

The Impact of Artificial Intelligence on Improving Oral English Proficiency of Undergraduates

-- A Case Study of Doubao

Wenjia Zheng *

School of Foreign Languages, Anhui Agricultural University, Hefei, Anhui, China

* Corresponding author Email: cola20665098@outlook.com

Abstract: The promotion of AI technology to English learning has attracted people's attention. This paper takes Doubao as an example to analyze how college students can improve their spoken English proficiency with the help of AI technology, and to identify the influencing factors. The research method adopts questionnaire survey and literature analysis, focusing on the perception of undergraduate students in China on AI-assisted spoken English learning. Most students said that their spoken English was significantly improved and that Doubao was effective. A positive correlation was found between perceived gains, self-efficacy, AI practice preference and other variables. The study deepens the understanding of AI language learning theory, and provides reasonable suggestions for the development of digital oral teaching.

Keywords: AI-Empowered English-Speaking Practice; English-Speaking; Enjoyable Foreign Language Speaking; College Students; Speaking Proficiency Improvement.

1. Introduction

1.1. Literature Review

1.1.1. Current Status of College Students' Oral English Proficiency

In the context of global economic integration, spoken English is an important content in the cultivation of college students, which can enhance their overall quality. In the 2020 edition of the Guide to College English Teaching, it is clearly pointed out that the standards set for college students in the English Language Ability Standards should be strictly followed, and college students' oral English proficiency should be improved, so that they can master more basic knowledge, learn logical thinking, and have strong logical expression ability. In the specific teaching, college students should be guided to analyze the social hot topics, express their opinions to others, and give reasonable answers [1]. Despite extensive scholarly research, Chinese college students still struggle with limited vocabulary, poor grammar, unclear logic, cultural gaps, and "speech anxiety" [2]. In the process of traditional classroom teaching, teachers usually emphasize knowledge transmission and do not pay attention to personalized feedback or language environment construction, which can not meet the communication needs of college students [3]. Interactive technology tools can play an important role in the improvement of students' oral English proficiency.

What needs to be paid attention to is that spoken English research focuses on primary and secondary schools, and the number of keywords at the university level is relatively low, which means that the work is not systematic enough, and the research on spoken English in colleges and universities is relatively single.

1.1.2. Enhancing English-Speaking proficiency Through Artificial Intelligence

At present, LLM has attracted the attention of scholars, and its influence is increasing. CNKI data show that the

integration of this model and education is getting closer. This model can support a series of conversational tasks and has outstanding performance in the process of speech generation. Generative AI can play an important role in oral teaching, but it also must face challenges. Government policy makers should pay attention to the role of AI in education.

In March 2017, China's State Council released the Next Generation Artificial Intelligence Development Plan. The document explicitly required intelligent technology to accelerate reforms in talent cultivation and teaching methods [4]. The Ministry of Education's 2022 work agenda called for a digital education strategy that would drive digital and intelligent transformation. Digital speaking training relies on multimedia, smart assessment, and virtual scenarios. This approach fosters divergent, adaptive, and multidirectional thinking [5]. International empirical studies on AI in second language acquisition cover multiple skill areas. A review by Wenting Bao and colleagues examined 281 studies published between 2015 and 2024. Their analysis revealed that writing and speaking together accounted for roughly 60% of the research [6]. In the process of writing, we should use artificial intelligence to give feedback and automate evaluation. In the process of speaking, we should make full use of intelligent dialogue and speech recognition technology. In the study of speaking, we should focus on solving the problem of pronunciation, reducing the sense of anxiety, and enhancing the motivation. In the process of empirical research, we should pay attention to the improvement of students' speaking ability, make good use of multimodal interaction, and provide personalized feedback [7][8]. One investigation employed the app "English Liulishuo" [9]. It found that AI, by simulating authentic situations and delivering immediate feedback, boosted learners' confidence and oral expression. Another tool, FIF, provided human-machine feedback that was linked to higher scores, stronger motivation, and reduced anxiety [10]. It can be seen that artificial intelligence technology can provide good input and output opportunities for students, motivate students to learn English, and improve their

speaking ability efficiently.

