ISSN:2308-5703
Impact factor:13.5 *International Multidisciplinary Online Journal* October, 2026

The Impact of AI Voice Software on Speaking Speed and Hesitation among EFL Students

Ibodullayeva Diyora Orifjon qizi
Student at the Samarkand State Institute of Foreign Languages
diyoraibodullayeva50@gmail.com
doi:**https://doi.org/10.5281/zenodo.23323466**

Abstract. Developing oral fluency remains a major challenge for many as a EFL due to limited classroom speaking time and fear of making mistakes. This six-week study examined whether combining two mobile AI applications – ELSA Speak for pronunciation drills and ChatGPT Voice Mode for conversational practice – could improve students' speaking speed and reduce hesitancy. Fifty B1-level undergraduate students in Uzbekistan were divided into an experimental group (n = 25), who completed 90 minutes of weekly AI practice alongside their regular classes, and a control group (n = 25), who received standard instruction only. Oral monologues recorded before and after the intervention showed that the experimental group increased their average speaking rate from 92.4 to 118.4 words per minute and cut hesitation pauses from 12.8 to 6.2 per 100 words. In contrast, the control group made much smaller gains. Survey and interview feedback revealed that practicing with AI reduced speaking anxiety by providing a judgment-free environment. These results suggest that AI voice tools work effectively as home practice aids to build confidence and fluency alongside traditional classroom teaching.

Keywords: EFL speaking fluency, ELSA Speak, ChatGPT Voice Mode, speech recognition, speaking anxiety, mobile language learning.

Introduction. Speaking fluency is one of the most challenging aspects for EFL learners [8; 30]. However, in many traditional classrooms, grammar and vocabulary are paid attention the most, improving speaking skills is secondary [6; 2682]. In classrooms, allocating speaking time for every student, giving real-time feedback immediately for their mistakes, and eliminating speaking anxiety are always difficult [5; 255-270]. Consequently, although many students are theoretically knowledgeable, they are faced with some problems like making pauses, fear and self-doubt [3; 1018-1027]. In recent years, the advent of Artificial Intelligence (AI)-based speech recognition and voice interaction technologies has brought about a fundamental shift in language teaching methodology [1; 180]. Apps, such as Elsa Speak, help assess intonation and speech speed by AI algorithmic analysis. In addition, it gives students clear and objective advices [2; 180]. Interactive voice models, such as ChatGPT's Voice mode, enable students to engage in natural and flexible real-time conversations. These digital technologies enable students speak fluently in their own speed and comfortable atmosphere.

Literature review. Existence literature review shows that AI-related researches are mainly about student's vocabulary, grammatical proficiency and accuracy. However, the empirical impact of the combined use of AI-based speech applications (such as ELSA Speak and ChatGPT Voice) on the specific speaking fluency of EFL learners – namely, words per minute (WPM), the frequency of inappropriate pauses, and overall speaking confidence – has not been sufficiently or systematically studied. Filling this research gap is of vital importance for improving modern digital language teaching methodology and integrating AI tools into the teaching process. This research aims to recognize AI speech, analyze the effect of speaking apps (ELSA speak and ChatGPT Voice) to students' fluency as well as study their thoughts about using these apps. This research relies on these two questions:

RQ1: To what extent does the integration of AI speech-recognition apps (ELSA Speak and ChatGPT Voice) improve EFL learners' speaking fluency in terms of Words Per Minute (WPM) and pause frequency?

RQ2: What are EFL learners' perceptions regarding the effectiveness, usability, and confidence-building aspects of AI voice tools in their speaking practice?

Methodology. Fifty undergraduate students majoring in English philology at a public university in Uzbekistan participated in this study. All participants were at the B1 intermediate level, which was verified using the Oxford Online Placement Test prior to the intervention. We split the students into two equal groups of 25: an experimental group that used AI tools alongside their regular classes, and a control group that followed the standard curriculum only. All students gave written consent before the project began, and their personal details were kept strictly confidential. This was a six-week study designed to test the practical impact of AI software on speaking fluency. Both groups attended the same two-hour weekly speaking class taught by the same instructor to keep classroom instruction consistent. Students in the experimental group completed three 30-minute self-study sessions per week using two mobile applications: ELSA Speak and ChatGPT Voice Mode. During each session, they spent 15 minutes on ELSA Speak practicing pronunciation and sentence stress through repetition drills. They spent the remaining 15 minutes having spontaneous conversations with ChatGPT Voice Mode on weekly class topics, such as education and technology. Students submitted weekly screen-time logs to confirm they completed the app practice.

The control group performed similar speaking activities through standard pair work in class, without any software-based practice. We evaluated oral fluency using a two-minute monologue task modeled on IELTS Speaking Part 2, administered before and after the six-week period. Each student received a topic card, had one minute to prepare, and spoke for two minutes in a quiet testing room while being recorded. The fluency was scored using two main measures:

1. Words Per Minute (WPM): The total number of words spoken divided by total speaking time in minutes.
2. Pause Frequency: The total number of hesitation sounds (such as "uh" and "um") and silent pauses longer than 0.5 seconds per 100 words.

