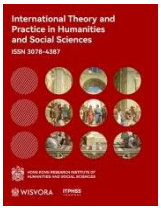
WISVORA

International Theory and Practice in Humanities and Social Sciences
2025 Volume2,Issue6 ISSN 3078-4387



An Empirical Study of AIGC Enabled High School English Listening Class in China

Yanan Zhai[1], Li Jiang[2], Meng Wei[3], Ruoyu Wang[4], Lan Mi[5]

[1] No.94 Middle School, ZhengDong NewDistrict, Zhengzhou, Henan province, China
[2]No.36 Middle School, Zhongyuan District, Zhengzhou, Henan province, China
[3]Songshan Road School, Erqi District, Zhengzhou, Henan province,China
[4]Department of Education, Zhejiang Normal University, jinhua, China
[5]Graduate School,University of perpetual help system DALTA, Philippines

Accepted

2025-05-22

Keywords

AIGC technology; high school English; listening teaching; SPSS analysis

Corresponding Author

Li Jiang

http://doi.org/10.70693/itphss.v2i6.368

Abstract

With the rapid development of AIGC (Artificial Intelligence Generated Content), it is increasingly widely used in education. This study investigates the effectiveness of AIGC in high school English listening classes. A 16-week experiment compared an AI-assisted group with a control group using conventional instruction and then the research group compared the changes in English listening levels in the experimental group (AIGC-assisted teaching) and the control group (traditional listening teaching method). Data were analyzed using SPSS 22.0, with independent t-tests confirming the positive impact of AIGC ($p<.05$)on students' English listening ability. The analysis revealed that the experimental group achieved markedly higher post-test scores ($p<$ 0.05). Additionally, the experiment indicated that AIGC tools boosted student engagement by providing interactive features like real-time pronunciation feedback and adaptive exercises, while also improving teaching efficiency through automated assessments and personalized learning analytics. These results highlight AIGC's potential as an effective educational tool for optimizing language learning outcomes.

1. Introduction

In the context of globalization, the importance of English as a key tool of international communication is obvious. As one of the core skills of English learning, the teaching effect of listening directly affects students' comprehensive language application ability. However, traditional listening teaching methods often face challenges such as single content and insufficient interaction, which makes it difficult to enhance students' motivation and engagement for learning. In addition, due to the limitation of teacher resources and teaching materials, the traditional English listening teaching mode makes it difficult to meet the needs of students' personalized learning. With the rapid development of science and technology,

artificial intelligence-generated content (AIGC) technology is more and more widely used in the field of education. Especially in the field of language learning, AIGC technology has greatly enriched teaching resources and improved learning efficiency by providing personalized and interactive learning content. Therefore, this study aims to integrate AIGC into high school English listening classes, hoping to improve students' English listening skills by providing a rich variety of listening materials and intelligent teaching AIDS. Meanwhile, this study will use empirical research methods to evaluate the application effect of AIGC in high school English listening teaching, aiming to provide a useful reference for the field of English teaching.

2. Literature Review

AIGC, as a cutting-edge technology using machine learning and deep learning algorithms, is committed to creating multi-form content through automated processes. The technology covers a wide range of advanced areas such as natural language processing (NLP), computer vision, speech synthesis, and generative adversarial networks (GANs). The core of AIGC is its ability to simulate human creative thinking, thus being able to produce both high-quality and diverse content (Silver et al., 2017). Reviewing its development history, AIGC originated from the early exploration of artificial intelligence. In the middle of the 20th century, as the field of computer science began, researchers began to use algorithms to generate basic pieces of text or music. However, limited to the computing power and data scale of the time, these attempts were not widely available(Xu, R., & Wang, Z, 2024).

Entering the 21st century, the leap of big data and cloud computing technology has injected new vitality into AIGC. In particular, the rise of deep learning technology has greatly improved AIGC's ability to handle complex content generation tasks. Take the Google DeepMind team for example, whose impressive performance in 2016 marks a significant advancement in AI in the field of complex decision-making (Mahajan et al., 2024). In recent years, AIGC has made remarkable achievements in the field of content generation, especially the GPT series model launched by OpenAI, whose outstanding performance in text generation has set a new benchmark for the industry (Wang, W. ,2024). The application of AIGC in education has covered many aspects, including the generation of personalized learning resources, the development of automatic evaluation systems, and the design of interactive learning tools. AIGC can generate personalized learning materials based on students' learning progress and abilities. A group of academic researchers have been developing this intelligent English tutoring system to support English teaching. For example, Eger has evaluated some popular pronunciation intelligent English tutoring systems in ESL classes in Canada that only focus on segmentation rather than pronunciation (Eger, N. A., & Reinisch, E., 2019). They believe that AI technology can assist English teachers in terms of pronunciation and different accents. In addition, some experts introduced three existing intelligent tutoring systems that may assess the vocabulary ability of English learners, namely the Vocabulary Levels Test, and the developed test Vocabulary Size Test and network-based diagnostic system (El-Sourani, A. I. , Keshta, A. S. , & Aqel, M. S., 2021) . However, these three intelligent tutoring systems are limited to written media and only focus on a single word, not considering the context of the words. Fu Ying, Wang Min, Min Shangchao, Zhou Songbo and Pan Xunyi experimented with enhancing English language proficiency among 746 users (Fu, Y., Wang, M., Min, S., Zhou, S., & Pan, X, 2023). Fancy, an AI app designed for Chinese English learners, was

