Curriculum and Teaching Methodology (2025)
Clausius Scientific Press, Canada

DOI: 10.23977/curtm.2025.080316
ISSN 2616-2261 Vol. 8 Num. 3

The Impact of AI Tools on English Speaking Learning: A Case Study of iFLYTEK Spark

Dong Xingrui[1,a], Liang Ling[1,b,*], Pan Zhiwei[1,c], Shi Xiaolin[1,d]

[1]School of Foreign Languages, Nanfang College, Guangzhou, Guangdong Province, China
[a]dxr825@126.com, [b]16450818@qq.com, [c]18620635703@163.com, [d]kkkatrina.lin@qq.com
**Corresponding author: LIANG Ling*

Keywords: English Speaking Learning, AI Tools, iFLYTEK Spark

Abstract: With the rapid advancement of technology, the integration of artificial intelligence (AI) technology and language learning has emerged as a prevailing trend in foreign language learning. This study selected 20 English major senior students of Nanfang College ·Guangzhou, who had already passed the Test for English Majors Band 4 (TEM4) as the research subjects, experimentally demonstrated exhibited significant improvements in their comprehensive oral English skills and fluency by following a one-month intervention with AI-assisted English Speaking learning tool iFLYTEK Spark. This finding further underscores the substantial potential of AI technology in enhancing the effectiveness of foreign language teaching and learning outcomes.

1. Introduction

In language learning, oral proficiency is one of the most challenging aspects, primarily due to the scarcity of opportunities for authentic practice, especially in monolingual environments. However, with the widespread use of personal mobile devices, and the advent of AI-driven mobile applications featuring automatic speech recognition, natural language processing, and speech-to-text functionalities, innovative solutions have been developed to surmount these challenges.

Incorporating artificial intelligence (AI) into language learning, particularly to enhance English-speaking skills, has attracted considerable attention in recent years. This literature review aims to systematically explore the existing research on the impact of AI-based language learning platforms on students' English-speaking abilities.

2. Literature Review

2.1. Previous Research on The Potential and Challenges of AI-Assisted Interaction in Second Language Acquisition (SLA)

Interaction constitutes an indispensable component of second language acquisition (SLA). Michael Long's Interaction Hypothesis underscores the pivotal role of language interaction in this process, positing that learners enhance their linguistic development through negotiated meaning and adaptive linguistic adjustments during communicative interactions. This theoretical framework

serves as the foundational underpinning of the present study[1].

Recent advancements in AI have revolutionized English language education, particularly in enhancing speaking skills through interactive platforms like chatbots and conversational agents[2]. These systems create adaptive learning environments that cater to individual learner needs[3], fostering fluency and confidence while mitigating anxiety[4]. Moreover, effective pedagogical frameworks ensure alignment with curriculum objectives[5]. Indeed, such innovations highlight AI's transformative role in personalizing language education and bridging skill gaps.

However, AI chatbot principles, while useful for designing elementary English-speaking lessons [6], face implementation challenges. Specifically, successful deployment hinges on learners' technological receptiveness and cultural adaptability[7]. While positive AI-mediated experiences boost self-efficacy[8], persistent challenges include nuanced pronunciation error detection and data privacy concerns[9]. Ultimately, to balance AI and human roles in education, hybrid models combining AI-driven practice with instructor feedback are promoted to be prioritized[10].

2.2. Research Gap and Questions

In existing research, differences in the intervention effectiveness of AI tools across various linguistic elements (vocabulary, grammar, pronunciation, fluency) have not been systematically investigated, and the unique characteristics of human-AI interaction remain underexplored. This study aims to examine the following two questions empirically. The first is whether AI tools have a positive impact on learners' English-speaking abilities. The second is in which aspect—vocabulary, grammar, pronunciation, or fluency—are AI tools most effective in helping learners improve?

3. Key Assessment Criteria of English Speaking

IELTS, a globally recognized English proficiency test, evaluates oral skills across four equally weighted criteria: fluency/coherence, lexical resource, grammatical accuracy, and pronunciation. This balanced framework ensures comprehensive assessment of language abilities[11].

