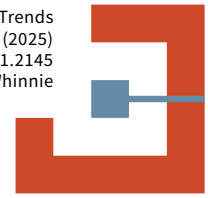
Trends

JALTCALL Trends
ISSN 2982-3749, 1, 2145 (2025)
https://doi.org/10.29140/jct.v1n1.2145



Castledown

Exploring the efficacy of ChatGPT's audio features in language learning

Steven G. B. MacWhinnie
Hirosaki Gakuin University
smacwhinnie@gmail.com

With the increased interest in integrating artificial intelligence (AI) into language education, this mixed-method study explored the use of ChatGPT's audio functionality to enhance Japanese students' oral communication skills. Japanese university students often complain that they have little chance to practice their English skills. To address this issue, the following research question was devised: Can ChatGPT effectively serve as a conversation partner for Japanese students aiming to improve their oral communication skills? Students participated in 10–20-minute chat sessions with ChatGPT, where they were required to interact through the speech-to-text function of the ChatGPT smartphone app. Following this, the students were given a reflective survey to draw out students' opinions on, and feelings about, their interactions. The results of this study indicate that while students have a generally positive attitude towards the use of Chatbots, there are concerns about privacy and ethical use.

Keywords: artificial intelligence, language learning, ChatGPT, smartphone applications, reflective analysis

Introduction

The ChatGPT audio-enabled app, introduced in late 2023 (OpenAI), has sparked interest among language teachers for its potential application as a conversation partner in language teaching (Dennisson & Ross, 2024). The current mixed-method study was designed to evaluate the ChatGPT audio-enabled app for its potential as a conversation partner; by using the app's audio functionalities to address a pervasive issue in language education—providing students in Japan with meaningful conversation partners. Students used the voice recognition and text-to-speech capabilities of the ChatGPT app to engage in conversation practice, followed by completing a questionnaire with both qualitative and quantitative items (Annamalai, 2023; Al-Abdullatif, 2023). This research specifically targeted the student-centric experience of using ChatGPT as a conversation partner. With the aim of understanding the nuances of students' perceptions, opinions, and feelings during interactions with AI.

Previous research has found that students recognize the value of AI for language learning, but also maybe be apprehensive of privacy and data security issues (Viktorivna et al., 2022; Antony & Ramnath, 2023). Other research has indicated AI's value for giving immediate feedback (Zou et al., 2023), other studies have focused on perceptions of usefulness of AI as a learning tool (Al-Abdullatif, 2023: Asif et al., 2023; Chan & Hu, 2023; Annamalai et al., 2023). All these studies point to the possible utility of AI for language learning but indicate that student perceptions play a role in chatbot acceptance.

While previous studies have explored text-based interactions with AI chatbots, real-time, spoken conversations using AI technology have only recently become possible (OpenAI, 2023). As such, little is known about how learners perceive and experience AI-driven audio interactions, particularly in mobile learning contexts. Addressing this gap, the present study investigates Japanese university students' experiences using ChatGPT's audio features for conversation practice, aiming to understand the nuances of students' perceptions, opinions, and feelings during interactions with AI.

Method

Participants

Six undergraduate English majors (three male, three female), aged 18 to 20, from a private university in Northern Japan participated in this study. All participants had intermediate English proficiency, approximately equivalent to CEFR levels B1 to B2, and were capable of engaging in basic spoken conversation using the ChatGPT audio interface. Selection criteria ensured that participants could reasonably understand and be understood by the speech recognition system. Informed consent was obtained from each participant prior to the study, including an explanation that ChatGPT may collect data during conversations.

Data collection

After installing and setting up the ChatGPT mobile app, participants completed one conversation session with ChatGPT using their personal smartphones. Five participants conducted their sessions in the researcher's office/lab, where the environment was quiet, and the researcher was available to assist with any technical difficulties if needed. One participant, who felt self-conscious about speaking in front of the researcher, completed the session alone in a nearby quiet classroom. While this meant the setting was not fully standardized, all participants were instructed to begin the conversation with the prompt "Let's chat" and to initiate the interaction by talking about their hobbies or other familiar topics. Each session lasted at least 10 minutes, with most extending to around 20 minutes.

Post-interaction, participants immediately completed a survey via Google Forms. The survey included 28 items rated on a 4-point Likert scale (1 = Does not apply, 4 = Applies) (Harzing, 2006), adapted from previous research on technology acceptance and chatbot interaction (e.g., Annamalai et al., 2023; Al-Abdullatif, 2023). These

items measured participants' perceived ease of use, perceived usefulness, perceived enjoyment, perceived value, attitude, perceived risk, and chatbot acceptance. The survey also included three open-ended questions, which prompted participants to elicit insights into (1) any difficulties encountered, (2) reflections on ChatGPT's effectiveness, and (3) suggestions for ethical AI use in language learning.

Qualitative data underwent thematic analysis informed by grounded theory methodology (Glaser & Strauss, 1967; Charmaz, 2006). Open-ended responses were read multiple times and initial codes were generated to capture recurring themes. These codes were organized into broader thematic categories that reflected patterns in participants' experiences and perceptions. No qualitative analysis software was used; all coding and categorization were conducted manually.