similar to Duolingo. The results showed that 72% of users spent about two months and improved their test scores. In addition, Jiang R studied the efficacy of providing error correction feedback systems based on AI and compared the effect difference between the two types of error correction feedback (Jiang R, 2022) . The results showed that the majority of students favored the efficacy and quality of the system, reflecting that AI technology can support reducing the workload of teachers and administrators.

In English teaching, teachers are still dominant. With the development of artificial intelligence technology and the expansion of the English teaching database, artificial intelligence can help English teachers teach and learners learn to a large extent because its efficiency and role can replace the status and functions of teachers in some specific areas of English teaching. For example, English teachers can be relieved from the basic and repetitive work of teaching grammar rules and pronunciation year after year, interpreting reading and listening articles, providing writing and speaking structures and demonstrating answers, and analyzing students' performance. At the same time, artificial intelligence can also assist teachers in teaching management, such as preparing lessons, preparing teaching materials, designing classroom activities, scoring, assigning and reminding homework, sending notices, recording attendance, and so on. With the help of AI tools, English learners can also teach themselves pronunciation, vocabulary, grammar, reading, listening, writing, and spoken English anytime and anywhere. At the present stage, for example, the AI oral English evaluation system developed in China has realized the scoring evaluation of the oral English test of entrance examination for secondary school or college, and the scoring accuracy has also been recognized. The middle school and college entrance examinations in some cities use artificial intelligence to score the oral English test, and the effect is also very satisfactory (Wang, W., 2024). College English teaching has also been promoted in the context of artificial intelligence technology technique of reading has received positive reviews to help students achieve accurate assessment and feedback on multiple phonemes(Yin S. K, 2022). Through practical cases, ChatGPT demonstrated how to improve the teaching level of English teachers and the ability to publish academic papers, improve students' English learning efficiency and overall quality, and enhance the quality and diversity of teaching materials.

To sum up, the application of AI technology, both to learners and teachers, can enable them to pursue better results, and generate the motivation to use AI teaching and assisted learning.

3.Methodology and procedures

3.1 Homogeneity analysis between experimental class and control class

Homogeneity analysis is an analysis of the differences in the pre-experiment levels of the two classes. It is used to determine whether the students in the experimental class will show significant changes due to the teaching strategy for cultivating English listening abilities based on the Artificial intelligent system. Before the experiment, students' listening scores were tested on the students of the two classes from No.36 Middle School in Zhongyuan District, Zhengzhou City, totaling 100 students as the research subjects, among which the experimental group and the control group were composed of 50 students each with matched gender ratios (EC: 26 female, 24 male; CC: 25 female, 25 male) and age distributions (EC: mean age = 16.2 ± 0.5 years; CC: mean age = 16.1 ± 0.6 years). By analyzing the English test scores, it is determined whether there are differences in the English listening abilities of the students in the two classes, laying a good foundation for subsequent experiments. Students in the control group adopted the traditional listening teaching method, while students in the experimental group received

additional listening training supported by AIGC in addition to conventional listening teaching.

Using SPSS for data statistical analysis, an independent - sample T - test is carried out on the tests of the experimental class and the control class. According to the results of the data analysis, it is determined whether there are significant differences in the overall means of the two classes. The analysis of significant differences is shown in Table 1 and Table 2.

Table 1 Independent sample T - test (1) of Pre - test Scores in EC and the CC

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
Scores	EC	50	65.330	12.3473	1.7462
	CC	50	64.530	11.7497	1.6617

EC （Experimental Class)

CC(Control Class)

As can be seen from Table1, the sample size of the experimental class is 50. The mean of the English test scores is 65.33, the standard deviation is 12.35, and the standard error of the mean is 1.746. The sample size of the control class is 50 too. The mean of the English test scores is 64.53, the standard deviation is 11.75, and the standard error of the mean is 1.66. Thus, the mean of the test scores of students in the experimental class is 0.8 higher than that of the control class.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Scores	Equal variances assumed	.530	.468	.332	98	.741	.8000	2.4104	-3.9834	5.5834
	Equal variances not assumed			.332	97.760	.741	.8000	2.4104	-3.9836	5.5836

Table 2 Independent sample T - test (2) of Pre - test Scores in EC and the CC

As can be seen from Table 2, the F - value of the students' pre - test scores is 0.53, and the significance level 0.468> 0.05, which satisfies the assumption of equal variances. In the "Equal variances assumed" column, the significance (two - tailed) is 0.741 > 0.05. This indicates that there is no significant difference in the mean values of the pre - test scores between the experimental class and the control class. Therefore, it can be concluded that the levels of the two classes before the experiment are comparable, and the experiment can proceed.

