Original Paper

Personalized Oral English Learning Based on Generative Artificial Intelligence

Guohui Rao[1]

[1] School of Foreign Languages, Chengdu University of Information Technology, Chengdu, Sichuan 610225, China

Abstract

In the context of the rapid advancement of Generative Artificial Intelligence (GAI) technology and the intelligent transformation of education, traditional oral English learning is plagued by such enduring dilemmas as a one-size-fits-all approach, insufficient practice opportunities, delayed feedback, and anxiety about speaking up, making it difficult to meet learners' personalized needs and the requirements for cultivating international competence. Generative Artificial Intelligence, however, offers a brand-new solution to these challenges with its robust natural language processing, multimodal interaction, and adaptive learning capabilities. Grounded in constructivist learning theory, second language acquisition theory, and metacognitive theory, this paper reviews the current research on the integration of Generative Artificial Intelligence into oral foreign language pedagogy. Focusing on six core dimensions—personalized learning demand analysis, customized learning corpus generation, interactive and authentic learning experience, intensive language application practice, intelligent assessment with real-time feedback, and dynamic adjustment of learning plans, this study systematically examines the application pathways and implementation strategies of Generative Artificial Intelligence in personalized oral English learning by combining specific cases of application tools. As an auxiliary tool, Generative Artificial Intelligence drives the transformation of education and the innovation of teaching models, providing an efficient solution for personalized oral English teaching. This research enriches the theoretical framework of GAI-empowered language education, and offers practical references for promoting the personalized and intelligent reform of oral English teaching.

Keywords

Generative Artificial Intelligence, Oral English Learning, Personalized Learning

1. Introduction

Generative Artificial Intelligence (GAI) refers to models and related technologies with the ability to generate content such as text, images, audio, and video (Cyberspace Administration of China. Interim Measures for the Administration of Generative Artificial Intelligence Services, 2023). It can independently learn existing data and information in a human-like manner, explore the inherent structure and patterns of data, and predict and generate various types of new data similar to the original data, thereby enabling more natural conversations and faster content creation. In contrast, traditional Analytical Artificial Intelligence focuses more on analyzing given data to identify patterns and relationships within it, and generate reports or provide suggestions based on this analysis, without the ability to create new content.

Since OpenAI released GPT-3.5 in November 2022, discussions in the educational circles at home and abroad about the impact of Generative Artificial Intelligence on education have become increasingly intense. On September 7, 2023, the United Nations Educational, Scientific and Cultural Organization (UNESCO) issued the Guidance for Generative Artificial Intelligence in Education and Research, the world's first guiding document on Generative Artificial Intelligence, aiming to facilitate the better integration of Generative Artificial Intelligence into education (UNESCO. Guidance for Generative Artificial Intelligence in Education and Research, 2023). In an era of deep integration of digital informatization and educational globalization, Generative Artificial Intelligence has achieved breakthrough progress in multimodal content generation and intelligent interaction by virtue of core technologies such as deep learning and natural language processing, injecting new impetus into the personalized transformation of the educational ecosystem.

Personalized Learning is a learning method that optimizes learning progress and approaches according to the needs of each individual learner (Han, 2022). Currently, Generative Artificial Intelligence is constantly changing learning methods and presenting a new form that better meets students' personalized needs. From one perspective, personalized learning based on Generative Artificial Intelligence can be understood as generative learning, which emphasizes the formation of diverse learning pathways through highly interactive activities such as questioning and follow-up questioning between the "high autonomy" of individual learners and the "high generativity" of Generative Artificial Intelligence as an integrated content resource and environment, thus catering to personalized needs. It features creativity, diverse pathways for solving complex problems, self-monitoring and self-regulation, among other characteristics (Zhu, Zhang, & Han, 2025).

Along with the rapid development of Generative Artificial Intelligence technology, the educational ecosystem is undergoing a profound intelligent transformation. As a core component of foreign language education, the demand for personalized reform of oral English teaching is becoming increasingly urgent. Traditional oral English teaching in China has long been limited by a one-size-fits-all teaching model, facing prominent problems such as a lack of oral practice scenarios, insufficient personalized feedback, obvious anxiety about speaking up among learners, and a single

79

evaluation system. These issues have led most learners to fall into the predicament of "dumb English", making it difficult for them to meet the demand for practical English application capabilities in the context of globalization. Generative Artificial Intelligence boasts advantages such as personalized adaptation, real-time interactive feedback, and multi-scenario simulation, which can accurately align with learners' individual differences, break the temporal and spatial constraints of traditional oral teaching, provide a brand-new technical pathway and solution for personalized oral English learning, and drive the transformation of oral English learning from passive input to active output.

2. Current Research on the Integration of Generative Artificial Intelligence into Oral Foreign Language Pedagogy

Foreign research on the integration of Generative Artificial Intelligence and oral English teaching and learning has mostly focused on the application of chatbots in oral practice, exploring their functions in simulated dialogue, pronunciation assistance, multimodal interaction, personalized feedback and potential problems. Kohnke et al. (2023) analyzed the affordances and teaching benefits of ChatGPT and proposed that Generative Artificial Intelligence can be applied to oral teaching scenarios such as the explanation of phonetics knowledge and the guidance of pronunciation skills, providing targeted phonetic assistance for learners (Kohnke, Moorhouse, & Zou, 2023). Wan and Moorhouse (2024) took Call Annie as the research object and found that this intelligent tool developed based on the GPT architecture can effectively improve learners' participation in oral English and fluency of expression with its authentic English accent, vivid animated images, and oral exam simulation functions, but also point out its shortcomings in terms of speech recognition accuracy, depth of emotional response, and technical stability (Wan & Moorhouse, 2024). López-Molines (2025) indicated in the research that with regard to cultivating oral English skills, Generative Artificial Intelligence can enhance pronunciation, fluency and learning motivation, make the learning process more personalized, dynamic and authentic, and ultimately improve overall English communicative competence. Nevertheless, this method has challenges and complexities in terms of reliability, coherence and interaction (López-Molines, 2025).

