

Unveiling AI's Accent in Spoken English Education

Xinjie Jiang^{1,a,*}

¹*School of Foreign Studies, Nanjing University, Nanjing, 210000, China*

a. 221108052@smail.nju.edu.cn

**corresponding author*

Abstract: Viewed against the backdrop of the technological landscape, artificial intelligence has become the shining point of the 21st century. In recent years, AI has been pervasively applied into various realms in modern society, among which the incorporation into spoken English education has become one of the most heated topics. To navigate the application of AI in spoken English learning, this paper evaluates the effectiveness based on the Interaction Hypothesis proposed by Michael Long. To be specific, this research focuses on the comprehensive application of AI in the realm of spoken English acquisition, underscoring why AI language models have gained in popularity in innumerable apps concerning language learning and how AI serves as a double-edged sword in spoken English education. By illustrating the language acquisition process, a strong connection between the Interaction Hypothesis and AI-based spoken English learning is constructed and clarified. Through all-round evaluation and analysis, the article concludes that despite challenges, AI has not only acted as a knowledgeable mentor for individual learners, but also provided an effective driving force for the innovation and improvement of spoken English education. In the future, more advancements concerning human-like interactions and immersive learning environment are likely to emerge.

Keywords: Spoken English education, AI language model, second language acquisition, language learning application.

1. Introduction

As AI has been introduced into spoken English education, numerous language learning applications have leveraged AI language models to enhance users' spoken English proficiency [1]. Through providing interactive platforms that simulate real-life conversations and offering instant feedback on pronunciation, grammar, and vocabulary usage, they offer learners unprecedented opportunities to access personalized learning experiences tailored to their individual needs and learning pace. Duolingo, Itongzhuo and other popular apps among students learning spoken English have gained in popularity because of their wise incorporation of AI assessment and speakers. Taking Duolingo as an example. It enables users to communicate one-to-one with intelligent robot which not only walks users through correcting their linguistic and grammatical mistakes, but also generates smooth and coherent conversations [2].

Based on the Interaction Hypothesis proposed by Michael Long, this research navigates the application of AI technology in spoken English learning, including its effectiveness, drawbacks and future development. Dividing the AI-based learning process into three parts, it respectively examines

the effectiveness of AI-driven tools in offering personalized learning experiences, constructing a positive learning environment as well as acting as a trustworthy practicing partner. Additionally, the study addresses potential drawbacks such as the limitations in contextual understanding and the need for human interaction to develop nuanced language skills. Looking forward, the research explores future developments in AI technology aimed at further refining its role in facilitating effective spoken English learning, emphasizing ongoing advancements in the application of AI virtual reality technology that creates an immersive environment for learners.

2. Viewing AI language Tools through Interaction Hypothesis

In the following case, the AI model in a popular app named TalkAI engages in a role-play practice about the weather, generating responses and giving feedback. The prompt is recorded, some crucial comments are selectively presented, and the final assessment is excerpted.

Table 1: Role-playing practice with TalkAI.

Prompt: Now let's start a role play exercise for spoken English practice. I am a college student preparing for IELTS oral test, while you are an IELTS examiner. Our discussion topic is about weather, one of the most common and typical IELTS oral test topics. Please give me some feedbacks and assessments of my language performance and try to point out what mistakes I have made in my conversations.
Comments: Great start! Just a minor tip: instead of "in sunny days," you could say "on sunny days." Keep going, and remember to use a variety of vocabulary. How do sunny days affect your mood? Well done! You're using a good range of vocabulary. Just a small correction: "upon my mood" can be "on my mood." Also, try to use more complex sentence structures occasionally. Like, "Moreover, the vibrant sunshine contrasts sharply with the gloom of rainy days, enhancing my sense of well-being." Do you think weather affects other people's moods too? Absolutely, the weather can lift spirits. Just a slight tweak: "when I enjoy sunny days, I notice that people around me are also very energetic and content with the weather." You're doing well. Let's switch topics slightly—do you think weather affects your productivity or ability to focus?
Final assessment: Overall, excellent job! Your pronunciation is clear, and you've used a wide range of vocabulary. Grammar is mostly correct, but be mindful of prepositions. Try varying sentence structures more. Great engagement! Keep practicing!

In the above example, AI language model acts as three kinds of roles in the conversation, language advisor, language partner, and language assessment specialist [3].

The AI model offer suggestions for specific use of words, phrases and sentence structures. For example, some fundamental grammar mistakes like misuse of prepositions in phrases “in sunny days” and “upon my mood” have been pointed out, while more diver sentence patterns like “the vibrant sunshine contrasts sharply with” are suggested.

AI model acts as language partner to provide assistance for students to practice spoken English. Instead of merely correcting linguistic mistakes and standardize expressions, AI generate positive comments and responses, such as “Great start!” “Well done!” and “Absolutely, the weather can lift spirits”.

In the final assessments, the AI model gives an overall comment upon language performance as a language assessment specialist, including pronunciation, grammar and sentence structure.