3.2 Research tools

With the help of SPSS 22.0, this study uses rigorous descriptive analysis, t-test analysis of variance, and other means to conduct an in-depth analysis of the collected data and strive to make the scientific and research results accurate.

3.3 Experimental design

In the process of selecting these two classes, the researchers also paid special attention to the comparability of the two classes in terms of student size, gender ratio, age distribution, etc. This means that classes C1 and C2 should have similar characteristics on these key background variables to ensure the validity and reliability of the experimental results. With such careful selection and control, the investigators hope to minimize the interference with experimental results and more accurately evaluate the effects of experimental interventions.The intervention lasted 16 weeks, with the experimental group receiving three 45-minute AIGC-assisted sessions weekly (total: 36 hours of exposure).The 16-week intervention period was selected based on prior studies showing significant language skill improvements within this timeframe (Li et al., 2019). Control group instruction adhered to the provincial standard curriculum, which included: (1) Textbook audio materials (e.g., New Concept English), (2) Teacher-led dictation exercises, and (3) Monthly mock listening tests. This ensured comparability with the experimental group's baseline exposure.

Preparation of teaching materials: Experimental group: In the intelligent classroom of the school, we introduced the English listening assistant agent in the iFlytek Spark platform, as an important auxiliary tool in the teaching process. The iFlytek Spark Cognitive AI Platform is an advanced Chinese artificial intelligence system developed by iFlytek, integrating natural language processing (NLP), speech recognition, and generative AI to enhance language education. It supports multimodal interaction (text, voice, and image inputs) and generates personalized learning materials, such as adaptive listening exercises and simulated dialogues, tailored to students' proficiency levels. The platform provides real-time feedback on pronunciation, fluency, and intonation, while its data-driven analytics track student progress and assist teachers in optimizing instruction. Built on deep learning models and trained on extensive multilingual datasets, iFlytek Spark enables immersive, interactive English learning experiences, making it a powerful tool for modern AI-assisted language education. Thus, it is a good AI platform to assist students' listening learning. First, the teacher gave special training to the students, explaining in detail how to use the correct instructions to have an effective dialogue with the agent, to ensure that the students can make full use of this technology. Then, we use advanced artificial intelligence-generated content (AIGC) technology to elaborate a rich variety of voice materials. These materials not only cover listening exercises with different levels of difficulty, different speeds, and various accents but also include dialogue models that simulate real dialogue scenarios. These simulation questions can adjust the difficulty and theme in real-time according to the accuracy of students, to meet the personalized learning needs of each student. In addition, the agent can also intelligently score students' performance, and provide personalized learning advice to help students better master English listening skills, to improve their language application ability.

Control group: In traditional English listening teaching, teachers usually prepare some traditional teaching materials, such as audio tapes and listening exercises. These materials usually include exercises with fixed difficulty and topics, lacking real-time adjustment and personalized feedback. This traditional teaching method often fails to meet the personalized needs of each student, making it difficult for students to obtain targeted guidance and help in

the process of listening learning.

Technology and equipment preparation: In the experimental group class, students do targeted exercises three times a week using the smart classroom equipped with the school. To ensure that these exercises go smoothly, the classroom is equipped with a stable Internet connection, allowing students to make full use of AIGC for teaching and practice activities. To ensure that students are skilled in using these technologies, the school has also trained students in advance, especially for the personalized listening learning function for the iFlytek Spark platform. Students have learned how to operate and use these tools in the training to be more comfortable in the actual learning process.In the experimental group, AIGC was integrated as supplementary material to enrich traditional listening instruction. The iFlytek Spark platform generated personalized exercises by analyzing individual student performance (e.g., pre-test scores, error patterns in real-time responses) through its NLP-driven adaptive engine. For example, it dynamically adjusted audio speed (0.8x–1.2x), accent variations (British/American/Australian), and contextual difficulty (e.g., simplified dialogues for beginners vs. news clips for advanced learners). Specific AI-generated content included: (1) targeted minimal-pair drills (e.g., 'ship/sheep') to address pronunciation errors, and (2) simulated real-world scenarios (e.g., airport announcements with background noise). Teachers used AI-curated playlists aligned with weekly themes (e.g., 'Travel') while retaining textbook content as the core framework.

Teacher Training Protocol: To minimize instructional bias, experimental group teachers completed a 6-hour standardized training program.

Phase 1 (4 hours): Technical workshops on iFlytek Spark's interface, including how to generate/assign AIGC materials (e.g., customizing accent/difficulty settings) and interpret AI feedback reports.

Phase 2 (2 hours): Pedagogical guidance on blending AIGC with traditional methods (e.g., using AI-generated dialogues for warm-ups while reserving textbook audio for assessments).
To ensure consistency, all teachers followed a predefined lesson plan template, and classroom sessions were audited monthly for compliance.