Based on this, the specific hypotheses of this study are as follows:

H1: AI self-efficacy has a significant positive effect on the improvement of college students' English speaking skills.

H2: A preference for AI-assisted speaking practice has a significant positive effect on the improvement of college students' English speaking skills.

1.2. Significance of the Study

1.2.1. Theoretical Value

This paper focuses on the growth of spoken English, focuses on undergraduate students, and focuses on the role of AI in the growth of spoken English. The research results enrich the theoretical system of oral acquisition, and students should use this tool to enhance understanding and achieve negotiation of meaning. In the process of development, a new understanding of emotion has been formed, which can be combined with the traditional teaching method to form a human-computer collaborative teaching mode, and to test the applicability of the mode in the intelligent era.

1.2.2. Practical Value

Because of the lack of speaking practice opportunities, the time allocated to each course is relatively short, and Doubao is an efficient and convenient AI conversation tool, which can provide students with a broader practice platform. This study records the learning difficulties and shortcomings of college students in the process of using Doubao, and puts forward reasonable suggestions for the teaching of college English, so that more students can use AI tools rationally in the process of learning, have a deeper understanding of the tools, and make up for their shortcomings.

2. Theoretical Framework

2.1. Doubao Oral Practice Based on Krashen's Input Hypothesis and the Affective Filter Hypothesis

Doubao Oral Practice is founded on Krashen's Input Hypothesis ($i+1$: materials slightly above current proficiency, determined by language test scores) and the Affective Filter Hypothesis (low-anxiety, low-barrier environment). These core SLA theories continue to be widely debated [11]. Krashen pointed out that when learners are in a good emotional state, they can lower the "filter", obtain more input, and absorb the acquired knowledge, which is the opposite of the above situation. In his view, the output hypothesis does not pay too much attention to the input hypothesis, and the input hypothesis only focuses on writing and listening, which can not produce "mute English" and lead to oral difficulties.

From the current development situation, the "silent English" problem will be solved with the help of Doubao. First, at the input level, learners can use Doubao to simulate English exams, evaluate their English proficiency through the platform, and obtain learning materials that are slightly higher than their level, which is enough to bring them a strong sense of comprehension. Second, the platform will not judge the user's behavior, and it will not bring psychological pressure to the user. It can create a good learning environment for users, so that users can learn English and constantly improve their performance.

2.2. Application of Swain's Output Hypothesis in AI-Enhanced English-Speaking Practice

Looking back on the learning process of Chinese students, it is not difficult to find that the students do not have enough speaking opportunities, the learning is too much paper, their thinking is Chinese, and they do not have a strong desire to speak, which is related to the teaching mode [12]. In this case, students should enhance their language practice, get positive feedback and speak actively.

Swain's Output Hypothesis, built on Krashen's Input, argues learners need output skills like logical reasoning for fluency, beyond just receiving input [13]. Input stresses comprehension; output adds further processing. In 1995, Swain proposed three functions: noticing, hypothesis testing, and metalinguistic reflection [14].

First, students can use commands to get a good expression, and they can compare with Doubao's examples after answering, so as to relieve their anxiety and achieve self-practice. The hypothesis testing method is adopted. Doubao gives feedback and exchanges with students in a friendly way, arranges a variety of situations for students to learn, such as hotel check-in and so on, so as to make students' language more refined. In terms of reflection, Doubao can analyze the language, students can find more problems in the transcript, and can master the grammatical rules.

3. Research Methods

3.1. Research Design

In the survey, questionnaires were distributed through WeChat, Wenjuanxing and other platforms, all of which were strictly confidential, and the data were only used for academic research purposes. There are no standard answers, which can encourage users to express their views honestly and objectively. This survey collected information on the use frequency, age, self-efficacy and so on, and analyzed the relationship between the use frequency of Doubao and the perceived improvement, in order to find out the psychological and behavioral path of AI-assisted spoken English learning.