Domestic research is closely integrated with the actual needs of English teaching in China, mainly involving the exploration of application models, analysis of practical effects, and discussion of existing problems. Li (2024) proposed that Generative Artificial Intelligence can be used to create foreign language conversations, act as a foreign teacher, construct a cross-cultural learning environment, provide personalized learning, and serve as a learning assistant (Li, 2024). Wu, Zhou and Cao (2024) confirmed through empirical research that the interactive environment created by Generative Artificial Intelligence technology helps learners obtain more authentic interactions and improve their language expression abilities, providing a new pathway for oral English teaching (Wu, Zhou, & Cao, 2024). Ge Nan and Zheng Chunping (2025) summarized 17 different forms of teaching activities in 5 categories applied in pre-class, in-class, after-class and extracurricular stages with the support of intelligent chatbots through a systematic review. They pointed out that intelligent chatbots generally help promote

80

the development of learners' learning psychology, learning behaviors, knowledge and skills, and the evaluation and feedback of their various system performances are good, but there are still deficiencies (Ge & Zheng, 2025). Jin and Sun (2025) found through a comparison between the experimental group and the control group that oral practice assisted by Generative Artificial Intelligence can not only stimulate the enjoyment of speaking English but also have a multifaceted impact on foreign language learning (Jin & Sun, 2025). Zou (2025) explored a human-machine collaborative evaluation system of "self-assessment+self-comparison" for oral English based on Generative Artificial Intelligence. By integrating criterion-referenced evaluation and self-comparison feedback mechanisms, the system realizes multi-dimensional dynamic assessment of oral English proficiency, provides a new paradigm for intelligent education evaluation, and promotes the deep integration of information technology and language teaching (Zou, 2025). In addition, domestic research has also paid attention to issues such as the transformation of teachers' roles and the integration of teaching resources, emphasizing that Generative Artificial Intelligence is not a substitute for teachers but an auxiliary tool to drive the innovation of teaching models.

Overall, the application of Generative Artificial Intelligence in personalized oral English learning has received extensive attention at home and abroad, and relevant research has achieved certain results in theoretical frameworks, application models and practical effects, but the research cycle is relatively short. In terms of application tools, most early studies focused on ChatGPT, with few involving other Large Language Models. Compared with the application of Generative Artificial Intelligence in English reading, writing and translation, the number of studies related to oral English is relatively small.

3. Theoretical Foundations of the Application of Generative Artificial Intelligence into Personalized Oral English Learning

3.1 Constructivist Learning Theory

Constructivist learning theory is one of the core theoretical foundations for the application of Generative Artificial Intelligence in personalized oral English learning. Constructivism holds that learning is a process of acquiring knowledge, but knowledge is not obtained through teachers' instruction; instead, it is acquired through meaning construction in a certain context, i.e., a socio-cultural background, with the help of others, the use of necessary learning materials. Therefore, constructivist learning theory argues that context, collaboration, conversation and meaning construction are the four key elements or attributes of learning. Multimedia computers and network communication technologies can serve as ideal cognitive tools in a constructivist learning environment, effectively promoting students' cognitive development. With the rapid development of the educational application of multimedia computers and the Internet, constructivist learning theory is showing its strong vitality and expanding its influence worldwide (He, 1997). This theory emphasizes that learning is a process in which learners actively construct knowledge rather than passively accept it, and learners' learning experiences, individual differences and learning contexts have an important impact on knowledge

81

construction. It is necessary to provide learners with a learning environment for independent exploration and active communication, and help them construct their own knowledge systems through interaction, practice and other means.

Generative Artificial Intelligence is highly consistent with the core concepts of constructivist learning theory. GAI can provide customized learning contexts and resources according to learners' individual differences, such as English proficiency, learning habits and interest preferences, enabling learners to improve their oral English abilities through independent choice and active practice. Unlike traditional fixed teaching contexts, GAI can flexibly adjust scenario details according to learners' feedback, making the learning context more in line with individual needs. GAI simulates oral dialogues in different scenarios, allowing learners to actively participate in interactions in an immersive context and construct knowledge and skills of oral expression through continuous attempts, feedback and adjustment. At the same time, the personalized feedback function provided by GAI can help learners identify their own shortcomings in a timely manner, guide them to reflect actively and optimize independently, so as to realize the active construction of knowledge. In addition, GAI supports the independent learning model which breaks the temporal and spatial constraints of traditional classrooms, allowing learners to arrange their learning time and content according to their own learning rhythm, which fully reflects the learner-centered teaching concept of constructivism and lays a solid theoretical foundation for the personalized application of GAI in oral English learning.

3.2 Second Language Acquisition Theory

Second language acquisition theory is the core theory of foreign language teaching, and its subordinate Input Hypothesis, Interaction Hypothesis, Output Hypothesis and Affective Filter Hypothesis provide important theoretical support for GAI-assisted personalized oral English learning, clarifying the key role of GAI in promoting oral language acquisition.

3.2.1 Input Hypothesis

Proposed by the American linguist Stephen D. Krashen in the late 1970s and early 1980s, the Input Hypothesis holds that the key to language acquisition lies in exposure to "comprehensible input", and the difficulty of the input should be slightly higher than the learner's current language proficiency (the i+1 principle). Excessively simple input (i) or overly difficult input (i+2) cannot promote acquisition (Krashen, 1997). Generative Artificial Intelligence can accurately diagnose learners' oral English proficiency through speech recognition and natural language processing technologies, construct personalized ability portraits, and then dynamically generate oral input content in line with the i+1 principle. This not only prevents basic learners from feeling frustrated when exposed to complex expressions but also avoids advanced learners from repeating simple dialogues without improvement. At the same time, GAI ensures the comprehensibility of input through contextualized dialogue and multimodal prompts, breaking the one-size-fits-all limitation of traditional oral input. For example, GAI can adjust the vocabulary difficulty and sentence complexity of dialogue content in real time according to the learner's pronunciation fluency and response speed, ensuring that the input is always

82

within the learner's acceptable range while promoting progress.