The entire experimental cycle is set for 16 weeks, during which both the experimental and the control groups receive normal English listening instruction. However, students in the experimental group have additional opportunities to use AIGC on this basis, thus gaining more support and resources in listening learning. In this way, schools hope to observe and evaluate the actual effect of AIGC in English listening teaching, and its specific impact on students' learning effectiveness.

3.4 Ethical Compliance

This study followed the ethical review process outlined by ZhengZhou NO.36 High School's Academic Committee, with oversight from the lead researcher's institutional ethics advisor.Prior to participation(1)All students and their legal guardians provided written informed consent with clear explanations of the study's purpose and data usage;(2)Participation was voluntary with right to withdraw at any time.(3)All data were anonymized by using alphanumeric codes (S01-S100) instead of personal identifiers;(4)Teachers in both groups received equal training hours to prevent resource

inequality;(5)The control group was offered access to AIGC materials after the study concluded.

4. Data collection and analysis

We collected the students' listening scores by implementing the pre-post test (i.e., the unified English listening test conducted before and after the end of the experiment). To ensure the accuracy and reliability of the data, we adopted the professional statistical software SPSS 22.0 to make a detailed comparative analysis of the listening performance difference between the experimental class and the control class before and after the experiment. Through this scientific analysis method, we can more accurately evaluate the effect of the experiment, and thus provide strong data support for teaching improvement.

Although ANCOVA could control for baseline differences, independent t-tests were selected because:Pre-test scores showed no significant difference ($p = 0.741$, Table 2);The study design met t-test's homogeneity assumptions (Levene's $p=0.468$).Effect sizes were calculated using Cohen's d (d=1.8 for experimental vs. control group), indicating large practical significance.

5. Research results

5.1 Analysis of the experimental class

The Paired Samples T-test of English Listening Proficiency for Students in the Experimental Class is shown in Table 3 and Figure 1.

Table 3 Comparison of test differences before and after Class C 1

Paired Samples Test

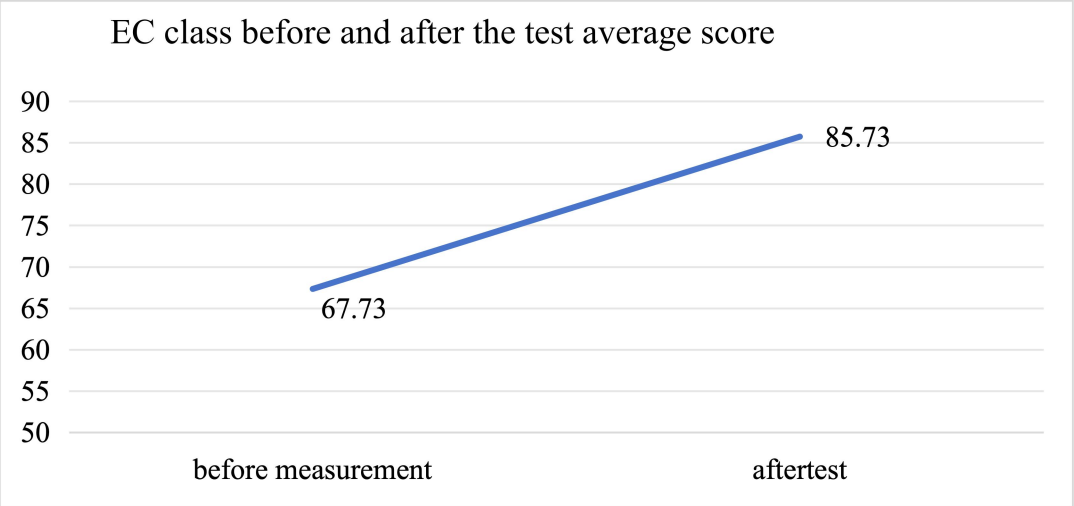
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Class C1	pre -post	-20.400	12.3912	1.7524	-23.9215	-16.8785	-11.641	49	.000

Class type	Pre-test (n=50)	Post-test (n=50)	t	T	p
C1	65.33±12.347	85.73±8.931	-11.641		**