This study did not control the laboratory environment, but used questionnaires to collect data in the real environment, and asked the participants to evaluate their pronunciation, fluency, confidence and so on in the process of using Doubao, so as to obtain more real data, avoid the shortcomings of experimental research, and objectively show the impact of this tool on the learning of college students.

3.2. Scale Development

his questionnaire is formulated according to the mature scale, and according to the actual situation of the study, the scale is adjusted, which has high reliability and validity. The questionnaire includes three dimensions: how often users use Doubao, self-efficacy of users, and whether they like AI technology. The Likert five-point scale is used in all dimensions of the questionnaire.

In this study, four dimensions are selected as the core content, and the instruments are used to measure them, which are taken from the research results of predecessors. Specifically, Doubao's frequency of use is measured. Garland et al. [15] [16]. With items focusing on specific daily scenarios in which students employ AI tools. Response options range from "Rarely use" to "Use for more than 1 hour daily," with several intermediate levels to capture different

intensities of AI engagement. The AI Self-Efficacy dimension employed the “AI Self-Efficacy Scale (AISE)” developed by Wang and Chuang [17], which includes seven items, and students can express their mastery of AI technology and confidence in their own abilities through this scale. For example, "AI technology can provide strong support for my learning" and "the interaction with AI technology is similar to that with human beings."The scale for measuring students' willingness to use AI is compiled by Li et al [16], focusing on students' acceptance of AI and their willingness to recommend it actively. Lastly, perceived improvement in speaking ability through AI assistance refers to students' subjective judgments—based on their actual Doubao usage—about changes in pronunciation accuracy, fluency, confidence, and overall practical speaking proficiency..

4. Data Analysis

4.1. Descriptive Statistical Analysis

4.1.1. Distribution of Basic Information on College Students

In this survey, questionnaires are used as a tool for online investigation, Wenjuanxing is the platform for creating and

distributing questionnaires, and WeChat and other online platforms are used to attract respondents. In this study, full-time undergraduate students in Chinese universities are the survey targets, and random sampling method is adopted. A total of 224 questionnaires were distributed, but only 145 were effectively recovered. The specific reasons for elimination are as follows: (1) The response time is too short, less than 60 seconds, 39 people in total; (2) The response time is too long, more than 300 seconds, 2 people in total; (3) The same answer was given to more than five questions in a row, 15 people in total; (4) The answers to the questions are not consistent, 13 people in total, the most prominent of which is that the answers to the questions are in conflict, such as high usage tendency and high worry at the same time; (5) The answer is not in line with the logical relationship, 10 people in total. From the gender of the respondents, the number of female students is relatively large, reaching 66.2%, while male students only account for 33.8%.The number of undergraduates in the survey results accounts for 84.8%, while graduate students account for only 15.2%.The samples in the survey are evenly distributed by grade, and the applicability of the survey results is relatively broad.

Table 1. Distribution of basic information on college students

Name	Option	Frequency	Percentage (%)	Cumulative Percentage (%)
Gender	Male	49	33.8	33.8
	Female	96	66.2	100
Grade	Freshman	32	22.1	22.1
	Sophomore	23	15.9	37.9
	Junior	34	23.4	61.4
	Senior	34	23.4	84.8
	Graduate Student and Above	22	15.2	100

4.1.2. Frequency of Doubao Use

According to the analysis of the frequency of Doubao use, most college students use it for 1 hour or less a day (42.1%), 30.3% for more than 1 hour, and 20.7% for 1-3 days. The inverse scoring method (A=8, B=1) is adopted, and the score is 6.8828 (SD=1.12733).It can be seen that most students use Doubao at least once every 1-3 days, which is more representative and has a wide range of users, but only 1% of

them are rarely used, which means that this part of users can only have limited exposure. Aldamen et al. [18] believe that some students are unwilling to use Doubao in the process of learning, which means that they need to have a strong understanding of the value of Doubao in learning, and guide them to use it independently to learn AI technology. The above data analysis results show that most students use Doubao for 1-3 days, but there are a few students who do not use it, which needs to be trained.