3.2.2 Output Hypothesis

Proposed by the Canadian linguist Merrill Swain in 1985, the Output Hypothesis is a supplement and improvement to Krashen's Input Hypothesis. Swain argues that language input is a necessary but not sufficient condition for language acquisition, and language output is also an important link in second language acquisition (Swain, 1985). Through language output such as oral expression and writing, learners can test their linguistic knowledge, identify their own deficiencies, and thus promote the consolidation and improvement of linguistic knowledge. In traditional oral English teaching, learners have limited opportunities for oral output, while GAI can provide a large number of oral output scenarios for learners. Learners can practice oral English with GAI at any time, consolidate what they have learned through continuous language output, and improve the fluency and accuracy of oral expression. The targeted feedback provided by GAI can help learners reflect on and revise their output, further strengthen the effect of language output, and promote the process of second language acquisition.

3.2.3 Affective Filter Hypothesis

Krashen, the proposer of the Affective Filter Hypothesis, points out that the affective filter is an emotional barrier that prevents learners from fully digesting and absorbing language input in the process of language learning. The affective filter can be divided into a high affective filter and a low affective filter. A low affective filter, such as self-confidence, relaxation and strong interest, can promote the absorption of comprehensible input and improve acquisition effects; while a high affective filter, such as anxiety, tension, inferiority and other negative emotions, will hinder the transformation of input into acquisition (Krashen, 1982). In other words, learners' emotional state will affect the effect of second language acquisition. In traditional oral English teaching, learners often suffer from anxiety about oral expression due to fear of making mistakes and worry about being evaluated, resulting in a high affective filter and reduced efficiency of oral English learning. The one-on-one, stress-free learning environment provided by Generative Artificial Intelligence can effectively lower learners' affective filter. Learners can practice oral English in a relaxed and free atmosphere without worrying about being criticized for making mistakes. At the same time, functions such as encouraging feedback and personalized adaptation of GAI can enhance learners' learning self-confidence, stimulate learning interest, and further promote oral language acquisition. In addition, GAI supports 24/7 uninterrupted interaction, and learners can arrange practice according to their own rhythm, which further relieves learning pressure, reduces the effect of the affective filter, and helps the efficient absorption and application of oral input.

3.2.4 Interaction Hypothesis

Proposed by the American linguist Michael Long in the early 1980s, the Interaction Hypothesis

emphasizes that language acquisition depends not only on comprehensible input but also on the interaction obtained by learners through "negotiation for meaning" in real communication (Long, 1981). In the process of interacting with others, learners can obtain comprehensible input through negotiating meaning, correcting mistakes and other ways, thereby improving their language abilities. Generative Artificial Intelligence provides learners with an all-weather interaction partner by simulating real oral interaction scenarios. Learners can have one-on-one simulated dialogues with GAI, which can timely adjust the difficulty of the dialogue and supplement relevant vocabulary and sentence patterns according to learners' expressions, helping them obtain comprehensible input. At the same time, GAI can correct learners' pronunciation and grammatical errors in real time, and help them negotiate meaning and optimize expressions through interactive feedback, which is in line with the core requirements of the Interaction Hypothesis. This kind of interactive learning not only simulates real communication scenarios but also ensures the pertinence of interaction, making it easier for learners to master oral expression skills.

3.3 Metacognitive Theory

Proposed by Flavell, metacognitive theory holds that metacognition is learners' cognition and regulation of their own cognitive processes, including three core components: metacognitive knowledge, metacognitive experience and metacognitive monitoring (Flavell, 1979). Metacognitive knowledge refers to learners' stored knowledge or beliefs about themselves and others as cognitive subjects, about tasks, about actions or strategies, and about how all these interact to affect the outcomes of any kind of intellectual activity. Metacognitive experiences are conscious cognitive or emotional experiences that occur during the learning process and relate to any aspect of it—often, how well the process is going. Metacognitive theory emphasizes that the cultivation of metacognitive ability is the key to improving learning effects. Teaching should guide learners to learn to reflect on their own learning processes, master methods of independent learning, and improve their self-regulation abilities. Generative Artificial Intelligence can effectively promote the cultivation of learners' metacognitive ability, providing support for personalized oral English learning. In terms of metacognitive knowledge, by analyzing learners' learning data, GAI can provide them with personalized learning diagnosis reports, helping them understand their oral English proficiency, learning strengths and weaknesses, clarify learning goals and directions, and accumulate metacognitive knowledge. These reports can also provide targeted strategy suggestions, helping learners master more effective oral practice methods. In relation to metacognitive experience, functions such as real-time feedback and phased assessment provided by GAI enable learners to feel their learning progress in a timely manner, enhance learning self-confidence, and gain positive metacognitive experiences. For learners who make slow progress, GAI can adjust feedback methods to avoid negative emotions and maintain positive learning motivation. With regard to metacognitive monitoring, GAI can dynamically remind learners to adjust their learning strategies according to their learning progress and performance, guide them to reflect on their oral practice process, timely discover and correct problems in learning, and improve their

84

self-regulation abilities, thus helping learners form independent learning awareness and lay a foundation for lifelong learning.