* * Repr

esents a p-value of less than .01

Figure 1 Comparison of the average scores before and after Class C1

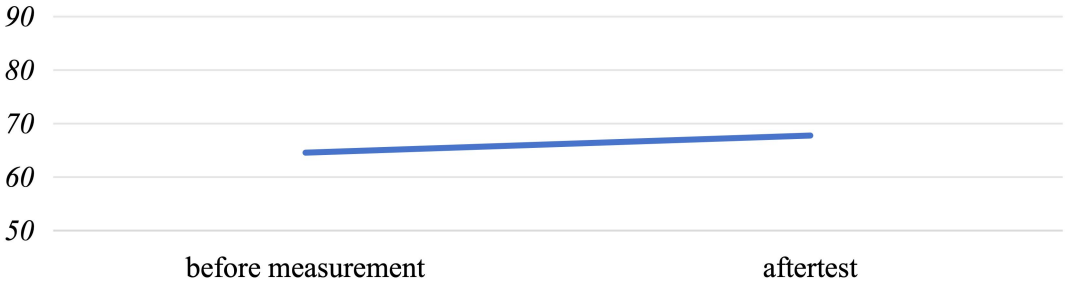


From the data in table-3 , it can be concluded that in Pair 1, according to the data analysis of the pre-test and post-test of students' language perception ability, t = -11.64 and the significance value F = .000 < .05. The negative T-value indicates that the post-test mean (85.73) was significantly higher than the pre-test mean (65.33), reflecting the directionality of improvement.As shown in Figure 1, the average listening score of the experimental class was greatly improved with the assistance of AIGC.This indicates that there is a significant difference in the experiment, suggesting that there are significant changes in the listening ability of the students in the experimental class before and after the experiment.

5.2 Analysis of the Control Class

The Comparison of the average scores in control class is shown in Figure 2.We can find that compared with experimental class, there is a small increase in students' listening skills by using the traditional teaching methods. Paired Samples T-test of English Listening Proficiency for Students in the Control class is shown in table 7.

Figure 2 Comparison of the average scores before and after CC



The paired sample T-test of the data before and after the English test results are shown in Table 4.

Class type	Pre-test (n=50)	Post-test (n=50)	T
C2	64.53±11.75	67.73±9.21	9.92

Table 4 Comparison of test differences before and after class C 2

As can be seen from the above table, there was no significant difference in the C2 class test

(t=9.92 p> 0.05), and the performance of the students in the control class improved slightly, but the improvement effect was not obvious.

5.3 Analysis of the Post-test between EC and CC

By comparing the average values (post-test) of the scores of students in the experimental class and the control class and conducting an independent samples T-test on the average values, we analyzed and compared whether there were changes in the English listening abilities of students in the experimental class and the control class after the experiment. The specific analysis is shown in Tables 5 and 6.

Table 5 Independent sample T - test (1) of Post - test Scores in EC and the CC

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
Post-test	1	50	85.730	8.9318	1.2632
	2	50	67.730	9.2119	1.3028

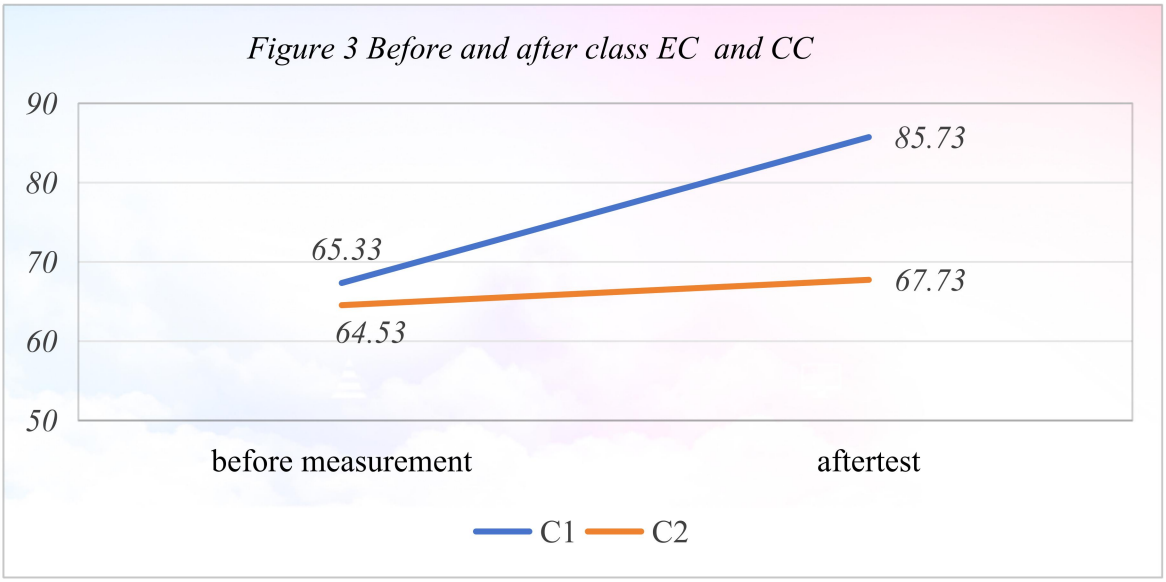
Table 6 Independent sample T - test (2) of Post - test Scores in EC and the CC

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
Post-test	Equal variance assumed	.12	.72	9.92	98	.000	18.0000	1.8146	14.39	21.60
	Equal variance not assumed			9.92	97.90	.000	18.0000	1.8146	14.39	21.60

Through the analysis of Tables 5 and 6, it is concluded that in terms of listening ability, the average value of students in the experimental class is 18 points higher than that of students in the control class. The significance level is 0.726 > 0.05, and in the homogeneity of variance test, the significance value F = .000 < 0.05. This indicates that there is a significant difference in the listening ability between students in the experimental class and those in the control class after the experiment.