The AI tool iFLYTEK Spark mirrors IELTS standards, offering precise scoring and targeted feedback. Its four modules address each evaluation dimension: Fluency & Cohesion analyzes speech rhythm and logical flow; Lexical Diversity checks vocabulary precision and richness; Grammatical Range & Accuracy identifies syntax errors; Pronunciation Accuracy Review diagnoses articulation issues. Through adaptive practice, the tool not only assesses but also guides learners toward holistic improvement, aligning with IELTS' rigorous standards to foster effective language development[12].

4. Empirical Research on AI-Assisted English Speaking Learning

This empirical study combines quantitative measures of linguistic performance with qualitative interviews on learner experiences, offering a holistic view of AI-enhanced language learning.

4.1. Test

This study evaluates the efficacy of the AI tool iFLYTEK Spark in improving English majors' speaking skills via pre/post-test design. Participants, selected from TEM4-certified students to ensure baseline proficiency and representativeness, completed an initial oral assessment. A one-month intervention using iFLYTEK Spark (3-5 weekly 30-minute sessions across four modules) preceded a post-test. Pre-test results showed consistent proficiency (IELTS 6-6.5), validating sample homogeneity. The structured protocol and TEM4-qualified sample enhance the findings'

generalizability, while mandatory consistent engagement ensures data reliability.

As Figure 1 shows, in the pre-test, a total of 20 participants were involved, among whom 4 achieved a score of 6.5, while the remaining participants all scored 6. According to the scoring criteria of the IELTS, all participants fell within the range of 7 to 6 points. Because all participants who had successfully passed the TEM-4 were competent in English, it can be concluded that the pre-test proficiency levels of the participants in this study were generally consistent.

Subject	Test	**Band**	FC	LR	GRA	PRON	Subject	Test	**Band**	FC	LR	GRA	PRON
1	Pre	**6.0**	6.0	6.0	5.5	6.5	11	Pre	**6.0**	6.5	6.0	5.5	6.0
	Post	**6.5**	6.5	6.0	6.0	6.5		Post	**6.0**	6.0	6.5	5.5	6.5
2	Pre	**6.0**	6.0	6.5	6.0	6.0	12	Pre	**6.0**	6.0	6.0	5.5	6.5
	Post	**6.0**	6.0	6.5	5.5	6.5		Post	**6.5**	6.5	6.5	5.5	6.5
3	Pre	**6.0**	6.0	6.0	6.0	6.5	13	Pre	**6.0**	6.0	6.0	6.0	6.5
	Post	**6.0**	6.0	6.5	5.5	6.5		Post	**6.5**	6.5	6.5	6.0	6.5
4	Pre	**6.0**	6.0	6.0	5.5	6.5	14	Pre	**6.5**	6.0	6.5	6.0	6.5
	Post	**6.0**	6.0	6.5	5.5	6.5		Post	**6.5**	6.5	6.5	5.5	6.5
5	Pre	**6.5**	6.0	6.5	6.0	6.5	15	Pre	**6.0**	6.0	6.0	6.0	6.5
	Post	**6.5**	6.0	6.5	6.5	6.0		Post	**6.0**	6.5	6.0	5.5	6.5
6	Pre	**6.0**	6.0	6.5	5.5	6.5	16	Pre	**6.0**	6.0	6.0	5.5	6.5
	Post	**6.5**	6.5	6.5	5.5	6.5		Post	**6.5**	6.5	6.5	5.5	6.5
7	Pre	**6.0**	6.5	6.0	5.5	6.5	17	Pre	**6.5**	6.5	6.0	6.0	6.5
	Post	**6.0**	6.5	6.0	5.5	6.5		Post	**6.5**	6.0	6.5	6.0	6.5
8	Pre	**6.0**	6.0	6.5	5.5	6.5	18	Pre	**6.0**	6.5	6.5	5.5	6.0
	Post	**6.0**	6.5	6.0	5.5	6.5		Post	**6.5**	6.5	6.5	5.5	6.5
9	Pre	**6.0**	6.0	6.0	5.5	6.0	19	Pre	**6.0**	6.5	6.0	5.5	6.5
	Post	**6.5**	6.5	6.5	5.5	6.5		Post	**6.0**	6.5	6.0	5.5	6.5
10	Pre	**6.5**	6.0	6.5	6.0	6.5	20	Pre	**6.0**	6.0	6.0	5.5	6.5
	Post	**6.0**	6.0	6.0	6.5	6.0		Post	**6.0**	6.5	6.0	5.5	6.5