Two experienced English teachers independently evaluated all audio recordings. The agreement between their scores was strong, with an intraclass correlation coefficient of 0.89. At the end of week six, the experimental group completed a 15-item survey about tool usability and speaking confidence. We also interviewed eight students from the experimental group for 15 minutes each to get detailed feedback about their experience using the apps. Quantitative data were analyzed using SPSS software. We used paired t-tests to measure progress within each group from pre-test to post-test, and independent t-tests to compare performance differences between the two groups. The significance level was set at $p < 0.05$. For the qualitative interview data, we transcribed the audio recordings and grouped the responses into main themes regarding student motivation, anxiety, and software feedback.

Results. At the start of the study, both groups were speaking at virtually the same pace. The experimental group averaged 92.4 words per minute, while the control group sat at 91.8. Hesitation patterns were also nearly identical, with both groups averaging around 13 pauses every 100 words. By the end of the six weeks, however, we saw a clear difference between the two groups. Students using the AI apps managed to increase their speaking speed to an average of 118.4 words per minute – an increase of roughly 26 words per minute. The control group also improved, reaching 98.2 words per minute, but their progress was much slower. The reduction in pauses was even more noticeable. The experimental group cut their hesitation rate by more than half, dropping from 12.8 down to 6.2 pauses per 100 words. The control group showed a slight decrease (from 12.9 to 10.4), but they still struggled with frequent hesitations and long silences during their monologues.

The survey and interview feedback helped explain why this happened. Students rated the software highly overall (4.38 out of 5), but the main theme in their comments was the lack of social pressure.

Many mentioned that speaking to an app removed the fear of making mistakes in front of classmates. As one student explained, "I didn't have to worry about looking foolish with ChatGPT, so I just kept talking without stopping to translate every sentence in my head." Others noted that ELSA Speak made them aware of how often they were pausing, which motivated them to speed up their delivery during practice drills.

Discussion. What these findings show is that mobile AI tools can solve a common problem in language teaching: getting students to practice speaking outside the classroom. In a standard university class with 20 or 25 students, no teacher can give each person enough individual speaking time. By using ELSA Speak and ChatGPT at home, our experimental group got an extra 90 minutes of dedicated speaking practice every week. Beyond just adding hours, the psychological side of using AI turned out to be critical. Anxiety is one of the biggest hurdles to fluency – students pause because they are double-checking their grammar before speaking. ChatGPT gave them a risk-free environment where they could focus on communicating ideas rather than getting every rule right. Meanwhile, ELSA provided the technical feedback on pronunciation and rhythm that ChatGPT lacks. Using both tools together balanced spontaneous conversation with targeted accuracy. That said, these tools shouldn't replace classroom teaching. AI cannot replicate authentic human interaction or group dynamics, so these apps work best as a bridge – building students' confidence and speed at home so they can participate more effectively during class activities. Since our study lasted only six weeks and focused on 50 students at one university, longer studies across different proficiency levels would help show whether these fluency gains hold up over time.

Conclusion. This study shows that integrating AI tools like ELSA Speak and ChatGPT Voice Mode into an EFL student's routine can noticeably improve speaking fluency and reduce hesitancy [4; 570-578]. By offering immediate feedback on pronunciation alongside a low-pressure environment for conversational practice, these applications help students bridge the gap between knowing English grammar and using it fluidly in real-time communication [7; 186]. Crucially, AI should not be viewed as a substitute for classroom teachers or real human interaction. Instead, these tools work best as a personal practice ground – a way for students to build confidence, expand their speaking time, and work through anxiety on their own schedule. As AI technology continues to evolve, incorporating such tools into language curricula offers a practical, scalable way to help learners become more continuous, confident speakers both inside and outside the classroom.

References:

1. Aryadoust, V., & Zakaria, A. (forthcoming). Formative assessment of speaking for young learners using Generative Artificial Intelligence. In C. C. M. Goh, R. L. Sabnani, & W. Renandya (Eds.), Second language oracy development in young learners. Springer Nature
2. Christine C. M. Goh and Vahid Aryadoust, Developing and assessing second language listening and speaking: Does AI make it better? Annual Review of Applied Linguistics (2025), 45, 179–199 doi:10.1017/S0267190525100111
3. Dheressa, M. A., Olana, T., & Bekele, E. (2024). The effects of task-based speaking instruction on students' speaking perception at Shambu Secondary School. Journal of Languages and Language Teaching, 12(2), 1018-1027. https://doi.org/10.36077/jllt.2024.1202.1018
4. Hadi, M. S., & Junor, R. S. (2022). Speaking to devices: Can we use Google assistant to foster students' speaking skills? Journal of Languages and Language Teaching, 10(4), 570–578. https://doi.org/10.33394/jollt. v10i4.5808
5. Sada, E. Y., Bulbula, D. A., & Bulti, T. A. (2023). EFL teachers' and students' attitudes and practices regarding oral communication in English classes: Ethiopian high school context. International Journal of Language Education, 7(2), 255-270. https://doi.org/10.26855/ijle.2023.02.022

6. Tanveer, M. (2007). Investigation of the factors that cause language anxiety for ESL/EFL learners in learning speaking skills and the influence it casts on communication in the target language, 2682. (Unpublished master dissertation). University of Glasgow, Scotland.
7. Xu, S., Qin, L., Chen, T., Zha, Z., Qiu, B., & Wang, W. (2024). Large language model-based situational dialogues for second language learning. arXiv:2403.20005. https://arxiv.org/abs/2403.20005
8. Win Listyaningrum Arifin, REGISTER JOURNAL, Language & Language Teaching Journals, Vol. 10, No. 1, 2017, pp.29-47 http://dx.doi.org/10.18326/rgt.v10i1.29-47