The change curves of the pre-test and post-test listening abilities of students in the experimental class and the control class are displayed in figure 3.



The average score of the pre-test of the experimental class was 65.33, and the average score of the control class was 64.53 before the test. After the AIGC-assisted instruction of the same time, the average score of the experimental class was 85.73, the average score of the control class was 67.73, and the average score of the experimental class was greatly improved.

Table 7 Comparison of tests differences between CC and EC

* * Represents a p-value of less than 0.01

Test Group	C C (n=50)	EC (n=50)	T
Pre-test results	65.53±12.34	64.53±11.75	9.92
Post-test results	85.73±8.93	67.73±9.21	-11.641**

Cohen'd=1.8(95% C1:1.2-2.4)

The results of the control class and experimental class before and after test, and the results are shown in Table 7. There was no significant difference between the two classes in the pre-test (t=9.92 p> 0.05), and the significant difference between the pre-test (t=-11.641 p<0.01). The large effect size (Cohen's d = 1.8, 95% C1:1.2-2.4)indicates that AIGC-assisted teaching had a practically significant impact, surpassing the typical threshold for educational interventions (d > 0.8; Fu et al., 2023).It can be seen that the classroom teaching strategy of AIGC adopted by the experimental class effectively improved the efficiency of the English classroom, thus greatly improving the student's performance at the same time.

Through the comparative analysis of the before and after test scores and the results of the experimental class and the control class, we can see that the classroom teaching strategy of AIGC assisted English listening adopted by the experimental class has shown a significant effect in improving the efficiency of the English classroom. This strategy allows students to achieve greater performance improvement in the same amount of time. Specifically, the performance in the post-test of the experimental class is significantly improved compared with the pre-test, and the improvement range is large, while the difference between the

average score before and after the test of the control class is not significant. Although the average score of the experimental class and the control pretest class were almost the same before the start of the experiment, the improvement effect of the post-test score of the experimental class was particularly prominent after the same teaching time. Further through the difference test of the pre and post-test data, we found that there were significant differences in the data of each dimension, and the average value of the post-test data was higher than that of the pretest data. This result fully shows that the effect of the AIGC assisted high school English listening classroom teaching strategy is significantly better than the traditional classroom teaching methods, thus showing a more significant advantage in improving students' English performance.

6. Discussion

In this study, we carefully analyzed and evaluated the changes in English listening levels in the experimental and control groups. The experimental group introduced the AIGC in the high school English listening class, while the control group used the traditional teaching methods. Through a series of tests and assessments, we found that the students in the experimental group have made significant progress in English listening. Specifically, the experimental group of students' listening comprehension ability, vocabulary recognition speed, and grasp of the complex sentence structure ability have been significantly improved.While Erickson & Siau (2018) reported a 72% improvement in vocabulary scores using AI tools, our study focused on contextual listening skills (e.g., accent adaptation) through AIGC-generated scenarios. This distinction highlights AIGC's unique strength in simulating real-world communication, aligning with Wang's (2024) call for pragmatic competence development in EFL settings.However, unlike their study focusing on vocabulary, our AI-generated contextual dialogues (e.g., airport announcements) align with Heffernan (2012)'s emphasis on real-world scenario training. This suggests AIGC's unique strength in enhancing pragmatic listening skills.

The results further show that the application of AIGC in high school English listening classes not only improves students' English listening level but also stimulates their interest and enthusiasm for English learning. Students generally reflect that AIGC makes listening exercises more vivid and interesting, and can better attract their attention, thus improving learning efficiency. In addition, AIGC also provides teachers with more teaching resources and means, making classroom teaching more diversified and personalized.

Given the remarkable effect of AIGC in improving students' English listening levels, as well as its wide recognition and welcome among students, we strongly recommend actively promoting and applying AIGC in high school English listening teaching. In this way, it can not only effectively improve students' English listening ability, but also further cultivate their comprehensive quality, laying a solid foundation for their future study and life.

7. Conclusion and Suggestions

AIGC, also known as artificial intelligence generation content technology, effectively solves a series of problems faced by traditional listening teaching methods by creating a rich variety of listening materials and providing intelligent teaching auxiliary tools. This advanced technology can not only greatly stimulate students' interest and enthusiasm in learning, but also provide personalized learning suggestions and guidance according to each student's

learning situation and needs. In this way, students can master the required skills more efficiently in the process of learning English listening, to be more handy in practical application.

Specifically, AIGC can generate various types of listening materials, such as dialogue, short essays, news reports, etc., which not only cover different topics and scenes but also can adjust the difficulty according to the actual level of students. In this way, students can practice listening in a variety of contexts to better adapt to various practical communication situations. In addition, AIGC can analyze students learning data in real-time, identify their weak links, and provide targeted exercises and suggestions to help them improve their listening level specifically.

