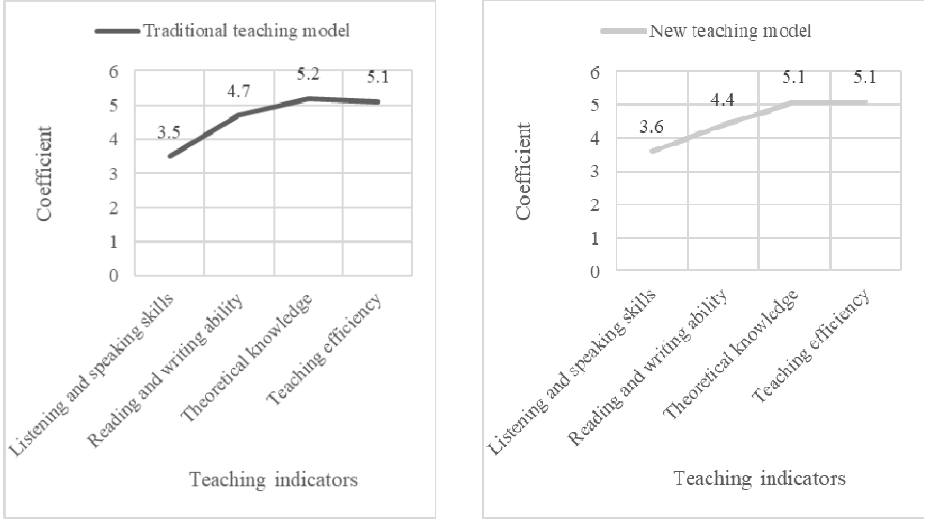
The mode of teaching generally refers to a relatively stable structure of educational activities in a specific educational environment under the guidance of the prescribed educational philosophy and teaching methods. It is also an important factor that affects the teaching quality and teaching effect of schools, and has been paid attention to by many researchers in education. The interactive audio-visual-oral TM in college English is a student-centred TM. In this TM, students are regarded as the subject objects of cognition in the teaching process, and the process of students' receiving education is the process of students' active construction of personal knowledge understanding under the interaction of existing learning experience and the perception of the surrounding environment. In the process of receiving education, students' interaction and cooperation should be two important factors that teachers should consider. Students discuss the theoretical knowledge imparted by the teacher efficiently with other students in class. At the same time of finding that the knowledge points they have mastered are insufficient, they can also carry out preliminary practice on this newly mastered skill, which further improves the teaching effect of English teaching. Through VR and AI technology, this text proposes an interactive audio-visual and oral teaching classroom for college English. This new TM focuses on students, with teachers and related technologies as the auxiliary, to improve students' learning efficiency and effect. Students with good foundations can choose their learning content and rhythm more independently to broaden their

knowledge; students with poor foundations can establish independent learning habits and improve efficiency. Although the starting speed of students at all levels is different, they are all improving their independent learning ability and forming a more positive learning attitude.

This project is based on VR and AI technology and aims to improve students' English audio-visual-oral skills. It adopts a split-class teaching method, divided into traditional college English TM and a newly proposed model for teaching English at the university. The teaching cycle is two months, and tests are conducted once a month. The test indicators include: listening and speaking skills, reading and writing ability, theoretical knowledge, and teaching efficiency. All test results are evaluated on a 10-point scale to track students' progress and optimise teaching effectiveness.

First, the indicators of the two classes are pre-tested, and the results are shown in Figure 5.

Figure 5 Test results of two classes before teaching, (a) traditional college English TM (b) a newly proposed model for teaching English at the university



(a) (b)