4. Application Pathways of Generative Artificial Intelligence in Personalized Oral English Learning

4.1 Personalized Learning Demand Analysis

Personalized learning demand analysis is the premise and foundation of Generative Artificial Intelligence-assisted personalized oral English learning, and also a concrete embodiment of the educational concept of "teaching students in accordance with their aptitude" in oral English teaching. Only by accurately grasping learners' individual differences and real learning needs can we break the traditional one-size-fits-all teaching model and provide learners with customized learning plans and resources that meet their actual situation. Relying on its advantages in efficient data collection and analysis, Generative Artificial Intelligence can achieve accurate diagnosis of learners' oral English learning needs, providing a solid data support for the orderly development of personalized learning.

Combined with the practice of oral English teaching, GAI first needs to conduct an initial assessment and diagnosis of learners to fully understand their English foundation, including specific English proficiency levels and core weak links in oral English, such as phonetic symbol pronunciation, grammatical application, fluency of expression, and vocabulary reserve, so as to clarify learning goals and lay a foundation for subsequent teaching adaptation. Then, it tracks learners' learning behaviors and collects various types of data in real time during their oral practice, including practice duration, focus of practice content, common types of errors, and the acceptance and implementation of teaching feedback. Through data analysis, it accurately captures learners' learning habits, interest preferences and learning difficulties. Furthermore, it conducts active feedback and guidance, guiding learners to actively express their own learning needs, learning feelings and puzzles through simple and easy voice interaction and option matching forms, further supplementing and improving the results of demand analysis, and making demand diagnosis more in line with learners' real demands.

On the basis of data collection, GAI conducts in-depth sorting and analysis of various data through algorithm models to construct personalized portraits that conform to learners' actual situation, clarifying their core needs and potential needs. For example, the core needs of learners with a weak English foundation are mostly to consolidate basic vocabulary, correct basic pronunciation, improve the ability of simple daily dialogue, and focus on solving the problem of "not daring to speak and not being able to express". Learners preparing for oral exams focus more on mastering the answering ideas of various oral topics, improving the fluency and logic of expression, being familiar with exam scoring criteria, and making targeted breakthroughs in exam difficulties. The core needs of workplace business English learners are to accumulate workplace-related vocabulary and sentence patterns, improve oral application abilities in scenarios such as business negotiation, work report and customer communication, and meet the practical scenario needs of the workplace. Through such accurate

demand analysis, GAI can provide a clear direction for the subsequent links such as customized learning corpus generation and dynamic adjustment of learning plans, ensuring that personalized learning is always in line with teaching needs and learners' actual situation, and improving the pertinence and effectiveness of teaching.

4.2 Customized Learning Corpus Generation

Customized learning corpus generation is the core link for GAI to realize personalized oral English learning, and also the key to solving the problems of single corpus and poor adaptability in traditional oral English teaching. With its strong content generation ability, GAI can generate oral English learning materials that meet learners' English proficiency, interest preferences and learning goals based on their personalized needs and learning portraits. These materials not only retain professionality but also enhance the pertinence and interest of learning, fully mobilizing learners' enthusiasm to speak up.

In conjunction with the actual needs and learning stages of oral English learning, the customized learning materials generated by GAI can be divided into four categories: basic practice corpus, scenario-based practice corpus, targeted improvement corpus, and multi-form immersive corpus. The basic practice corpus is designed in a hierarchical manner according to learners' English proficiency to ensure that the difficulty of the corpus matches their abilities. Scenario-based practice corpus designs diverse real oral dialogue scenarios combined with learners' learning goals and interest preferences, allowing learners to practice oral English in scenarios close to real life and realize the integration of learning and application. Targeted improvement corpus designs special practice materials for learners' weak links in oral English, making precise breakthroughs in learning difficulties. For example, it generates pronunciation correction tutorials and phonetic symbol practice dialogues for learners with weak pronunciation, and grammar error correction exercises and sentence pattern expansion dialogues for learners with weak grammar. Multi-form immersive corpus combines text, audio, video, animation and other forms to generate immersive learning materials, such as dialogue videos with real-person pronunciation and pronunciation tutorials with animation demonstrations, enhancing the intuition and interest of learning.

The learning corpus generated by GAI also has the advantage of dynamic update. GAI can track learners' learning progress and performance in real time, and dynamically adjust the difficulty, content and focus of the corpus according to their ability improvement. For example, when learners have a solid grasp of basic vocabulary and sentence patterns, GAI will automatically increase the difficulty of the corpus and incorporate more complex sentence patterns, advanced vocabulary and expression skills. When learners' oral expression ability in a certain scenario is significantly improved, it will generate more challenging scenario corpus and supplement relevant extended content, ensuring that the learning corpus is always matched with learners' learning needs and ability levels, and gradually promoting the steady improvement of their oral English ability.

4.3 Interactive and Authentic Learning Experience

Interactive and authentic learning experience is the key to improving the effect of oral English learning

86

and addressing the challenges in traditional oral teaching, in particular, insufficient interaction, single scenario, and learners' reluctance to speak up. In classroom teaching, limited by factors such as class size, teaching time and teaching resources, many learners have difficulty obtaining sufficient one-on-one interactive practice opportunities. GAI, however, can provide learners with an immersive and interactive oral English learning experience by simulating real oral interaction scenarios and building an intelligent interactive interface, effectively improving learners' participation and expression ability in oral English and making up for the deficiencies of traditional teaching.