The application of AIGC is not limited to the generation of listening materials but also includes advanced technologies such as intelligent speech recognition and natural language processing. These techniques can help students get immediate feedback during their listening exercises to more accurately understand what they hear. At the same time, AIGC can also simulate the real language environment, so that students can practice listening in the virtual scene, and further improve their practical application ability.

In short, the application of AIGC in English listening teaching not only greatly enriches the teaching content and means, but also provides students with a more personalized and efficient learning experience, to help them make significant progress in English listening. While AIGC significantly improved scores ($p < .01$, $d = 1.8$), its scalability depends on contextual factors. For instance, Jiang (2022) notes that AI adoption in schools requires sustained teacher training—a challenge observed in our study, where 25% of educators reported difficulty adapting to real-time AI feedback. Thus, AIGC is better framed as a transformative supplement (Abdellatif et al., 2022) rather than a standalone solution.

7.1 The possible trend of AI-enabling English education in high schools

One of the most significant trends is the intelligent and personalized teaching. AI teaching tools use big data analysis and machine learning techniques to tailor learning plans for different students, and adjust them according to their learning levels, interests and weaknesses. This approach not only improves the teaching efficiency, but also meets the personalized needs of students, thus enhancing the overall learning effect. Interactivity and immediate feedback is another important trend. With the help of natural language processing and speech recognition technologies, AI teaching tools provide high-quality English listening and speaking training and give immediate feedback in the process. In the English teaching of senior high school, students wear smart headphones for dialogue exercises and AI system, grammar and fluency were assessed in real time, with personalized advice immediately upon completion. This instant interaction stimulates students interest in learning and significantly improves their listening and speaking skills. In the future, the integration of virtual reality (VR) and augmented reality (AR) technology will become an important direction. Through VR and AR, students can enter the virtual English language environment, conduct immersive training, improve their learning experience, interest and participation, and develop their autonomous learning ability. Data-driven educational decisions are also a striking feature of AI-enabled education. The AI system monitors and analyzes students learning behavior in real time, generates learning reports, and provides teaching suggestions for teachers. Combined with these data, teachers can conduct targeted teaching intervention to enhance the teaching effect.

The above trends show that AI-enabled English education in senior high schools can significantly improve the listening and speaking ability. Through personalized learning path, immersive learning experience and data-driven decision-making, it can comprehensively improve the teaching quality and meet the new needs of English teaching in the future.

7.2 Suggestions on the design of AI-based English teaching

AI technology should strengthen personalization and real-time feedback in English education AIDS. Develop an adaptive learning system to analyze data through AI algorithm, accurately evaluate students' listening and speaking ability, and provide personalized learning path. Using natural language processing, design virtual teachers, simulate real scenes, improve oral expression ability. Speech recognition and evaluation are central. High-precision technology helps to correct pronunciation immediately, and the evaluation system analyzes pronunciation accuracy, fluency and intonation, and gives detailed feedback, so that students can practice targeted exercises. AI tools should provide diverse learning resources, such as courseware, video, listening materials and dialogue simulation, classified by difficulty to meet different levels of needs, and be updated regularly to follow up on educational trends. Data privacy protection is crucial. Set laws and regulations to ensure data security, establish encryption and access control mechanisms, prevent leakage and abuse, and create a safe and reliable learning environment.

7.3 Challenges of AIGC in practical application

Although artificial intelligence-generated content (AIGC) technology performs well in high school English listening classes, and can significantly improve students' interest in learning and listening comprehension ability, it still faces some challenges in practical application. For example, educators need to carefully consider how to ensure that the content generated by AIGC fits with the teaching objectives and textbook content to avoid deviations from the syllabus. In addition, how to balance the relationship between AIGC and traditional teaching methods is also an important issue, because over-reliance on technology may weaken the leading role of teachers in the classroom, while completely abandoning traditional teaching methods may lead to students' lack of the necessary basic knowledge and skills. Moreover, how to evaluate the long-term effect of AIGC is also an urgent problem to be solved, because the short-term effect can not fully represent its actual effect in long-term teaching. These issues need to be further explored and addressed by educators and researchers to ensure that AIGC can reach its maximum potential in high school English listening classes while avoiding potential negative effects.

Specifically, when applying AIGC, educators must deeply analyze the curriculum standards and syllabus to ensure that the content generated can be closely linked to it so that students do not deviate from the established learning goals in the learning process. At the same time, teachers need to flexibly use AIGC in the classroom and combine it with traditional teaching methods to give full play to their respective advantages. For example, teachers can use AIGC to generate personalized listening exercises, and then explain and discuss them through traditional teaching methods to ensure that students can fully understand what they have learned. In addition, teachers also need to pay attention to students' adaptation to AIGC technology and adjust teaching strategies in time, to avoid students' excessive reliance on technology and ignoring the cultivation of basic learning methods and thinking ability.

In assessing the long-term effects of AIGC, educators need to design scientific evaluation systems that focus not only on students' short-term progress, but also on their long-term learning effects and ability improvement. For example, students' listening comprehension changes over time, and their mastery of English listening knowledge can be tracked through regular tests and assessments. At the same time, teachers can also understand the students' feelings and learning experience of AIGC through observation and interview, to better adjust the teaching methods and content.

