



“HEY SIRI, ALEXA, OK GOOGLE—HELP ME LEARN ENGLISH!”: VOICE ASSISTANTS IN EFL LEARNING

Sabrina^{1*}

¹English Education Department, Faculty of Teacher Training and Education, Universitas Serambi Mekkah, Indonesia

*sabrina@serambimekkah.ac.id

ABSTRACT

This study investigated the impact of voice assistant technologies (VATs), specifically Google Assistant, Amazon Alexa, and Apple Siri, on English as a Foreign Language (EFL) speaking practice among undergraduate students at Universitas Serambi Mekkah. Using a mixed-methods approach, the study involved 40 EFL students over a four-week intervention. Quantitative data were obtained through pre- and post-tests measuring speaking fluency, pronunciation accuracy, and lexical range, while qualitative data were gathered via surveys and semi-structured interviews. Results showed significant improvements across all three oral language domains, with learners exhibiting greater fluency, enhanced pronunciation, and a wider lexical repertoire after interacting with VATs. Learners also have positive perceptions towards the use of VATs in their English-speaking practice, particularly in terms of reducing their speaking anxiety and increasing their learning motivation and confidence. Despite acknowledging that all VATs were easier to use for speaking practice than traditional pedagogical approaches, they expressed a clear preference for Google Assistant over the other options due to its accessibility on Android devices, user-friendly interface, capability to recognize non-native accents more accurately, capability to process a wide range of spoken prompts, multimodal input flexibility in circumstances where voice recognition struggled, and ease integration into daily routine. The findings underscore the effectiveness of VATs in supporting speaking practice, particularly in low-resource contexts, and offer valuable insights into how digital voice technologies can be integrated into EFL instruction to promote learner independence and engagement.

Keywords: *voice assistant technologies (VATs), EFL speaking, Google Assistants, Alexa, Siri, fluency, pronunciation, lexical range, learner motivation*

1. INTRODUCTION

The proliferation of artificial intelligence (AI) in education has brought substantial changes to how languages are taught and practiced. Among the most accessible AI innovations are voice assistant technologies (VATs), such as Siri, Alexa, and Google Assistant. These technologies have become a part of daily life and offer new affordances for education by simulating natural conversations and providing interactive feedback. In the EFL context, speaking is often cited as the most difficult skill to acquire, largely due to the scarcity of authentic speaking environments and the anxiety learners experience in real-time interactions. VATs present a novel avenue for overcoming these obstacles by enabling learners to engage in oral practice without the fear of judgment or error correction from human interlocutors.

A growing body of literature supports the integration of VATs in language learning. Some studies have reported that AI-driven voice assistants offered valuable opportunities for listening and repeating new words which eventually improved their pronunciation (Abimanto & Sumarsono, 2024; Alharthi, 2024). Similarly, Rahman & Tomy (2023) reported improvements in speaking fluency and listening comprehension among Indian university students who used Google Assistant in L2 classrooms. Another study by Tregubov (2021) and Alrajhi (2024) emphasized the role of AI-powered VATs in improving students' lexicons, including those used in academic contexts. However, the majority of such research has taken place in technologically advanced contexts, leaving a gap in understanding how these tools function in low-resource settings.

In Indonesia, and particularly in Aceh, access to native speakers and advanced language labs remains limited. At Universitas Serambi Mekkah, many students come from rural backgrounds and face constraints in accessing immersive English-speaking environments. Nonetheless, the widespread ownership of smartphones among students presents an opportunity to utilize built-in VATs like Siri, Google Assistant, or third-party



applications like Alexa. While these tools are not specifically designed for educational purposes, their conversational capabilities can be leveraged to provide consistent speaking practice.

Despite their potential, there is scant research evaluating the pedagogical outcomes of using different forms of VATs—Siri, Alexa, and Google Assistant—in EFL speaking instruction within developing country contexts. Previous studies mostly focused on the comparison of those forms of VAT performance in terms of customer satisfaction (Brill et al., 2019), capability of answering health questions (Alagha & Helbing, 2019), and security (Aakanksha et al. 2025). As such, this study aims to fill this gap by exploring the impact of voice assistants on the speaking skills of undergraduate students at Universitas Serambi Mekkah.