Coupled with oral English teaching scenarios, the interactive and authentic learning experience rendered by GAI is mainly reflected in two aspects: intelligent virtual teaching assistant interaction and multimodal immersive scenario experience. On the one hand, intelligent virtual teaching assistants can realize one-on-one interaction. GAI builds a virtual teaching assistant role with intelligent guidance functions, and conducts one-on-one simulated dialogue practice with learners through voice interaction, which functions as an exclusive oral English companion. Virtual teaching assistants have strong natural language understanding ability, which can accurately identify the content of learners' oral expression, pronunciation errors and grammatical problems, respond to learners' questions in a timely manner, and flexibly adjust the dialogue rhythm, content difficulty and guidance methods according to learners' expression, simulating real interpersonal interaction scenarios and allowing learners to speak boldly in a stress-free environment. For example, learners can have simulated dialogues with virtual teaching assistants in scenarios such as daily communication, oral Q&A and business negotiation. Virtual teaching assistants will put forward corresponding questions, refutations or supplementary suggestions according to learners' expression ideas, guiding them to improve their expression and enhancing the flexibility, fluency and response ability of oral expression in interaction. On the other hand, multimodal immersive scenario experience relies on multimodal technology of GAI to build virtual oral English learning scenarios. Learners can interact with the scenarios in simple and convenient ways, such as voice and gestures, gaining an immersive learning experience. For example, building virtual scenarios such as overseas travel, workplace office and English interview, learners can have real-time dialogues with virtual characters in the scenarios, complete oral practice such as asking for directions, ordering meals, shopping, work communication and interview responses, accumulate scenario-based vocabulary and sentence patterns in real scenario simulation, improve oral application ability, and realize the application of what has been learned.

In addition, the interactive learning experience of GAI also has the advantages of being stress-free and available 24/7, which is highly in line with learners' independent learning needs. In traditional teaching, some learners suffer from anxiety about oral expression and dare not take the initiative to speak for fear of being criticized for mispronunciation or inarticulate expression. In GAI interaction scenarios, however, learners can carry out oral practice at any time without worrying about making mistakes, which effectively reduces expression anxiety and enhances learning self-confidence. At the same time, GAI can deliver targeted feedback and guidance in real time according to learners' interactive

performance, helping them discover problems and optimize expressions in a timely manner, truly realizing learning in interaction and improvement in practice.

4.4 Intensive Language Application Practice

The improvement of oral English ability is inseparable from sufficient practical practice. Intensive language application practice is an important link to consolidate the achievements of oral English learning and improve oral application ability, and also one of the core goals of oral English learning. Traditional oral English teaching is characterized by monotonous practice settings and insufficiently tailored practice tasks, consequently resulting in a "high-input-low-output" phenomenon where learners struggle to apply acquired knowledge flexibly in authentic contexts. This systemic limitation hinders the development of communicative competence, as the disconnect between structured learning environments and real-world language demands prevents effective transfer of linguistic skills. However, GAI can provide diverse practical scenarios and personalized practical tasks along with learners' personalized needs, guiding them to apply the learned oral English knowledge and skills to real expression, and realizing the all-round improvement of oral English ability.

In accordance with the idea of hierarchical practice and precise reinforcement, the intensive language application practice which GAI furnishes can cover various aspects. The arrangement of personalized practical tasks combines learners' learning goals, weak links and ability levels to assign highly targeted oral practical tasks for them, clarifying task requirements, completion criteria and evaluation dimensions, so that learners have a clear practical direction. For example, assigning targeted tasks such as topic speeches, viewpoint discussions and scenario Q&A for learners preparing for oral exams, in line with exam requirements; allocating tasks such as imitating daily dialogues, applying vocabulary and sentence patterns, and expressing simple topics for primary learners, focusing on consolidating basic abilities. Real scenario simulation practice generates various types and difficulty levels of real oral English application scenarios through GAI, allowing learners to carry out practical practice in a realistic environment, familiarizing them with the oral expression norms, tone characteristics and communication skills of different scenarios, and improving their scenario adaptability and practical application ability. It enables learners to accumulate relevant experience in simulated practice, break through difficulties and ensure that they can flexibly respond in real scenarios. GAI-powered multimodal practice feedback mechanisms enable precise, data-driven assessment of learners' oral performance across multidimensional criteria, including phonetic accuracy, grammatical competence, fluency, semantic coherence, and logical structuring. By delivering comprehensive, context-specific feedback and generating actionable recommendations, GAI systems facilitate metacognitive awareness of individual proficiency gaps while delineating targeted improvement pathways to optimize practice efficacy. Furthermore, GAI supports both autonomous and collaborative learning paradigms: learners can engage in self-regulated scheduling of practice sessions aligned with personal learning rhythms, while the system simultaneously constructs virtual collaborative clusters for interactive speaking tasks.

Through simulated peer groups enabling collaborative discourse, e.g., structured debates, role-playing simulations, GAI fosters the development of pragmatic communication skills and interlocutor adaptability, thereby enhancing the transferability of oral competencies to real-world communicative contexts.

4.5 Intelligent Assessment with Real-time Feedback

Intelligent learning assessment and feedback is an important guarantee for GAI to assist personalized oral English learning. In traditional oral teaching, the assessment method is relatively single, mostly based on teachers' subjective evaluation. The assessment results lack scientificity and comprehensiveness, and the feedback is delayed and the content is general, making it difficult to help learners accurately identify their own deficiencies. With its strong speech recognition and natural language processing capabilities, GAI can realize real-time, scientific and comprehensive assessment of learners' oral English learning effects, and provide personalized feedback and improvement suggestions, helping learners timely optimize learning strategies and improve learning effects.

With respect to learning assessment, aligned with the assessment criteria of oral English teaching, GAI adopts diversified assessment indicators to build a scientific and perfect oral assessment system, and strives to ensure that the assessment results are comprehensive, objective and accurate. Specifically, pronunciation accuracy compares learners' pronunciation with standard pronunciation through professional speech recognition technology, comprehensively assesses their mastery of pronunciation skills such as phonetic symbols, intonation, liaison and weak form, and accurately identifies pronunciation errors and weak points. Grammatical normativeness carefully analyzes the content of learners' oral expression with natural language processing technology, accurately identifies common grammatical problems, for instance, tense errors, subject-verb disagreement, preposition misuse and mixed sentence patterns, and assesses the normativeness of learners' application of grammatical knowledge. Expression fluency objectively assesses the fluency of learners' oral expression by analyzing data such as the speed of oral expression, number of pauses, number of repetitions and duration of hesitation, and identifies the key factors affecting fluency. Content integrity and logic conduct in-depth analysis of the content of learners' oral expression, assess the integrity of content, clarity of viewpoints and coherence of logic, and judge learners' language organization ability and expression ideas. Through the comprehensive analysis and quantitative assessment of various indicators, GAI generates a personalized assessment report for each learner, clearly presenting their learning progress, strengths and deficiencies, so that both learners and teachers can fully grasp the learning situation.