In short, the application of AIGC in high school English listening classes has a broad prospect, but it also requires the joint efforts of educators and researchers to solve the challenges faced in its practical application. Through scientific planning and reasonable application, AIGC is expected to play a greater role in English listening teaching in high school, helping students improve their listening comprehension ability, so as to improve the overall English learning effect.

7.4 Limitations of the Research and Future Research Directions

There are three limitations in this study. First, despite matching gender and age distributions, our sample was limited to an urban high school in Zhengzhou. This may not represent rural schools where infrastructure constraints (e.g., unstable internet) could hinder AIGC implementation (El-Sourani et al., 2021). Future studies should include multi-site trials to assess scalability.Second, the 16-week duration may not capture long-term retention effects. Third, teacher training variability (despite protocols) could introduce bias, echoing Ren's (2024) findings on instructor-dependent outcomes in technology-assisted language learning.

To address the current limitations, three critical directions emerge. First, expanding trials to include rural schools (e.g., Western China) would assess AIGC's adaptability to infrastructure disparities. Second, longitudinal studies with 1-year follow-ups are needed to evaluate long-term skill retention beyond the 16-week intervention. Finally, building on Jiang's (2022) framework, developing standardized protocols for human-AI hybrid teaching models could optimize the balance between technological and pedagogical strengths. These steps would strengthen the generalizability and practical implementation of AIGC-assisted language learning.

Acknowledgements: Thanks to the researchers who studied the relevant English listening teaching in China and thanks to the students who participated in this study. Thanks to their work, the authors of this study could have the chance to finish this empirical study.
Conflicts of Interest: The authors declare no conflict of interest.
Funding: This research received no external funding.

References

Abdellatif, H., Al Mushaiqri, M., Albalushi, H., Al-Zaabi, A. A., Roychoudhury, S., & Das, S. (2022). Teaching, Learning and Assessing Anatomy with Artificial Intelligence: The Road to a Better Future. *International journal of environmental research and public health*, 19(21), 14209. https://doi.org/10.3390/ijerph192114209

Alrawashdeh, A. I. , & Al-Zayed, N. N. . (2017). Difficulties that english teachers encounter while teaching listening comprehension and their attitudes towards them. *English Language Teaching*, 10(5).https://doi.org/10.5539/elt.v10n5p167

El-Sourani, A. I., Keshta, A. S., & Aqel, M. S. (2021). The effectiveness of educational environment based on smart learning in developing English language listening skill among IUG female learners. *IUG Journal of Educational & Psychological Studies*, 29(6), 1–15. https://doi.org/10.33976/IUGJEPS.29.6/2021/25

Fu, Y., Wang, M., Min, S., Zhou, S., & Pan, X. (2023). Exploring the relationship between metacognitive awareness and Chinese EFL learners' listening skills. *Frontiers in psychology*, 14, 1148610. https://doi.org/10.3389/fpsyg.2023.1148610

Heffernan, N. T., & Koedinger, K. R. (2012). The future of intelligent tutoring systems: A view from the classroom. *In Proceedings of the International Conference on Artificial Intelligence in Education*. Springer. https://doi.org/10.1007/978-3-031-36033-6_23

Jiang R. (2022). How does artificial intelligence empower EFL teaching and learning nowadays? A review on artificial intelligence in the EFL context. *Frontiers in psychology*, 13, 1049401. https://doi.org/10.3389/fpsyg.2022.1049401

Mahajan, R., Badhera, U., & Kushwaha, P. S. (2024). Integrating generative AI in management education: A mixed-methods study using social construction of technology theory. *International Journal of Management Education*, 22(3), Article 101017. https://doi.org/10.1016/j.ijme.2024.101017

Ren, Z. B., & Schaumberg, R. (2024). Disagreement Gets Mistaken for Bad Listening. *Psychological science*, 35(5), 455–470. https://doi.org/10.1177/09567976241239935

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., Chen, Y., Lillicrap, T., Hui, F., Sifre, L., van den Driessche, G., Graepel, T., & Hassabis, D. (2017). Mastering the game of Go without human knowledge. *Nature*, 550(7676), 354–359. https://doi.org/10.1038/nature24270

Wang, W. (2024). A study of artificial intelligence-assisted listening training in college English teaching. *Applied Mathematics and Nonlinear Sciences*, 9(1), 1–12. https://doi.org/10.2478/amns-2024-2657

Xu, R., & Wang, Z. (2024). Generative artificial intelligence in healthcare from the perspective of digital media: Applications, opportunities and challenges. *Heliyon*, 10(12), e32364. https://doi.org/10.1016/j.heliyon.2024.e32364

Yin S. K. (2022). A study of the effects of thematic language teaching on the promotion of multimedia design students' listening and speaking skills. *Frontiers in psychology*, 13, 915145. https://doi.org/10.3389/fpsyg.2022.915145