This research was guided by the following questions:

1. To what extent does the use of voice assistant technologies (Google Assistant, Siri, and Alexa) enhance EFL students' speaking fluency, pronunciation accuracy, and lexical development?
2. How do EFL students at Universitas Serambi Mekkah perceive the use of voice assistant technologies in their English-speaking practice, particularly in relation to speaking anxiety, learning motivation, and perceived ease of use?
3. Which voice assistant technology (Google Assistant, Siri, or Alexa) do students prefer for speaking practice, and what factors influence their preferences?

By addressing these questions, this study aims to fill a critical gap in current EFL research and provide actionable insights for integrating digital voice technologies into speaking instruction in similar developing educational contexts.

2. METHODS

This study employed a mixed-methods research design to explore the effects and perceptions of using voice assistant technologies in EFL speaking practice. Both qualitative and quantitative data were gathered to ensure a holistic understanding of students' experiences.

The participants of this study were 40 undergraduate students enrolled in the English Education program at Universitas Serambi Mekkah. Prior to the intervention, participants completed a demographic survey to confirm their technological accessibility and frequency of VAT usage.

The data collection instruments included pre- and post-speaking fluency assessments, a structured questionnaire, and semi-structured interviews. The fluency tests, administered before and after the intervention, evaluated students' speaking rate, pronunciation accuracy, and vocabulary. The questionnaire, based on a five-point Likert scale, assessed learner motivation, speaking confidence, and perception of VAT usefulness in improving the aspects of their speaking skills. Semi-structured interviews were conducted with 20 randomly selected students from the larger sample, focusing on their subjective experiences using VATs in speaking tasks as well as their preferred VATs.

Participants engaged in a structured four-week intervention designed to incorporate regular speaking practice using voice assistant technologies (VATs), specifically Google Assistant, Amazon Alexa, and Apple Siri. Each week, students were assigned a set of speaking tasks that required direct interaction with their chosen VAT. These tasks were intentionally varied to reflect real-world communication scenarios and to promote different aspects of speaking skills.

The tasks included three main types of interaction:

1. Reading Aloud, in which students were instructed to read prepared sentences or short paragraphs aloud to their VAT. This task aimed to enhance pronunciation and rhythm by encouraging learners to articulate clearly enough for the VAT to recognize their input accurately.
2. Open-Ended Questioning, in which students engaged the VATs by asking and answering open-ended questions (e.g., *"What is climate change?"* or *"Tell me about famous Indonesian dishes"*). These



exchanges required learners to generate spontaneous responses, fostering fluency, lexical recall, and the ability to structure ideas verbally.

3. Dialogue Simulations, in which the students practiced short simulated dialogues with the VATs, such as making a restaurant reservation, giving directions, or asking for advice. These situational dialogues mimicked everyday conversations, allowing learners to apply language in meaningful contexts.

To monitor progress and encourage reflection, the students were asked to record their spoken interactions each week using mobile audio recording apps. These recordings were submitted to the instructor for tracking purposes and were used to assess changes in fluency, pronunciation, and lexical diversity. Each week's tasks increased in complexity, from simple questions and responses in Week 1 to more extended, multi-turn interactions by Week 4, to gradually challenge students and support continuous language development.

Quantitative data from the fluency assessments and questionnaires were analyzed using SPSS. Paired-sample t-tests were conducted to measure differences in pre- and post-intervention speaking scores and motivational ratings. Qualitative interview data were transcribed and analyzed thematically using Braun and Clarke's (2006) six-step approach, including familiarization, generating initial codes, theme development, reviewing themes, defining and naming themes, and producing the report.

3. RESULTS

3.1 Speaking Performance Before and After Using VATs

The quantitative analysis revealed statistically significant improvements in students' English-speaking fluency, pronunciation accuracy, and lexical range following the four-week voice assistant technology (VAT) intervention. A combination of pre- and post-intervention speaking assessments and survey responses were analyzed using paired-sample t-tests, and the results are presented in the following table:

Table 1. Pre- and Post-Test Quantitative Speaking Performance

Measure	Pre-Test Mean	Post-Test Mean	Mean Difference	p-value
Speaking Fluency (WPM)	88.4	97.2	+8.8	0.003
Pronunciation Accuracy (out of 10)	6.3	8.1	+1.8	0.002
Lexical Diversity (TTR)	0.47	0.58	+0.11	0.005

Table 1 shows the students' improvement in three speaking indicators: fluency, pronunciation, and lexical diversity. The fluency test scores, measured in words per minute (WPM), indicated that students' average speaking speed increased from 88.4 WPM before the intervention to 97.2 WPM afterward. This increase of 8.8 WPM reflects a statistically significant improvement in fluency with a p-value of 0.003.