In terms of learning feedback, GAI delivers real-time and targeted feedback, which is different from the delayed and single feedback in traditional oral assessment. GAI can intervene in real time during learners' oral practice and correct pronunciation, grammatical and other errors in a timely manner. At the same time, according to the comprehensive assessment results, GAI provides learners with personalized improvement suggestions that are in line with their weak links and have strong operability.

For example, for learners with insufficient expression fluency, it is suggested that they strengthen follow-up and imitation practice, accumulate common sentence patterns and connectives, and reduce expression pauses. For learners with insufficient content logic, it is suggested that they sort out expression ideas in advance, divide expression levels reasonably, and flexibly use connectives to improve the coherence of expression. In addition, GAI will dynamically adjust the feedback method, frequency and content according to learners' acceptance of feedback and improvement effects, ensuring the effectiveness and pertinence of feedback, and truly helping learners solve learning problems.

4.6 Dynamic Adjustment of Learning Plan

Traditional oral English learning plans are mostly fixed models, while learners' oral English ability, learning needs and learning status are in a dynamic state. Therefore, dynamic learning plan adjustment provides strong support for the realization of personalized oral English learning. Based on learners' real-time learning data, assessment results and changes in learning needs, GAI can dynamically adjust learning plans to ensure that they are always matched with learners' actual situation, improve the pertinence and effectiveness of learning, and promote the steady improvement of their oral English ability.

The process of GAI's dynamic adjustment of learning plans mainly includes three stages: real-time data tracking, learning status analysis and dynamic plan optimization. In the real-time data tracking stage, GAI continuously collects various types of data throughout the learning process, including learning progress, oral practice situation, assessment results, acceptance and implementation of feedback, learning duration and changes in learning interest, to grasp learners' learning status, ability improvement and changes in learning needs in real time. In the learning status analysis stage, GAI conducts in-depth sorting and analysis of the collected real-time data through algorithm models to comprehensively judge learners' learning effects and accurately identify problems in the learning process. For instance, whether learners have an unsteady grasp of a certain knowledge point, overly fast or slow learning progress, declining learning interest, or failure to effectively break through a certain weak link, providing a scientific basis for plan adjustment. In the dynamic plan optimization stage, targeted adjustments are made to the learning plan according to the results of learners' learning status analysis, focusing on the three core aspects of learning content, learning rhythm and learning methods to ensure that the adjusted plan is in line with learners' actual situation.

Specifically, the adjustment of learning content optimizes the difficulty, focus and emphasis of learning materials combined with learners' assessment results and weak links, ensuring that the learning content is always in line with their ability level and improvement needs. For example, when learners master basic scenario dialogues, learning content involving complex scenarios and extended topics will be added to expand the scope of their oral expression. The adjustment of learning rhythm reasonably adjusts the quantity, difficulty and completion time limit of learning tasks combined with learners' learning progress, learning efficiency and learning status, avoiding excessive or insufficient learning pressure. The adjustment of learning methods optimizes learning methods combined with learners'

learning interest, learning habits and learning effects to stimulate their learning motivation. For example, when learners' interest in simulated dialogue declines, GAI will introduce diversified learning methods such as gamified learning, topic debate and situational interpretation to enhance the interest of learning. Through such dynamic adjustment of learning plans, GAI can provide each learner with the most suitable learning pathway for their actual situation, ensure the sustainability and effectiveness of personalized learning, truly achieve "teaching students in accordance with their aptitude", and efficiently help learners improve their oral English ability.

5. Application Cases of Generative Artificial Intelligence in Personalized Oral English Learning

The rapid development of Generative Artificial Intelligence brings forth a variety of language learning tools based on Generative Artificial Intelligence large models. To further illustrate the application of GAI tools in personalized oral English learning, four representative GAI oral English learning tools—Call Annie, Doubao, iFlySpark, and Duolingo—are selected for in-depth analysis, providing practical references for the practical application of GAI-assisted personalized oral English learning and helping learners choose suitable tools according to their own needs.

Case 1: Call Annie

Based on the GPT architecture, Call Annie is a professional Generative Artificial Intelligence oral English learning tool whose core advantages lie in high-quality multimodal interaction and realistic scenario simulation. With its authentic native English accent, vivid animated images and natural dialogue rhythm, it provides learners with an immersive oral English learning experience that is close to real communication. Its most prominent features are the authenticity of scenario simulation and the idiomaticness of oral expression, which is mainly suitable for learners with intermediate and advanced oral English foundation who pursue fluency and idiomaticness of oral English, especially for candidates preparing for oral exams such as IELTS and TOEFL, which can help them adapt to the exam scenario and improve their exam performance.

Call Annie does not require complex initial settings and can quickly identify learners' English proficiency, oral expression characteristics and accent habits through their first dialogue, constructing accurate personalized learning portraits without tedious operation. In terms of scenario simulation, Call Annie supports the simulation of a variety of real and detailed scenarios, including daily communication, business negotiation, academic speech and English interview. Learners can freely choose scenarios they are interested in or have urgent needs for, and have one-on-one simulated dialogues with Call Annie. According to the scenario characteristics and learners' expression content, Call Annie can flexibly adjust the dialogue content, tone and speaking speed, simulate real and natural interpersonal interactions, and even simulate interlocutors with accents from different English-speaking countries, such as American, British and Australian accents, helping learners adapt to English expressions with different accents and improve their listening and speaking adaptability. In terms of pronunciation and expression guidance, Call Annie adopts high-fidelity real-person recordings of native

91

speakers. Learners can improve the idiomaticness of their pronunciation and intonation through follow-up and imitation. At the same time, Call Annie can correct learners' expression errors in real time, provide more natural and idiomatic sentence patterns and vocabulary replacement suggestions, and help them optimize their oral expression and get rid of the interference of Chinglish. In terms of personalized practice, Call Annie can generate targeted practice tasks according to learners' dialogue performance, assign special correction exercises for high-frequency errors in expression to strengthen learning effects, and the difficulty of practice will gradually increase with the improvement of learners' abilities, ensuring that learners can make steady progress.