In the above case of AI-based spoken English practice, it is obvious that the participation of AI has contributed to the personalized learning experience and real-time feedback of spoken English

practices. Michael Long's Interaction Hypothesis gives in-depth theoretical explanations for the positive impacts of AI in spoken English education.

This research is conducted on the basis of the Interaction Hypothesis proposed by professor Michael Long in 1996. Applied into the acquisition of a second language, this theory has combined the Input Hypothesis and the Output Hypothesis [4]. According to Krashen, the Input Hypothesis highlights the significance of the More Knowledgeable One during language learning. The MKO can be a teacher, parent, coach, or anyone who provides guidance [5]. In this context, the MKO is AI language model which provides learners with personalized grammatical and linguistic knowledge of second language [6]. The Output Hypothesis, on the other hand, posits the indispensable role of continuous output in second language learning [7]. When engaging with AI, users find themselves presented with ample opportunities to practice and enhance their language skills. Based on the above-mentioned theories, Michael Long underscores how interaction works throughout the input and output of language acquisition process. The AI-powered applications, according to each user's language performance and learning targets, provide personalized guidance and create virtual environments rich in language input and interactive opportunities, simulating real-life conversations and promoting active engagement. For example, Itongzhuo, which stands out as a prominent tool, harnesses AI to assess and provide feedback on spoken English proficiency. Based on the communications between users and their foreign teachers, it can provide instant assessment. For example, Itongzhuo is the first choice for people preparing for the IELTS oral test. It gives advices primarily on four aspects, including fluency, vocabulary, grammar and pronunciation, alleviating users' stress and embarrassment to a great extent.

During the communication, AI always tend to offer positive comments and feedbacks instead of criticism, preventing possible negative effects of anxiety or tension and stimulating learners' passion for continuous learning. Another example is Duolingo. Duolingo Max represents a collaborative effort between Duolingo and OpenAI, offering users a unique platform for practicing conversational English. By integrating OpenAI's language generation models, Duolingo Max enables learners to engage in lifelike dialogues with AI, providing a positive learning environment, for that every comment and feedback of AI are designed to be comforting and encouraging. Therefore, compared to interactions with human speakers, conversing with AI unequivocally involves less mental pressure and a more relaxed environment conducive to language output. What's more, free of time and access limit, users can utilize the easily-accessible AI tool to practice their spoken English at any time, so as to enhance their language ability with intense output.

3. Drawbacks of AI-based Language Learning

Despite power and promises, there are unavoidable drawbacks of AI-based spoken English learning, especially some common issues that might hinder the success of AI-based oral English learning.

Among all the limitations and challenges faced by the AI language model, over-reliance on technology is the one that poses formidable threats to the cultivation of students' learning habits and abilities. As convenient as AI tools, learners might prefer to communicate with them more frequently than human speakers. if one cannot learn without the help of AI tools, there is high chance that they may get accustomed to a passive learning approach, hindering their ability to engage in face-to-face interactions.

Apart from that, dialogues with AI differ significantly from human-to-human interactions, particularly in terms of emotional resonance and cultural nuances. In terms of semantic understanding, though AI models are capable of generating conversations that are similar to human language, problems arise because they response based on their database instead of semantic interpretation [8]. Such limitation is particularly evident when dealing with texts that convey affection and express abstract conceptions. While AI takes more responsibility for conversational tasks nowadays, such as

asking questions about a topic or answering a question on an online forum, when compared to interacting with an actual person, computers are not as understanding or intelligent as humans. While AI can simulate conversations, they may lack the empathy and cultural awareness inherent in human communication. AI also tends to be less sensitive towards specific emotions such as fear, anger, sadness, disgust and so on. For example, IELTS examiners usually prefer more humorous and vivid conversations, while AI tends to help learners use standard and even rigid language patterns for its lack of emotional resonance.

The limitations of understanding semantic and emotional meaning can also be attributed to its ability of interpreting human tones and moods [9]. Despite their high accuracy and fluency, speech recognition system still lags behind human speakers. As a result, they are incapable of responding to students' emotions reflected in conversations as sensitive as human.

Thirdly, no database is entirely error-free. It is based on the internet that AI database is collected and organized, which inevitably encompasses some fake or biased statistics and researches. When generating feedbacks and comments based on these questionable and stereotyped information, AI models themselves won't try to check its origins and authenticity. Therefore, when students communicate with AI tools, some of the feedbacks and responses they get may be untrustworthy, culminating in the second spreading of misleading data and research results. Additionally, language is incredibly nuanced and context-dependent, which can sometimes pose challenges for AI models to interpret accurately. Mistakes occur due to errors in the training data, ambiguous language usage, or evolving language conventions.

Overall, AI language models are still incomplete replacement for human speakers in the realm of spoken English education that still needs further enhancements and improvements.