Pronunciation was assessed using a pronunciation accuracy rubric scored by three independent raters, with a maximum score of 10. Students' average scores improved from 6.3 to 8.1 post-intervention, indicating better articulation, stress patterns, and intonation. This improvement was statistically significant ($p = 0.002$), suggesting that interaction with VATs enhanced learners' real-time pronunciation.

Lexical range was evaluated using lexical diversity measures, specifically the Type-Token Ratio (TTR). Students' mean TTR increased from 0.47 to 0.58, reflecting a greater variety of vocabulary in post-intervention speech. The observed difference was statistically significant with a p-value of 0.005.



These objective improvements were complemented by the self-reported survey data. The Likert-scale questionnaire explored students' levels of speaking anxiety, enjoyment of English-speaking activities, and motivation for daily practice. The analysis revealed significant increases in student confidence and enjoyment post-intervention.

3.2 Questionnaire Results

The results of the questionnaire assessing learners' motivation, speaking confidence, and their perception of VATs usefulness in improving their speaking skills are presented below:

Table 2. Learners' Perceptions Before and After VAT Use

Item	Pre-Test Mean	Post-Test Mean	Mean Difference	t-value	p-value
<i>I feel nervous speaking English.</i>	3.8	2.6	-1.2	4.92	<0.001
<i>I enjoy practicing English speaking.</i>	2.9	4.2	+1.3	-5.15	<0.001
<i>I feel motivated to use English daily.</i>	3.1	4.4	+1.3	-4.88	<0.001
<i>I find VATs easy to use for speaking practice.</i>	3.4	4.6	+1.2	-4.76	<0.001
<i>Using VATs helps me improve my pronunciation.</i>	2.7	4.3	+1.6	-5.02	<0.001
<i>Using VATs helps me improve my vocabulary (e.g., everyday and academic vocabulary).</i>	2.7	4.3	+1.6	-5.02	<0.001
<i>Using VATs helps me improve my fluency.</i>	3.1	4.4	+1.3	-4.88	<0.001
<i>I prefer using VATs over traditional speaking exercises.</i>	2.5	4.1	+1.6	-5.33	<0.001
<i>VATs make me more confident to speak English.</i>	2.8	4.5	+1.7	-5.21	<0.001

Table 2 illustrates the shifts in learner perceptions before and after using voice assistant technologies (VATs) in English-speaking practice. The data reveal statistically significant changes across all measured items, indicating a positive overall impact of VAT use on students' attitudes and self-perceived language development.

One of the most notable changes was a reduction in speaking anxiety. Before using VATs, students reported a relatively high level of nervousness when speaking English, with a mean score of 3.8. This decreased significantly to 2.6 after the intervention, suggesting that VATs helped create a more comfortable and less intimidating environment for oral practice. This reduction in anxiety is statistically significant, with a t-value of 4.92 and a p-value of less than 0.001.

Enjoyment and motivation also increased substantially following the integration of VATs into students' speaking routines. The mean score for enjoying English-speaking practice rose from 2.9 to 4.2, while motivation



to use English daily increased from 3.1 to 4.4. These results indicate that learners not only became more engaged but also more inclined to incorporate English into their everyday lives. The improvements in these areas were statistically significant, suggesting that VATs can positively influence learners' affective responses to language learning.

In terms of usability, students found VATs easier to use for speaking practice after the intervention. The mean score increased from 3.4 to 4.6, reflecting a stronger sense of comfort and accessibility with the technology. This suggests that the tools were perceived as user-friendly and did not pose significant technical barriers to language practice.

Additionally, learners perceived significant improvements in key areas of language development. The use of VATs was associated with notable gains in pronunciation, vocabulary, and fluency. The mean scores for pronunciation and vocabulary each increased from 2.7 to 4.3, while fluency rose from 3.1 to 4.4. These results imply that VATs provided valuable opportunities for repetition, exposure to new language, and consistent oral engagement—factors essential for skill development in EFL contexts.