Table 2. Frequency of Doubao use

Name	Frequency	Percentage	Mean (M)	Standard Deviation (SD)
A. Average daily usage of 1 hour or more	44	30.3	6.8828	1.12733
B. Average daily usage of 0–1 hour	61	42.1		
C. Average usage once every 1–3 days	30	20.7		
D. Once a week on average	4	2.6		
E. Once every 2–3 weeks on average	3	2.8		
F. Once a month on average	2	2.1		
G. Once every 2–3 months on average	0	1.4		
H. Rarely used	1	0.7		

4.2. Reliability and Validity

4.2.1. Reliability Analysis

Cronbach's α is used to evaluate reliability. The reliability coefficient of the questionnaire is 0.838, and the consistency

among items is relatively high.In terms of single dimension, the reliability coefficient of AI in improving spoken English is 0.767, which is relatively reasonable. The other two dimensions are between 0.6 and 0.7. Both Hair et al. (2010) and Numally (1978) pointed out that this is a reasonable value.

Combined with the overall $\alpha = 0.838$, it can be seen that these two dimensions can still be used in the next analysis. The use frequency of Doubao is a single scale, and the value of

Cronbach's α should be at least 2 items, so it is analyzed in the overall reliability.

Table 3. Reliability analysis

Dimension	Number of Items	Sample Size	Cronbach's α Coefficient
AI Self-Efficacy	7	145	0.656
Tendency to Use AI-Empowered English Speaking Practice	5	145	0.681
Perceived Improvement in Speaking Ability Through AI Assistance	5	145	0.767
Total Reliability	18	145	0.840

4.2.2. Validity Assessment

In this study, Bartlett's sphericity test and KMO sampling adequacy test are used to analyze whether the data is suitable for factor analysis. The KMO value is 0.853, which is a good value. If the KMO value is greater than 0.7, it means that there are many common factors among variables. The approximate

chi-square value is 678.864, the degree of freedom is 136, and the significance level is less than 0.001 (the ideal value is less than 0.05), which means that the null hypothesis is rejected and the correlation between variables is strong. Using these two indicators, it is known that the data collected in this study meets the requirements of factor analysis.

Table 4. Validity assessment

KMO Sampling Adequacy Statistic	0.853	
Bartlett's Sphericity Test	Approximate Chi-Square	678.864
	Degrees of Freedom	136
	Significance	<.001

4.3. Correlation Analysis

Correlation analysis is a statistical method used to examine relationships between variables to understand association

nature. This paper uses SPSS 32 to verify the correlations among the variables, as well as the strength and direction of these relationships. The results of the correlation test are shown in table 5 below.

Table 5. Correlation analysis

Correlations				
		AI Self-Efficacy	Tendency to Use AI-Empowered English Speaking Practice	Perceived Improvement in Speaking Ability Through AI Assistance
AI Self-Efficacy	Pearson	1	.450***	.552***
	Significance (two-tailed)		<.001	<.001
	Number of cases	145	145	145
Tendency to Use AI-Empowered English Speaking Practice	pearson correlation	.450***	1	.606***
	Significance (two-tailed)	<.001		<.001
	Number of cases	145	145	145
Perceived Improvement in Speaking Ability Through AI Assistance	pearson correlation	.552***	.606***	1
	Significance (two-tailed)	<.001	<.001	
	Number of cases	145	145	145
*** The correlation is significant at the 0.001 level (two-tailed).				

Correlation results show significant relationships among the three core variables. Specifically, perceived improvement in speaking ability through AI assistance correlates positively with AI self-efficacy, and AI self-efficacy also relates positively to the tendency to use AI-enabled speaking practice. Among these, the Pearson correlation coefficient between perceived improvement and usage tendency is the largest, at 0.606 ($p < 0.001$), suggesting a strong positive link. In other

words, students who more frequently adopt AI for oral practice tend to report greater perceived gains. The coefficient between perceived improvement and self-efficacy is 0.552, also significant at $p < 0.001$, reflecting a moderate positive correlation. This implies that strengthening learners' confidence in using AI might serve as a key leverage point for enhancing speaking outcomes. The correlation between self-efficacy and usage tendency stands at 0.450, which is

relatively weaker yet still positive, indicating that the two factors remain mutually reinforcing.