Figure 1: Band Score Comparison Between Pre-test and Post-test for 20 Test Examinees.

Twelve examinees remain the same band score, seven examinees have improved, one examinee has decreased performance. Thirty-five percent of the subjects improved their total oral test score over one month.

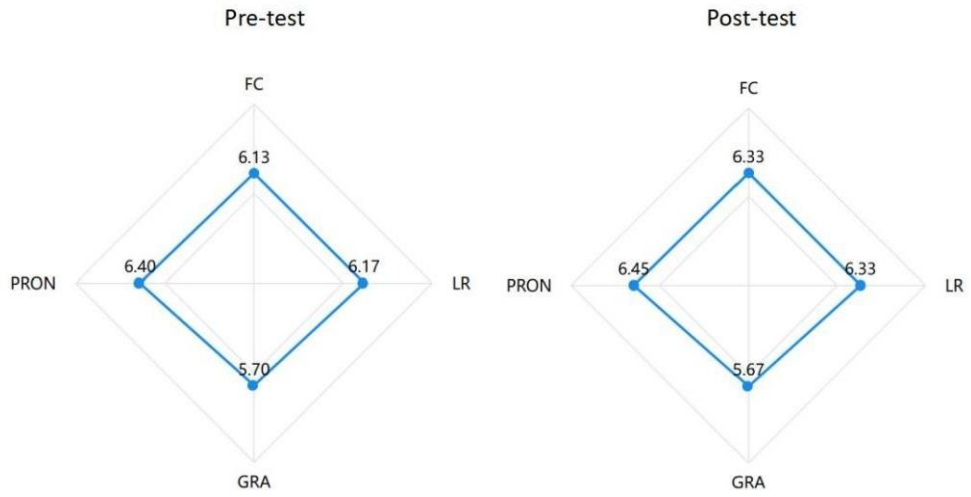

Figure 2: Pre-test vs. Post-test Radar Charts of Four Cognitive Dimensions.

The radar diagram in Figure 2 can visually show the score changes. In lexical resource, the

examinee group has increased by 0.15 band scores. In grammatical range, the examinee group has decreased by 0.03 band scores. In pronunciation, the examinee group has increased by 0.05 band scores. In fluency and coherence, the examinee group has increased by 0.2 band scores. While the examinee group has made progress in fluency and coherence, they have shown a decline in grammatical range.

Test	Objects	Mean	SD	SEM	t	p
Pre-Test	FC	6.13	0.22	-0.2	-2.629	0.017*
Post-Test		6.33	0.24			
Pre-Test	LR	6.17	0.24	-0.15	-2.042	0.055
Post-Test		6.33	0.24			
Pre-Test	GRA	5.7	0.25	0.03	0.37	0.716
Post-Test		5.67	0.34			
Pre-Test	PRON	6.4	0.21	-0.05	-0.809	0.428
Post-Test		6.45	0.15			
Pre-Test	Band Score	6.1	0.21	-0.15	-2.349	0.030*
Post-Test		6.25	0.26			

* $p<0.05$ ** $p<0.01$

Figure 3: Paired t-test of Band Scores: Pre-test vs. Post-test Comparison.

The figure 3 above reveals that out of a total of five paired datasets, two demonstrate significant differences ($p<0.05$). Detailed analysis shows that there is statistical significance at the 0.05 level between the FC groups ($t=-2.629$, $p=0.017$) and between the Band Scores ($t=-2.349$, $p=0.030$), with p-values below 0.05 in both cases. This suggests that the use of AI tools for oral English can enhance band scores, particularly showing notable improvements in fluency and coherence.