Students also reported a clear preference for using VATs over traditional speaking exercises, with the mean score increasing from 2.5 to 4.1. Furthermore, their confidence in speaking English grew significantly, rising from a pre-mean of 2.8 to a post-mean of 4.5. These findings highlight the potential of VATs not only to support language acquisition but also to boost learners' confidence and autonomy in speaking.

3.3 Students' Preferred VAT

The results of semi-structured interviews revealed that, among the three voice assistant technologies (VATs) introduced during the intervention—Google Assistant, Amazon Alexa, and Apple Siri—Google Assistant emerged as the most preferred tool, with 79% of students reporting it as their primary choice for English-speaking practice. This strong preference was shaped by a combination of technological, contextual, and linguistic factors that made Google Assistant more accessible and practical for learners in the study.

One of the primary reasons behind this preference was device accessibility. The majority of participants owned Android smartphones, which come pre-installed with Google Assistant, allowing for immediate use without the need for additional downloads or setup. In contrast, Apple's Siri is only available on iOS devices, limiting access to students who owned iPhones—a minority within the participant group. Amazon Alexa, meanwhile, typically requires a separate Echo device or the installation of a dedicated app, and was thus less familiar and less available to the learners. This disparity in availability significantly influenced the frequency and ease with which students could engage with each VAT.

Another factor that contributed to the popularity of Google Assistant was its user-friendly interface and perceived responsiveness. Students reported that Google Assistant more accurately recognized non-native accents and allowed for smoother interaction, particularly when compared to Siri and Alexa. The assistant's ability to process a wide range of spoken prompts, ranging from casual to academic language, further enhanced its suitability for varied speaking tasks. Moreover, students appreciated the assistant's multimodal input flexibility, including the option to type or speak, which was useful in situations where voice recognition struggled.

Additionally, learners valued the way Google Assistant could be integrated into their daily routines, such as asking questions, receiving reminders, or searching for information. This integration made speaking English with the VAT feel more natural and less like a formal task, thereby lowering anxiety and encouraging repeated use.

4. DISCUSSION

The results of this study indicate that the integration of voice assistant technologies (VATs) into English as a Foreign Language (EFL) speaking practice produced significant improvements in both objective speaking performance and learners' perceptions. These findings are particularly relevant in low-resource contexts such as



Universitas Serambi Mekkah, where access to immersive English environments and native-speaking interlocutors is limited. The positive outcomes observed across multiple dimensions, including fluency, pronunciation, lexical diversity, and learner confidence, support the viability of VATs as supplementary tools for oral language development.

Quantitatively, the improvements in fluency (as measured by words per minute), pronunciation accuracy, and lexical diversity were all statistically significant, suggesting that regular engagement with VATs facilitated real gains in spoken English proficiency. The increase in fluency is particularly important, as it reflects learners' improved ability to produce language more fluidly and with less hesitation. The enhancement of pronunciation skills is consistent with prior research (e.g., Abimanto & Sumarsono, 2024; Alharthi, 2024), which suggests that VATs' capacity to model pronunciation and provide responsive interaction may help learners develop better articulatory habits. Likewise, the observed increase in lexical diversity suggests that students were not only using more words, but were also incorporating a wider range of vocabulary, potentially due to the exposure to varied prompts and responses from the VATs.

These objective outcomes are reinforced by students' self-reported perceptions, which show significant shifts in their emotional and cognitive engagement with English speaking. Most notably, there was a marked reduction in speaking anxiety. This supports the idea that VATs can serve as low-stakes interlocutors, allowing learners to practice speaking without the fear of judgment or embarrassment that often accompanies human interaction. The increase in motivation and enjoyment indicates that students found the experience not only beneficial but also enjoyable, which is a critical factor in sustaining long-term language learning. Motivation and reduced anxiety are known predictors of successful language acquisition, particularly in speaking, which many learners perceive as the most challenging skill (Arif, 2019; Stolvoort et al., 2024).

The usability of VATs also emerged as a strength. Students reported finding these tools easy to use, suggesting that technological barriers were minimal. This is an important consideration, especially in developing contexts where learners may have limited digital literacy or access to formal training. That students were able to use built-in or freely available VATs on their smartphones with such positive results implies that VAT-based speaking practice could be widely implemented without the need for expensive infrastructure or specialized software.