4.4. Regression Analysis

This study examines the interdependence among AI self-efficacy and the tendency to use AI-enabled English speaking practice, with the latter serving as the dependent variable and the former two as independent variables. As shown in the

table 6, AI self-efficacy and the tendency to use AI-empowered English speaking practice explain approximately 45.7% of the variance in “perceived improvement.” The F-value is 61.653, and the P-value is less than 0.001, indicating that the overall regression model is significant. The D-W statistic is 2.064, which is close to 2, suggesting no autocorrelation and indicating that the model is well-specified.

Table 6. Regression analysis

Model Summary ^b						
R	R-squared	Adjusted R-squared	Standard error of the estimate	Durbin-Watson	F	Significance
682 ^a	0.465	0.457	0.38402	2.064	61.653	<.001

It can be seen from the table that the p value of self-efficacy is less than 0.001, and the hypothesis H1 is accepted; The p-value of usage tendency is less than 0.001, and the hypothesis H2 is accepted. It can be seen that the two variables have a positive impact on the improvement of speaking ability. In the

process of testing whether the variables are collinear, the VIF of the two variables is 1.254, which is far less than 5, and the model will not produce collinearity problems, and the research results are reliable.

Table 7. Regression model for perceived improvement

Regression Model for Perceived Improvement							
	Unstandardized Coefficient		Standardized Coefficient	t	Significance	Multicollinearity Statistics	
	B	Standard Error	Beta			Tolerance	VIF
Self-Efficacy	0.403	0.079	0.350	5.091	<.001	0.797	1.254
Usage Tendency	0.424	0.065	0.448	6.519	<.001	0.797	1.254

5. Discussion

In this study, the impact of Doubao on the oral English level of Chinese college students is analyzed to obtain the following conclusions.

(1) In the study, the three variables are positively related, that is, there is a positive correlation between AI self-efficacy and usage tendency ($r = 0.450$, $p < 0.001$), between self-efficacy and perceived improvement ($r = 0.552$, $p < 0.001$), and between usage tendency and perceived improvement ($r = 0.606$, $p < 0.001$). The three variables are related to each other, and there is a clear path from "perceived gains → attitude → confidence".

(2) Doubao provides users with a self-directed oral practice environment, which is different from the traditional group Q&A or large-class teaching mode. It can help users relieve their nervousness and actively participate in oral communication.

(3) Some students do not believe in the feedback given by artificial intelligence [19]. Although artificial intelligence technology is constantly updating, there are still some users who worry about pronunciation correction, information update and so on.

6. Conclusion and Outlook

6.1. Conclusion

In this research, the three variables of self-efficacy, willingness to use AI, and speaking ability have a significant relationship with each other, and the three variables are combined into a path diagram, which shows that the relationship between willingness and improvement is the most prominent, which means that the higher the willingness

to use AI, the more obvious the improvement.

The creation of this scenario can make students not feel anxious, reduce emotional filters, encourage students to self-correct, and put forward reasonable suggestions, which are consistent with the views of Krashen and Swain. In the process of using artificial intelligence, some students have doubts about the feedback results, believing that it can only play a supporting role in learning, not the only learning method. This research clarifies the role of artificial intelligence in reducing students' speaking anxiety, and can provide reference for oral teaching.

6.2. Limitations of the Study

Because there is no objective measurement, only self-evaluation is carried out to understand the improvement. In this case, we should pursue ecological validity, and there is a certain deviation in memory and social preference. Biased, we should organize objective tests in the next research to ensure the accuracy of the conclusion.

6.3. Prospects

In the next stage of research, we should use more reasonable ways to evaluate the effectiveness of the interview, including pronunciation, fluency, grammar and other aspects, and use pre-test and post-test design to analyze the effectiveness of the use of AI tools. In addition, we need to develop a self-efficacy scale for AI, and be able to distinguish between learning efficacy and task efficacy. In the process of using AI tools, we should pay attention to the accuracy of feedback, adjust the difficulty dynamically according to the user's performance, use multimodal interaction to promote user immersion, and enhance user's sense of participation.