Quantitative data were analyzed in Microsoft Excel using basic descriptive statistics to summarize participants' perceptions of ChatGPT. Mean scores were calculated for each item on the Likert-scale questionnaire, offering a broad overview of perceived ease of use, usefulness, enjoyment, value, risk, and chatbot acceptance.

Results

The mean scores of the students (n=6) were calculated to give an overall summary of student perceptions (see Table 1). Perceived ease of use (3.38), perceived usefulness (3.37), perceived value (3.17), and attitude (3.21) all had a positive perception among the students. The results also indicate that there is not a strong sense of perceived risk (2.22), nor chatbot acceptance (2.7), suggesting concerns regarding data privacy and the readiness for integrating AI tools in language learning.

Table 1. Cumulative results

Category	Average score
Perceived ease of use	3.38
Perceived usefulness	3.37
Perceived enjoyment	3.11
Perceived value	3.17
Attitude	3.21
Perceived risk	2.22
Chatbot acceptance	2.70

The qualitative data (see Table 2) added depth to these findings. Of the comments related to technical difficulties, one student noted the difficulty in cellphone reception as a distractor, while two students noted the lack of control over the AI output. Notably, these control issues appeared to be related to problems with pacing. Students reported that ChatGPT sometimes failed to wait long enough for them to finish speaking. In terms of the actual experience of having a conversation, it was noted by all but one student that the chatbot served as good listening and/or speaking practices, with an emphasis on listening. Additionally, reflected in the data is ethical considerations with students voicing concern over the handling of their data. Two students noted that the chatbot did not understand or display emotions.

Table 2. Qualitative responses

Name	Technical difficulties	Effectiveness of conversations	Ethical concerns
A.T.	None	The chat itself was very enjoyable, but sometimes it was hard to understand by listening alone, so I want subtitles. Also, it would be smoother if ChatGPT could confirm I finished speaking.	I'm worried about the leakage of personal information. While I agree to use it in education, teachers must protect students' safety, so I want to know how personal information is protected.
R.Y.	Difficult to use smoothly with a poor signal, and it captures noise other than the voice.	It's very useful for listening.	Whether for good or bad, it remembers everything it's taught, so there should be standards for teaching it.
S.M.	It's trendy, and after all, it's a machine.	It contributed well. It would be a good practice partner before speaking with native speakers.	Cannot feel emotions.
M.S.	It would be helpful if there were a function to adjust speaking for learners.	Not only learning English but also learning how to behave as a listener in communication.	A system that allows the deletion of provided personal information is necessary.
Y.S	It was a bit difficult to speak freely without a topic.	Able to fully utilize speaking and listening abilities.	Unable to understand emotions.
S.O	It felt like it was asking too many questions at once.	The pressure to respond immediately is too great, and it's hard to settle on the content of the response.	Even if information is deleted, there's still a concern about information leakage.

Discussion

This study explored Japanese university students' perceptions of using the ChatGPT mobile app's integrated audio features as a conversation partner for language learning. While previous research has largely focused on text-based chatbot interactions (e.g., written chats), the ability to engage in real-time, spoken conversations with AI has only recently become available (OpenAI, 2023). As a result, no empirical studies have yet examined how ChatGPT's integrated audio features might impact language learning. Although a recent review has discussed the potential of ChatGPT's voice capabilities (Pratiwi et al., 2024), to the best of the researcher's knowledge, this study represents the first empirical investigation of ChatGPT's audio interface in a real-world language learning context, specifically with Japanese university students.

The quantitative results revealed generally positive attitudes towards the perceived ease of use and usefulness of ChatGPT's audio capabilities. These findings align with previous research highlighting the potential benefits of AI tools in enhancing interactivity and providing personalized language support (Chan & Hu, 2023; Annamalai et al., 2023). The scores for perceived enjoyment and value suggest room for improvement in fostering student engagement and perceived educational benefits, echoing Labadze et al.'s (2023) emphasis on chatbot interactivity as a critical factor influencing learning outcomes.

Lower scores for perceived risk and chatbot acceptance reflect the participants' cautious stance towards the integration of AI tools in language learning. This result mirrors studies that showed data privacy concerns (Viktorivna et al.,2022; Antony & Ramnath, 2023). In the qualitative feedback, students voiced concerns about personal information safeguards and the inability of AI systems to understand and display emotions, aligning with Huang's (2023) call for robust ethical frameworks in educational AI applications.

This study's findings contribute to the growing body of literature on the integration of AI in language education, underscoring the multifaceted nature of students' acceptance and perceived usefulness of AI chatbots. While acknowledging the potential benefits of these tools, the results emphasize the need to address technical limitations, enhance user engagement, and prioritize ethical considerations, such as data privacy and the preservation of authentic human-to-human interactions.

Crucially, the study's results support the notion that AI should complement, rather than replace, human judgment in educational settings (Seo et al., 2021). By fostering an environment where personalized learning experiences are balanced with the development of independent learning skills, AI tools like ChatGPT can serve as valuable resources without compromising critical thinking abilities or creating an over-reliance on technology.