Case 2: Doubao

Developed independently by ByteDance based on the Doubao-Seed large model base, Doubao is a multi-functional Generative Artificial Intelligence tool with strong adaptability. In terms of personalized oral English learning, relying on its powerful natural language processing, high-precision speech recognition and intelligent content generation capabilities, Doubao has built a full-process closed-loop personalized learning system of "demand analysis—material generation—interactive practice—assessment feedback—plan adjustment", which is suitable for learners with different English proficiency levels, from beginners to advanced learners, and diverse learning goals, such as daily communication, exam preparation, and business negotiation, and has strong versatility and a wide range of adaptation.

First of all, through a scientific initial assessment including basic pronunciation questions, sentence pattern application questions and learning goal matching questions, Doubao accurately grasps learners' English proficiency, weak links in oral English and specific learning goals, and quickly constructs detailed personalized learner portraits. Then, based on the learner portraits, it generates highly customized oral English learning materials, including basic pronunciation practice, scenario dialogues and special improvement materials. It can flexibly match different types of oral materials, such as life-oriented, academic and business materials according to learners' specific needs, and the difficulty of the materials will be dynamically adjusted according to learners' learning performance, ensuring that the learning content is neither too simple nor too difficult. In terms of interactive practice, Doubao supports one-on-one real-time voice dialogue, can simulate interpersonal interactions in a variety of common scenarios, accurately identify learners' oral expression errors, respond to questions and correct errors in a timely manner, and provide positive and encouraging feedback at the same time, effectively reducing learners' expression anxiety and enhancing their confidence in speaking English. In terms of assessment and feedback, Doubao can conduct real-time and comprehensive assessment of core indicators such as learners' pronunciation, grammar, fluency and idiomaticness, generate a detailed assessment report including error details and improvement directions, and provide targeted improvement suggestions, helping learners accurately solve their oral English shortcomings. In terms of plan adjustment, Doubao can dynamically adjust learning content and rhythm according to learners' learning progress and assessment results, ensuring the pertinence and effectiveness of learning. In

92

addition, Doubao supports multimodal interaction (voice, text, picture) and 24/7 unlimited use, facilitating learners to use fragmented time for oral practice, further improving the convenience of learning, fully covering learners at different levels, and meeting various oral English learning goals.

Case 3: iFlySpark

Developed by iFlytek, a leading enterprise in speech recognition, iFlySpark is a Generative Artificial Intelligence tool with speech technology as its core. In terms of personalized oral English learning, relying on its industry-leading speech recognition and speech synthesis technologies, it focuses on precise pronunciation correction and targeted special improvement. It pays more attention to consolidating the foundation of oral English, suitable for learners with different English proficiency levels, especially for those with weak pronunciation foundation who need to improve their oral English foundation in a targeted manner, and is a core auxiliary tool for breaking through the bottleneck of oral English foundation and improving the accuracy of pronunciation.

The core advantage of iFlySpark lies in its high-precision pronunciation recognition and professional correction. Its advanced speech recognition technology can accurately capture the subtle details of learners' pronunciation, for example, the difference between /θ/ and /ð/, and the stress of words, compare it with standard native pronunciation, accurately identify various common pronunciation problems, and provide personalized feedback such as real-person pronunciation demonstrations, pronunciation mouth shape guidance and slow-speed reading, helping learners intuitively understand their own problems and correct pronunciation errors efficiently. In terms of customized learning materials, iFlySpark can generate targeted special pronunciation practice materials according to learners' English proficiency and weak pronunciation links, and can match corresponding follow-up and comparison exercises for different weak points such as phonetic symbols, stress and intonation, helping learners consolidate their weak links in a targeted manner. In terms of interactive practice, iFlySpark supports one-on-one voice dialogue and follow-up practice. Learners can follow the standard pronunciation to practice, and the system evaluates the follow-up effect in real time, gives accurate pronunciation accuracy feedback and targeted improvement suggestions, and helps learners adjust their pronunciation in time. In terms of learning plans, iFlySpark can dynamically adjust practice content and difficulty according to the improvement of learners' pronunciation, gradually helping them consolidate the foundation of oral English, improve the accuracy of pronunciation, and lay a solid foundation for fluent oral expression.

Case 4: Duolingo

Integrating advanced Generative Artificial Intelligence technology, Duolingo is a popular language learning tool that features gamified learning in personalized oral English learning and focuses on improving basic oral English ability. It is especially suitable for primary English learners and beginners who have just started to learn oral English, as it can effectively stimulate learning interest, reduce the difficulty of learning, and help learners consolidate the foundation of oral English in a relaxed and pleasant atmosphere.