The fraction of significant differences has been presented in figure 4.

Pre-test - Post-test	Mean Deviation	95% CI of the difference	df	sig.(2-tailed)	Cohen's d Effect Size
FC - FC	-0.2	-0.359 ~ -0.041	19	0.34	0.588
LR - LR	-0.15	-0.304 ~ 0.004	19	0.328	0.457
GRA - GRA	0.03	-0.117 ~ 0.167	19	0.302	0.083
PRON - PRON	-0.05	-0.179 ~ 0.079	19	0.276	0.181
Band Score - Band Score	-0.15	-0.284 ~ -0.016	19	0.286	0.525

Figure 4: In-depth Analysis of Band Score's Effect Size Indicator.

Cohen's d value represents the effect size, with larger values indicating greater magnitudes of difference. The thresholds for distinguishing small, medium, and large effect sizes are 0.20, 0.50, and 0.80, respectively. An effect size of $0.2 <$ Cohen's d ≤ 0.8 signifies a medium magnitude of difference. Both FC (0.588) and Band Score (0.525) exhibit medium effect size improvements when compared between regions. For LR, the Cohen's d value is 0.457, which, according to Cohen's empirical rule, is slightly below 0.5, suggesting that there is a certain difference between the two groups, albeit not statistically significant. However, with a Cohen's d value of 0.457, which approximates the medium effect size threshold (0.5), the existence of a medium-level difference may still hold practical significance in educational practice.

The experimental results show that overall AI tools can play a positive role in the improvement of oral English, specifically in fluency and coherence, but the functionality in lexical resource aspect remains to be verified.

4.2. Interview

To examine cognitive variations and learning strategy differences across students with diverse academic performance levels, this study employed a purposive sampling approach to select three representative participants for semi-structured interviews. Specifically, the sample included one top-performing student, one average-achieving student, and one struggling student. Three respondents were selected based on performance trends: (1) Examinee 9, the sole participant showing across-the-board improvement; (2) Examinee 10, the only case of score decline among 20 participants requiring analysis; and (3) Examinee 20, representing average performance with fluency gains offset by scoring conventions. All emphasized spoken English's centrality in language learning, with two highlighting fluency's critical role. They advocated hybrid AI-human oral teaching approaches, citing complementary strengths/weaknesses, and the fact of though standardized monitoring caused perceived stress. Recommendations included expanding AI grammar content, dynamic avatar customization, real-time collaboration features, and improved speech recognition. Overall, respondents recognized the value of monitored spoken practice and AI's transformative potential when continuously optimized for learner needs.

5. Conclusions

5.1. Research Review

AI tools simulate dialogues to provide real-time feedback, correct errors, and continue topics, creating an interactive negotiation environment. Test tasks with interactive elements, combined with pre- and post-test designs, can verify the effectiveness of AI interaction in promoting learning abilities. Interviews reveal varying priorities among learners of different proficiency levels regarding their needs for explicit feedback versus contextual adaptation capabilities (meaning-focused negotiation).

5.2. Key Findings

In this experiment on AI-assisted English-speaking learning, students' oral proficiency scores and fluency demonstrated significant improvements. This not only attests to the positive role of AI tools in oral language learning but also supports the practicality and effectiveness of their actual application. Furthermore, respondents acknowledged the importance of English speaking skills and expressed recognition and anticipation towards personalized interaction and immersive experiences. They also believed that AI tools have the potential to assist in English-speaking learning, albeit continuous improvement and refinement are necessary to meet learners’ needs. In sum, the effectiveness of AI in enhancing students' oral language abilities especially fluency has been confirmed. Subjects who underwent AI-tool-assisted learning exhibited positive feedback, which offers a certain reference value for the construction of future AI-assisted English-speaking learning models.