Furthermore, learners' increased confidence and their expressed preference for VATs over traditional speaking tasks are particularly encouraging. Confidence is a key element of speaking performance, and the data suggest that VATs may help create a supportive environment where students feel safe to experiment with language. This aligns with learner autonomy theories, which emphasize the importance of giving students tools and strategies they can use independently (Dickinson, 1987, as cited in Daflizar et al., 2022). The fact that learners preferred using VATs to more conventional classroom speaking exercises implies a potential pedagogical shift—educators might consider blending these technologies into their curriculum to enhance engagement and learner agency.

Qualitative data from interviews revealed that Google Assistant was the most frequently used VAT, preferred by 79% of participants. This preference was driven primarily by accessibility; most students owned Android smartphones with Google Assistant pre-installed—as well as its linguistic flexibility, recognition of non-native accents, and task integration features. Unlike Alexa, which required separate devices, and Siri, which was restricted to iOS users, Google Assistant provided a seamless, low-barrier option for regular speaking practice.

The students' ability to select and comfortably use a VAT of their choice suggests a growth in digital literacy and learner autonomy. Many learners expressed interest in continuing to use VATs beyond the study, indicating that such tools may contribute not only to immediate speaking gains but also to the long-term development of independent learning habits.

Overall, the findings underscore the pedagogical value of incorporating VATs into EFL instruction, especially in resource-limited contexts such as Aceh. Given their wide availability, ease of use, and ability to simulate conversation, VATs can function as accessible speaking partners, supplementing formal instruction with



authentic, low-stakes language practice. They offer a cost-effective way to extend speaking opportunities beyond the classroom, particularly for learners who have limited access to native speakers or English-speaking environments.

For educators, these findings suggest that VATs can be strategically integrated into speaking curricula to promote fluency development, learner confidence, and autonomy. Moreover, VAT-based activities can be adapted to various proficiency levels and learning styles, making them versatile tools in blended or remote learning scenarios.

5. CONCLUSION

This study investigated the impact of voice assistant technologies (VATs)—Google Assistant, Siri, and Alexa—on the speaking performance of EFL students at Universitas Serambi Mekkah. The findings revealed that integrating VATs into English-speaking practice significantly improved learners' fluency, pronunciation accuracy, and lexical diversity. These gains were supported by both quantitative speaking assessments and qualitative data on students' perceptions, indicating not only enhanced linguistic competence but also increased confidence, motivation, and engagement.

Learners experienced a notable reduction in speaking anxiety, reporting that VATs offered a low-pressure, non-judgmental environment that encouraged repeated oral practice. They also found VATs easy to use and enjoyable, making them more likely to practice consistently outside of the classroom. This is particularly important in low-resource contexts like Aceh, where access to immersive English environments and native speakers is limited.

Among the VATs examined, Google Assistant was the most frequently used and preferred tool, primarily due to its widespread availability on Android devices, user-friendly interface, and responsiveness to non-native speech. This suggests that learner preference is shaped not only by functionality but also by accessibility, ease of integration into daily routines, and familiarity with the platform.

The study demonstrates that VATs can be effective supplementary tools for promoting oral language development in EFL contexts. Their accessibility, adaptability, and potential to support learner autonomy make them particularly valuable in under-resourced educational settings. For educators, these findings highlight the importance of leveraging widely available technologies to extend speaking practice beyond classroom boundaries.

Future research should explore the long-term effects of VAT use on language retention, the impact of different VATs on other language domains such as listening or grammar, and the comparative effectiveness of VAT-assisted learning across proficiency levels. Additionally, studies incorporating classroom-based implementation and teacher-led VAT integration would further enrich the understanding of how these technologies can complement traditional pedagogical approaches.