At the beginning, through a simple and interesting initial assessment, Duolingo quickly determines learners' English proficiency and learning rhythm, and formulates a scientific and personalized basic oral English learning plan for them, which is in line with the learning rules of beginners. In terms of learning methods, Duolingo adopts an innovative gamified design, integrating oral practice into interesting game links such as level breaking, point accumulation, daily clock-in and friend ranking, which greatly enhances the interest of oral learning, stimulates learners' learning motivation and enthusiasm, and helps beginners develop the habit of persistent learning. In terms of customized learning content, Duolingo gradually generates simple and practical learning content such as basic vocabulary, common sentence patterns and short dialogues according to learners' learning progress, focusing on cultivating their ability of simple oral English expression, such basic topics as daily greetings, numbers, colors, diet and family. In terms of interaction and feedback, Duolingo supports simple and easy-to-operate interactive practice, such as voice follow-up and oral blank filling. The system evaluates learners' oral expression in real time and provides immediate feedback and prompts. At the same time, it flexibly adjusts the difficulty of learning content according to learners' level-breaking situation, ensuring that learners can improve steadily without feeling too much pressure. In addition, Duolingo also supports personalized learning reminders and progress records, helping learners develop regular learning habits and better promote the improvement of basic oral English ability.

6. Conclusion

This paper conducts a systematic research on the application of Generative Artificial Intelligence (GAI) in personalized oral English learning. In combination with the core connotations of constructivist learning theory, second language acquisition theory and metacognitive theory, it reviews the current research of the integration of GAI with oral English teaching at home and abroad, explores the application pathways of GAI-assisted personalized oral English learning, and indicates the feasibility of GAI in solving the predicament of traditional oral teaching and meeting learners' personalized needs through the case analysis of typical GAI oral English learning tools.

With its strong natural language processing, speech recognition and content generation capabilities, Generative Artificial Intelligence builds a learner-centered full-process personalized learning system through six core pathways: personalized learning demand analysis, customized corpus generation, interactive and authentic experience, intensive application practice, intelligent assessment and feedback, and dynamic plan adjustment. It realizes the transformation of oral English learning from passive input to active output, and provides learners with a highly adaptable, convenient and efficient solution for oral English learning.

Generative Artificial Intelligence has injected new impetus into personalized oral English learning, and its deep integration with oral teaching is an inevitable trend and important approach of educational intelligence. Personalized oral English learning based on Generative Artificial Intelligence conforms to

94

learners' actual oral English learning needs, is conducive to improving their oral English learning effects, cross-cultural communicative competence and international literacy, and has broad application prospects. Future research can further focus on the existing limitations in the application of GAI, such as insufficient cultivation of cross-cultural communicative competence and lack of guidance on critical thinking, explore more targeted optimization strategies, and combine empirical research to quantify the impact of GAI on oral English learning effects, providing a more solid theoretical support and practical reference for the standardized and in-depth application of Generative Artificial Intelligence in the educational ecosystem.

References

Cyberspace Administration of China. Interim Measures for the Administration of Generative Artificial Intelligence Services. (2023). http://www.cac.gov.cn, 2023-07-13.

Flavell, J. H. (1979). Metacognition and Cognitive Monitoring: A New Area of Cognitive–developmental Inquiry. *American Psychologist*, *34*(10), 906-911.

Ge, N., & Zheng, C. P. (2025). A Systematic Review of Intelligent Chatbots Supporting Oral English Teaching. *Foreign Language Teaching*, *46*(03), 56-63.

Han, X. T. (2022). Misinterpretation, Clarification and Breakthrough Pathway of Personalized Learning Technology in the Big Data Era. *Educational Technology Research*, 43(6), 25-31, 60.

He, K. K. (1997). Constructivism: The Theoretical Foundation for Reforming Traditional Teaching (I). *Educational Technology Research*, *1997*(03).

Jin, Y. X., & Sun, F. J. (2025). A Study on the Intervention of College Students' Enjoyment in Oral English Based on Generative Artificial Intelligence. *Foreign Language Education Research Frontiers*, *8*(03), 97-109.

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). Chat GPT for Language Teaching and Learning. *RELC Journal*, *54*.

Krashen, S. D. (1982). Principles and Practice in Second Language Acquisition. Oxford: Pergamon Press

Krashen, S. D. (1997). *Some issues relating to the monitor model[C]//Brown H, Yorio C, Crymes R. Teaching and Learning English as a Second Language: Trends in Research and Practice: On TESOL 77: Selected Papers from the Eleventh Annual Convention of Teachers of English to*

Speakers of Other Languages, Miami, Florida, April 26 – May 1, 1977 (pp. 144-158). Washington, DC: Teachers of English to Speakers of Other Languages/

Li, Z. W. (2024). Empowering Foreign Language Teaching with ChatGPT: Scenarios and Strategies. *Journal of Beijing International Studies University*, *2024*(01).

Long, M. H. (1981). Input, Interaction, and Second-Language Acquisition. *Annals of the New York Academy of Sciences*, *379*(1), 259-278.

López-Molines, L. (2025). A Systematic Review of Generative Artificial Intelligence-based Tools to Improve Oral Skills in English as a Foreign Language. *The EuroCALL Review*, *32*(2), 183-195.

Swain, M. (1985). *Communicative competence: Some roles of comprehensible input and comprehensible output in its development[C]//Gass S, Madden C. Input in Second Language Acquisition* (pp. 235-253). Rowley, Mass: Newburry House.

UNESCO. Guidance for Generative Artificial Intelligence in Education and Research. (2023). https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research, 2023-09-07.

Wan, Y., & Moorhouse, B. L. (2024). Using Call Annie as a Generative Artificial Intelligence Speaking Partner for Language Learners. *RELC Journal*, *2024*(2).

Wu, J. H., Zhou, W. T., & Cao, C. (2024). An Empirical Study on Generative Artificial Intelligence Technology Empowering Oral English Teaching. *China Educational Technology*, *2024*(04), 105-111.

Zhu, Y. H., Zhang, J. X., & Han, X. B. (2025). A New Form of Personalized Learning Based on Generative Artificial Intelligence. *Educational Technology Research*, *46*(04), 58-64.

Zou, J. J. (2025). Research on Human-machine Collaborative Evaluation of "Self-assessment+Self-comparison" for Oral English Supported by Generative Artificial Intelligence. *Modern English*, *2025*(14), 68-71.