5.3. Limitation of the Study

This study was constrained by the primary objective of ensuring homogeneous English proficiency among participants, resulting in a small sample group, which may compromise the generalizability of its findings. Additionally, time limitations of the graduating participants precluded thorough investigation into the long-term efficacy of AI tools. Future research should prioritize expanding participant diversity across demographics, age groups, and educational levels

while extending experimental durations to months or years. Such enhancements would enable a comprehensive evaluation of AI's pedagogical value, elucidation of its mechanistic impacts on oral English acquisition, and elevation of the study's scientific rigor and practical relevance.

Acknowledgement

This research was funded by the 2024 Undergraduate Innovation and Entrepreneurship Training Program of Nanfang College ·Guangzhou.

(1) AI-Assisted Language Learning: Exploration of the Innovative Path and Effectiveness Evaluation of iFLYTEK Spark (Project Number: X202412619028; Project Leaders: Dong Xingrui, Su Mingqi; Supervisor: Liang Ling).

(2) Voice of the Motherland: NLP-Based Social Analysis of Patriotic Sentiments and Dictionary Construction (Project Number: S202412619004; Project Leaders: Cai Kaiying, Pan Zhiwei; Supervisor: Liang Ling).

(3) Knowledge Graph Completion Based on Numerical Reasoning - Applications in Social Media (Project Number: X202412619046; Project Leader: Shi Xiaolin; Supervisor: Liang Ling).

References

[1] Long, M. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. Applied Linguistics, 4 (4), 126-141.

[2] Cahyono, A. E.,& Rosita, R. (2023). The Impact Of Using Ai-Based Language Learning Platforms On English Speaking Skills Of College Students. Transformational Language Literature And Technology Overview In Learning (TRANSTOOL), 2 (2), 1–8. https://doi.org/10.55047/transtool.v2i2.1352

[3] Yang, H., Kim, H., Lee, J. H.,& Shin, D. (2022). Implementation of an AI chatbot as an English conversation partner in EFL speaking classes. ReCALL, 34 (3), 327–343. https://doi.org/10.1017/s0958344022000039

[4] Makhlouf, M. K. I. (2021). The Effect of Artificial Intelligence-Based Application on Saudi Preparatory -Year Students' EFL Speaking Skills at Albaha University. International Journal of English Language Education, 9 (2), 36.

[5] Kim, H., Kim, N. Y.,& Cha, Y. (2021). Is it beneficial to use AI chatbots to improve learners' speaking performance? The Journal of AsiaTEFL, 18 (1), 161–178. https://doi.org/10.18823/asiatefl.2021.18.1.10.161

[6] Han, J.,& Lee, D. (2024). Research on the development of principles for designing elementary English speaking lessons using artificial intelligence chatbots. Humanities and Social Sciences Communications, 11 (1).

[7] Huang, F.,& Zou, B. (2024). English speaking with artificial intelligence (AI): The roles of enjoyment, willingness to communicate with AI, and innovativeness. Computers in Human Behavior, 159, 108355. https://doi.org/10.1016/j.chb.2024.108355

[8] Ericsson, E.,& Johansson, S. (2023). English speaking practice with conversational AI: Lower secondary students' educational experiences over time. Computers and Education Artificial Intelligence, 5, 100164.

[9] Rusmiyanto, R., Huriati, N., Fitriani, N., Tyas, N. K., Rofi'i, A.,& Sari, M. N. (2023b). The role of artificial intelligence (AI) in developing English language learner's communication skills. Journal on Education, 6 (1), 750–757. https://doi.org/10.31004/joe.v6i1.2990

[10] Nhan, L. K. (2024). Exploring students' perceptions of AI-Powered assistants in enhancing English speaking proficiency. International Journal of Innovative Science and Research Technology (IJISRT), 1299–1307. https://doi.org/10.38124/ijisrt/ijisrt24sep792

[11] IELTS. (n.d.). Understanding IELTS scoring. IELTS for Organisations. https://ielts.org/organisations/ielts-for-organisations/understanding-ielts-scoring

[12] iFLYTEK Spark (Version 2.7.0) [Mobile application]. (2025). Spark Education. https://apps.apple.com/cn/app/%E6%98%9F%E7%81%AB%E8%AF%AD%E4%BC%B4/id6449857311?l=en-GB