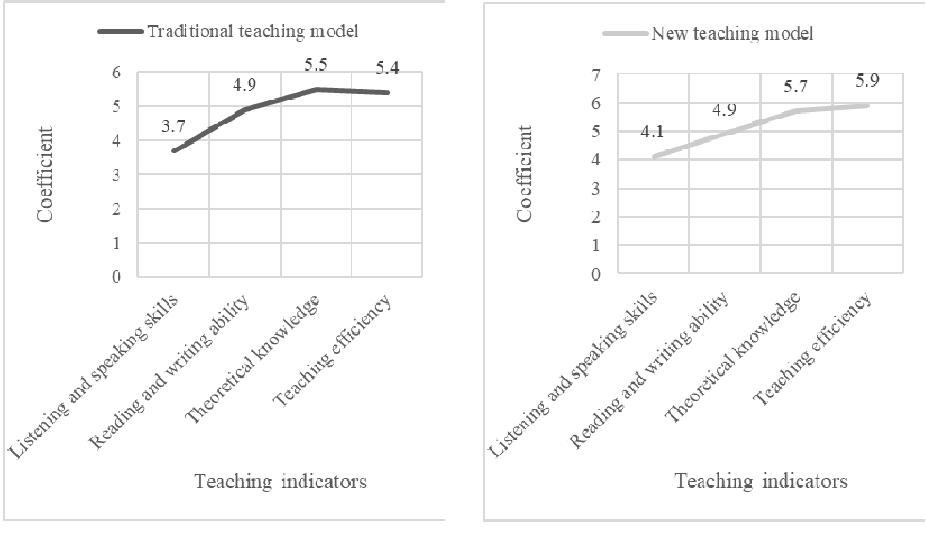
As can be seen from Figure 5, before the group teaching, the difference between the test indicators of the two classes is not large, which shows that the class selection is reasonable, and the differences between the groups are small, so they can be compared.

After one month of group teaching, the teaching effect of the two classes is tested, and the results are shown in Figure 6.

As can be seen from Figure 6, after one month of group teaching, the performance of both classes has improved, but it can be clearly seen that the indicators of the classes using the new method are higher than those of the traditional teaching method.

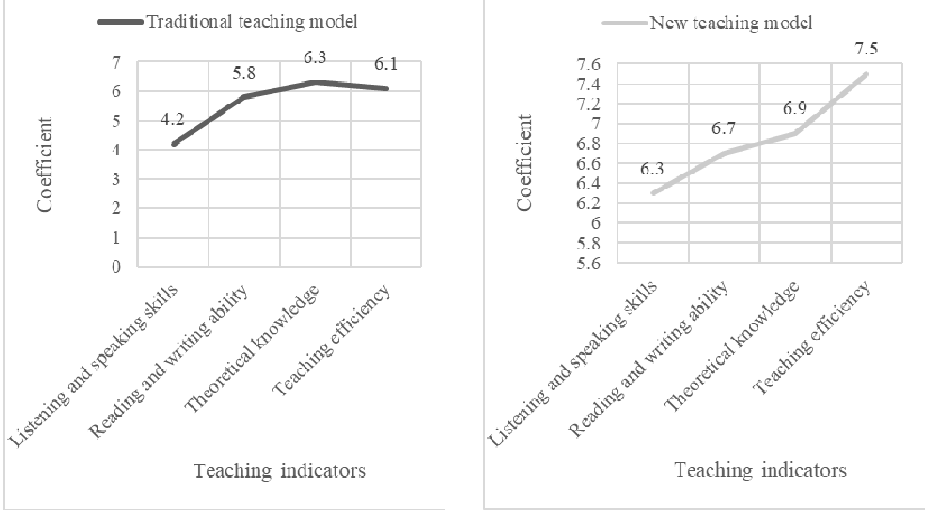
Finally, after two months of teaching, the teaching situation of the two classes is tested, and the results are obtained as shown in Figure 7.

Figure 6 Results of one month of teaching in two classes, (a) traditional college English TM (b) a newly proposed model for teaching English at the university



(a) (b)

Figure 7 Comparison between the new ITM of college English based on VR and AI technology and the traditional English TM in terms of several performances, (a) traditional college English TM (b) a newly proposed model for teaching English at the university



(a) (b)

Finally, the performance differences between the new college English interactive audio-visual-oral TM based on VR and AI technology proposed in Figures 7(a) and 7(b) and the traditional college English TM in multiple teaching indicators are analysed. Through the analysis of the data in Figure 7, it can be determined that the performance of the new college English interactive audio-visual-oral TM based on VR and AI technology proposed in this text has increased by about 35.5% on average compared with the traditional college English TM in terms of various teaching indicators. The new college English ITM has greatly improved the training of English listening and speaking abilities. The main reason is that the new college English ITM is a student-centred TM. It increases students' oral interactive practice when receiving English education, which also promotes students' mastery of theoretical knowledge in English teaching.