4. Future Advancements of AI Tools in Spoken English Learning

To address these challenges, future enhancements in AI models should focus on increasing adaptability and fostering a more nuanced understanding of learner responses. By incorporating sentimental analysis and contextual understanding capabilities, AI systems can deliver more personalized and empathetic interactions, bridging the gap between virtual and real-world communication.

What's more, teaching mode with both teachers and AI language models may complement each other, sparking students' interests and passion [10]. For example, in traditional English classes, students hardly have the chance to engage in one-to-one communication with their teachers. However, with the incorporation of AI tools, students have the opportunity to practice the expressions they learn from teachers and acquire fundamental communication skills during interactions with AI. Such cooperation with AI requires English teachers to adjust their teaching mode, transforming from knowledge transmitters to educational guides. In conjunction with AI tools, teachers can personalize learning methods and suggestions for students and set individualized goals for each learner according to their language performance in the communications with AI.

Concerning AI's limitations in understanding sentimental and contextual conversations, the participation of teachers also plays an indispensable role in providing learners with emotional support and guidance [10]. In case learners depend too much on AI language models, teachers can provide students with weekly chances to communicate with human speakers such as their peers. For example, English corner that holds every week for students to exchange ideas and express feelings.

Furthermore, advancements in AI technology may enable the development of immersive virtual environments that replicate authentic cultural experiences, enriching language learning through multisensory engagement. The recent breakthrough made by BUPT (Beijing University of Posts and Telecommunications) has echoed with the future expectations of AI-based spoken English learning. In the Situational English Virtual Simulation experiment conducted by their research team, the

artificial intelligence virtual reality technology was incorporated to cultivate students' language communication skills adjustable to the context. They designed travel routes from China's Capital International Airport to the major cities in the United States, encompassing tactile vocabulary recognition, situational interaction, summary writing and role-playing. Relying on the immersive language learning environment and human-computer interaction experience, learners enjoy personalized language teaching.

5. Conclusion

In conclusion, the integration of AI in spoken English teaching represents a significant shift in language education. Based on the theoretical framework of the Interaction Hypothesis proposed by Michael Long, the application of AI in spoken English learning has been proved to be conducive for the input and output of language learners. By leveraging AI-powered language models, educators can create dynamic environments to help linguistic development and promote cultural understanding. As AI technologies evolve, they offer opportunities to enhance language acquisition through personalized tutoring systems, adaptive learning platforms, and real-time communication tools.

Still, some problems arise during the practice of AI language model. For example, over-reliance on technology during learning process, alienation from human communications and errors in AI database. Considering the further development of AI technology in spoken English education, some measures must be taken. Instead of repeatedly providing rigid standard patterns of spoken English practice, AI must be able to cultivate sentimental and contextual abilities through constant imitation and advancements. Moreover, to apply AI in real living environments might refrain learners from being limited in the mindset and language patterns of virtual AI conversations.

While challenges persist, the potential for future advancements holds promise for further enriching the language learning journey. More advanced versions of ChatGPT are invented, promoting the application of AI language model in spoken English education. Additionally, more institutions and organizations begin exploiting potentialities of AI language education. Ongoing research and development in AI-driven language learning are paving the way for more effective methodologies and tools that address diverse linguistic needs and challenges.

References

- [1] Zhang, S. 2024. *From questioning, rejecting to accepting, AI language teaching products are competent enough to serve as "practice companions"*. Computer Newspaper, 008 Edition.
- [2] Duolingo Team. 2023. *Introducing Duolingo Max, a learning experience powered by GPT-4* <https://blog.duolingo.com/duolingo-max/>
- [3] Xu Jiajin, Zhaochong. *The role of language model in English education[J]. Foreign Language Education in China*. 2024, pp.3-10.
- [4] Shawn L., Masatoshi S. 2018. *Language Teaching, Volume 51, Issue 3. pp. 285-329* DOI: <https://doi.org/10.1017/S0261444818000125>
- [5] Zafar, Manmay. 2009. *Monitoring the 'monitor': A critique of Krashen's five hypotheses*. Dhaka University Journal of Linguistics 2(4). 139-146.
- [6] Luo, Zixu. 2024. *A Review of Krashen's Input Theory*. Journal of Education, Humanities and Social Sciences.
- [7] Krashen, Stephen D. 1982. *Principles and practice in second language acquisition*. Oxford: Pergamon. http://www.sdkrashen.com/Principles_and_Practice/Principles_and_Practice.pdf.
- [8] Baskara F & Mukarto F. *Exploring the implications of ChatGPT for language learning in higher education[J]. Indonesian Journal of English Language Teaching and Applied Linguistics*, 2023, 7(2) : 343-358.
- [9] Kim S, Shim J & Shim J. *A study on the utilization of OpenAI ChatGPT as a second language learning tool[J]. Journal of Multimedia Information System*, 2023, 10(1) : 79-88.
- [10] Suqi. *The application effectiveness analysis of large language model in second language teaching[J]. Foreign Language World*. 2024, pp.35-41.