These findings suggest that educators should consider using AI chat tools like ChatGPT as a supplement to traditional conversation practice, particularly in contexts where opportunities for real-time English communication are limited. Developers are encouraged to improve the pacing of AI responses to better accommodate the needs of lower-proficiency learners, and to ensure that clear privacy policies are in place to build user trust.

Conclusion

This study set out to explore the efficacy of ChatGPT's newly integrated audio features in facilitating language learning among Japanese university students. The findings corroborate the potential of AI tools, like ChatGPT, to serve as conversation partners, addressing the issue of limited opportunities for linguistic interaction in rural Japan. The quantitative data revealed generally positive attitudes towards the perceived ease of use and usefulness of ChatGPT's audio capabilities, underscoring their potential to enhance interactivity and provide personalized language support.

However, the study also highlighted the importance of addressing technical limitations, fostering student engagement, and prioritizing ethical considerations to ensure the responsible and effective integration of AI in educational contexts. Variations in individual participant responses underscore the complexity of user experiences with AI tools in education, reinforcing the critical importance of personalization and adaptability. The study's findings align with the broader consensus that AI should complement, rather than replace, human judgment in educational settings, fostering an environment where personalized learning experiences are balanced with the development of independent learning skills (Seo et al., 2021).

References

Al-Abdullatif, A. M. (2023). Modeling Students' Perceptions of Chatbots in Learning: Integrating Technology Acceptance with the Value-Based Adoption Model. *Education Sciences, 13,* 1151. https://doi.org/10.3390/educsci13111151

Annamalai, N., Rashid, R.A., Hashmi, U.M., Mohamed, M., Alqaryouti, M.H., Sadeq, A.E. (2023). Using chatbots for English language learning in higher education. *Computers and Education: Artificial Intelligence,* Volume 5, 2023, https://doi.org/10.1016/j.caeai.2023.100153.

Antony, S., & Ramnath, R. (2023). A phenomenological exploration of students' perceptions of AI chatbots in higher education. *IAFOR Journal of Education: Technology in Education, 11*(2). https://iafor.org/journal/iafor-journal-of-education/volume-11-issue-2/article-1/

Asif, M., Bibi, R., & Shahzaib, S. (2023). A survey of students' perceptions about the role of chatbots in English language learning at graduation level. *International Research Journal of Management and Social Sciences, 4,* 225–233. https://www.researchgate.net/publication/374812028_A_Survey_of_Students'_Perceptions_about_the_Role_of_Chatbots_in_English_Language_Learning_at_Graduation_Level

Chan, C., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education, 20*(1). https://doi.org/10.1186/s41239-023-00411-8

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis.* Sage.

Dennisson, J., & Ross, G. (2024). Utilizing ChatGPT for spoken task-based practices. In *JALTCALL 2024 Conference/Book of Abstracts* (p. 63). https://jaltcall.org/wp-content/uploads/2024/06/book-of-abstracts.pdf

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research.* Aldine.

Harzing, A.-W. (2006). Response styles in cross-national survey research: A 26-country Study. *International Journal of Cross Cultural Management, 6*(2), 243–266. https://doi.org/10.1177/1470595806066332

Huang, L. (2023). Ethics of artificial intelligence in education: Student privacy and data protection. *Science Insights Education Frontiers, 16,* 2577–2587. 10.15354/sief.23.re202.

Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: Systematic literature review. *International Journal of Educational Technology in Higher Education, 20*(56). https://doi.org/10.1186/s41239-023-00426-1

OpenAI. "ChatGPT Can Now See, Hear, and Speak." *Openai.com,* 25 Sept. 2023, https://openai.com/blog/chatgpt-can-now-see-hear-and-speak

Pratiwi, A. N., Hermayawati, H., & Saputra, D. (2024). Speaking practice using ChatGPT's voice conversation: A review on potentials and concerns. *Journal of Languages and Islamic Culture, 6*(1), 68–77. https://doi.org/10.54045/jlic.v6i1.149

Seo, K., Tang, J., Roll, I., Fels, S., & Yoon, D. (2021). The impact of artificial intelligence on learner–instructor interaction in online learning. *International Journal of Educational Technology in Higher Education, 18*(54). https://doi.org/10.1186/s41239-021-00292-9

Viktorivna, L. K., Oleksandrovych, A. V., Oleksandrivna, I. K., & Oleksandrivna, N. K. (2022). Artificial intelligence in language learning: What are we afraid of. *Arab World English Journal, 8*, 262–273. https://doi.org/10.24093/awej/call8.18

Zou, B., Reinders, H., Thomas, M., & Barr, D. F. (2023). Editorial: Using artificial intelligence technology for language learning. *Frontiers in Psychology, 14*. https://doi.org/10.3389/fpsyg.2023.1287667

JALTCALL Trends
Vol. 1
2024 JALTCALL Conference Papers
Special Issue

Author biodata

Steven G. B. MacWhinnie has been living and working in Japan since 2007. He holds a master's in TESOL from Shenandoah university. He is an associate professor in the English Department at Hirosaki Gakuin University. His research interests include AI, Timed Writing, and Motivation.