References

- [1] He, L. Z. (2020). New requirements for college English teaching in the new era: Basis and key points for revising the College English Teaching Guide. *Foreign Language World*, (4), 13–18. (in Chinese)
- [2] Tan, Q. Y., & Hu, M. T. (2022). A study on the effectiveness of college students' self-directed English speaking learning under the background of "Internet Plus": A case study of the School of Foreign Languages, Neijiang Normal University. *English Square*, (12), 54–56. (in Chinese)
- [3] Zhu, B. (2022, February 10). Innovative pathways for improving college students' English speaking competence. *Chinese Social Sciences Today*, p. 008. (in Chinese)
- [4] The State Council. (n.d.). New Generation Artificial Intelligence Development Plan. *Science & Technology Review*. (in Chinese)
- [5] Wu, W. (2024). A discussion on digital English speaking training and the cultivation of college students' innovative ability. *Journal of Inner Mongolia University of Finance and Economics*, 22(4), 55–59. (in Chinese)
- [6] Bao, W., Wang, T., Zhang, L., et al. (2025). A systematic review of AI in second language acquisition using the expanded SAMR model (2015-2024). *Discover Computing*, 28, 292. <https://doi.org/10.1007/s10044-025-01421-8>.
- [7] Huang, F., & Tang, Z. Q. (2026). Artificial intelligence empowering English speaking learning: A study on the relationship between perceived enjoyment and continuance intention. *Journal of Xi'an International Studies University*, 34(1), 53–59. (in Chinese)
- [8] Zou, B., & Wang, C. H. (2026). A study on the pathways through which generative artificial intelligence promotes personalized English speaking learning among college students. *Foreign Language Research*, (1), 61–69. (in Chinese)
- [9] Qin, A. H. (2025). Artificial intelligence and the improvement of business English speaking competence: A case study based on the Liulishuo App. *University of International Business and Economics*. (in Chinese)
- [10] Zhou, H. M. (2024). A study on the impact of human-machine collaborative feedback strategies based on problem diagnosis on the cultivation of college students' English oral expression ability. *Wenzhou University*. (in Chinese)
- [11] Huang, J., & Yang, Y. (2026). A re-exploration of second language acquisition strategies under the theoretical background of Krashen's input hypothesis. *University*, (5), 176–179. (in Chinese)
- [12] Hao, J. L. (2025). A study on college English speaking teaching under the guidance of Krashen's affective filter hypothesis. *Overseas English*, (21), 131–133. (in Chinese)
- [13] Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In *Input in Second Language Acquisition*. Newbury House.
- [14] Shi, Y. (2022). Optimization strategies for assignment design in blended college English listening and speaking teaching under the output hypothesis. *Journal of Suzhou Education Institute*, 25(6), 46–50. (in Chinese)
- [15] Garland, K. J., & Noyes, J. M. (2004). Computer experience: A poor predictor of computer attitudes. *Computers in Human Behavior*, 20, 823–840. <https://doi.org/10.1016/j.chb.2003.11.002>.
- [16] Li, W. T., & Wang, Y. (2025). A study on contemporary college students' attitudes toward generative artificial intelligence-empowered English speaking practice. *Cross-Science Bulletin*, 9(3), 221–231. (in Chinese)
- [17] Wang, Y., & Chuang, Y. (2023). Artificial intelligence self-efficacy: Scale development and validation. *Education and Information Technologies*, 29, 4785–4808. <https://doi.org/10.1007/s10644-023-02247-1>.
- [18] Aldamen, H., Almashour, M., Deaibes, A. M., et al. (2025). Testing Krashen's input hypothesis with AI: A mixed-methods study on reading input and oral proficiency in EFL.
- [19] Liu, Z. W. (2025). Perceptions and evaluations of Doubao-assisted English speaking training among non-English-major college students using generative artificial intelligence tools. *Sichuan International Studies University*. (in Chinese)