8 Conclusions

English can carry a lot of information, which makes it more and more used in various cross-border activities and cross-border daily exchanges. Therefore, it can be predicted that English can be the preferred language for cross-border communication in the future. Therefore, the education field in various regions has paid more attention to the English education model, and English is also an important carrier to enhance the economic strength and image of different regions. However, the current English TM in various regions still attach too much importance to the teaching of theoretical knowledge and grammar, and pay less attention to the cultivation of listening, speaking, reading, and other abilities in English teaching. Therefore, the students cultivated by the current English TM usually cannot master and use oral English skilfully. At this time, the importance of interactive English TM is reflected. The ITM not only makes better use of the essence and characteristics of language teaching, but also combines this feature with the advantages of the IT theory to play a more powerful teaching role. In the interactive English teaching classroom, teachers are more likely to provide students with places to answer various questions and interact with students in oral English. The ITM of college English audio-visual-oral class based on VR and AI technology proposed in this text not only focuses on listening, speaking, and reading abilities in English teaching, but also promotes the same status as theoretical knowledge education. It also optimises the current performance of English teaching, so as to provide students with better education. This better English TM can also provide a better supporting role for the future development of students. Although combining VR and AI technology has advantages, it also faces challenges. The immersion of VR may allow users to confuse virtual and reality, affecting physical and mental health. At the same time, AI technology is still immature, and general AI development is difficult. Although the article discusses the feasibility and advantages of these technologies in English teaching, the solutions to the challenges are insufficient and need further research.

Declarations

This work was supported by Brand Course: College English Band A (ZK2310).

All authors declare that they have no conflicts of interest.

References

Ahmadi, D. and Reza, M. (2018) 'The use of technology in English language learning: a literature review', *International Journal of Research in English Education*, Vol. 3, No. 2, pp.115–125.

Davenport, T.H. and Ronanki, R. (2018) 'Artificial intelligence for the real world', *Harvard Business Review*, Vol. 96, No. 1, pp.108–116.

El Beheiry, M., Doutreligne, S., Caporal, C., Ostertag, C., Dahan, M. and Masson, J.B. (2019) 'Virtual reality: beyond visualization', *Journal of Molecular Biology*, Vol. 431, No. 7, pp.1315–1321.

Fahimirad, M. and Kotamjani, S.S. (2018) 'A review on application of artificial intelligence in teaching and learning in educational contexts', *International Journal of Learning and Development*, Vol. 8, No. 4, pp.106–118.

Ghaemi, F. and Golshan, N.S. (2017) 'The impact of telegram as a social network on teaching English vocabulary among Iranian intermediate EFL learners', *International Journal of Information and Communication Sciences*, Vol. 2, No. 5, pp.86–92.

Ibrahim, A. (2017) 'Advantages of using language games in teaching English as a foreign language in Sudan basic schools', *American Academic Scientific Research Journal for Engineering, Technology, and Sciences*, Vol. 37, No. 1, pp.140–150.

Kavanagh, S., Luxton-Reilly, A., Wuensche, B. and Plimmer, B. (2017) 'A systematic review of virtual reality in education', *Themes in Science and Technology Education*, Vol. 10, No. 2, pp.85–119.

Khotimah, K., Widiati, U., Mustofa, M. and Ubaidillah, M.F. (2019) 'Autonomous English learning: teachers' and students' perceptions', *Indonesian Journal of Applied Linguistics*, Vol. 9, No. 2, pp.371–381.

Kim, N.Y., Cha, Y. and Kim, H.S. (2019) 'Future English learning: chatbots and artificial intelligence', *Multimedia-Assisted Language Learning*, Vol. 22, No. 3, pp.32–53.

Macaro, E., Curle, S., Pun, J., An, J. and Dearden, J. (2018) 'A systematic review of English medium instruction in higher education', *Language Teaching*, Vol. 51, No. 1, pp.36–76.

Shin, M.H. (2021) 'Development of English teaching model applying artificial intelligence through maker education', *Journal of the Korea Convergence Society*, Vol. 12, No. 3, pp.61–67.

Tan, M. (2020) 'Research on English teaching system based on artificial intelligence and WBIETS wireless network system', *EURASIP Journal on Wireless Communications and Networking*, Vol. 2020, No. 1, pp.1–15.

Turan, Z. and Akdag-Cimen, B. (2020) 'Flipped classroom in English language teaching: a systematic review', *Computer Assisted Language Learning*, Vol. 33, Nos. 5–6, pp.590–606.

Villafuerte, J. and Mosquera, Y.M. (2020) 'Teaching English language in Ecuador: a review from the inclusive educational approach', *Journal of Arts and Humanities*, Vol. 9, No. 2, pp.75–90.

Zhao, Y. (2021) 'Research on the blended teaching model of university English based on artificial intelligence', *Tobacco Regulatory Science*, Vol. 7, No. 5, pp.1107–1112.