REFERENCES

- Aakanksha, A., Jain, K., Giri, N., & Gupta, A. (2025). Assessing vulnerabilities in voice assistants: Comparative analysis of Google Assistant, Siri, and Alexa. *Advances in Knowledge-Based Systems, Data Science, and Cybersecurity*, 2(1), 10.
- Abimanto, D. & Sumarsono, W. (2024). Improving English pronunciation with AI speech-recognition technology. *Acitya: Journal of Teaching and Education*. 6. 146-156. 10.30650/ajte.v6i1.3810.
- Alagha, E. & Helbing, R. (2019). Evaluating the quality of voice assistants' responses to consumer health questions about vaccines: An exploratory comparison of Alexa, Google Assistant and Siri. *BMJ health & care informatics*. 26. 10.1136/bmjhci-2019-100075.
- Alharthi, S. (2024). Siri as an interactive pronunciation coach: its impact on EFL learners. *Cogent Education*. 11(1). doi:10.1080/2331186X.2024.2304245.



- Alrajhi, A. (2024). Intelligent personal assistants in self-access L2 vocabulary learning. *Education and Information Technologies*, 30(1), 801-831. doi: 10.1007/s10639-024-12967-7. https://www.researchgate.net/publication/383667416_Intelligent_personal_assistants_in_self-access_L2_vocabulary_learning
- Arif, M. (2019). The correlation of anxiety and motivation and speaking performance of English foreign learners, *LangEdu Journal*, 1(1).
- Aryanti, R. D., & Santosa, M. H. (2023). A systematic review on artificial intelligence applications for enhancing EFL students' pronunciation skill. *The Art of Teaching English as a Foreign Language (TATEFL)*, 5(1), 1-10. <https://doi.org/10.36663/tatefl.v5i1.718jurnal.stkipahsingaraja.ac.id>
- Brill, Tom & Munoz, Laura & Miller, Richard. (2019). Siri, Alexa, and other digital assistants: A study of customer satisfaction with artificial intelligence applications. *Journal of Marketing Management*, 35. 10.1080/0267257X.2019.1687571.
- Daflizar, Sulistiyo, U., & Kamil, D. (2022). Language Learning Strategies and Learner Autonomy: The Case of Indonesian Tertiary EFL Students. *LEARN Journal: Language Education and Acquisition Research Network*, 15(1), 257-281
- Dizon, G. (2017). Using intelligent personal assistants for second language learning: A case study of Alexa. *TESOL Journal*, 8(4), 811-830. https://www.researchgate.net/publication/321692268_Using_Intelligent_Personal_Assistants_for_Second_Language_Learning_A_Case_Study_of_Alexa
- Fathi, J., Rahimi, M., & Derakhshan, A. (2024). Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions. *System*, 121, 103254. <https://doi.org/10.1016/j.system.2024.103254BioMed Central>
- Junaidi, B., Hamuddin, B., Julita, K., Rahman, F., & Derin, T. (2020). Artificial intelligence in EFL context: Rising students' speaking performance with Lyra virtual assistance. *International Journal of Advanced Science and Technology*, 29(5), 6735-6741. <https://sersc.org/journals/index.php/IJAST/article/view/17726sersc.org>
- Nurmayasari, D. (2024). Leveraging artificial intelligence technology: Introducing speak.google as a digital speaking assistant for EFL students. *STRATEGY: Jurnal Inovasi Strategi dan Model Pembelajaran*, 4(1), 45-52. <https://doi.org/10.51878/strategi.v4i1.2927Jurnal P4I>
- Rahman, A. & Tomy, P. (2023). Voice assistant as a modern contrivance to acquire oral fluency: An acoustical and computational analysis. *World Journal of English Language*, 13(1). doi:10.5430/wjel.v13n1p92
- Stolvoort, J., Mackaaij, M., & Tribushinina, E. (2024). Age of onset, motivation, and anxiety as predictors of grammar and vocabulary outcomes in English as a foreign language learner with developmental language disorder. *Journal of Communication Disorders*, 108, 106407, ISSN 0021-9924, <https://doi.org/10.1016/j.jcomdis.2024.106407>.
- Tregubov, V. (2021). Voice assistants with artificial intelligence for improving academic English. *Conference: 10th International Conference on Information Technology Convergence and Services (ITCSE 2021)*, 01-12. 10.5121/csit.2021.110901.
- Zhang, M. (2024). Enhancing self-regulation and learner engagement in L2 speaking: Exploring the potential of intelligent personal assistants within a learning-oriented feedback framework. *BMC Psychology*, 12(421). <https://doi.org/10.1186/s40359-024-01917-0SpringerLink+1BioMed Central+1>