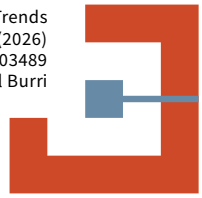
Trends

JALTCALL Trends
ISSN 2982-3749, 2(1), 103489 (2026)
https://doi.org/10.29140/jct.v2n1.103489



Castledown

Japanese EFL students' evolving use of GenAI tools: A two-year study

Kasumi Arciaga
Dokkyo University, Japan
k26003@dokkyo.ac.jp

Joel Neff
Utsunomiya University, Japan
joeldavidneff@gmail.com

Michael Burri
Wenzhou-Kean University, China
mburri@kean.edu

Considering the rapid development of generative artificial intelligence (GenAI) tools, this study examines the extent to which Japanese university students' use of such tools may change over two academic years. Using survey data from 1,034 respondents across two cohorts (2023 and 2024), independent sample t-tests (Welch's t-test) were conducted to compare mean differences in students' GenAI usage. Of the nine applications researched, only three (translation of instructions and textbooks, and AI use in class) demonstrated statistically significant differences, suggesting modest changes within a pattern of general stability. The study contributes to understanding AI integration in foreign language education by discussing how students' engagement with GenAI tools may change over time.

Keywords: artificial intelligence, English as a foreign language, educational technology

Introduction

Recent advances in Generative Artificial Intelligence (GenAI) have accelerated the integration of GenAI tools in educational contexts. Numerous GenAI products now create countless opportunities for English as a foreign language (EFL) learners to engage with authentic language input, receive personalized feedback, and access translation features beyond traditional classroom boundaries (Huang & Mizumoto, 2024). At the same time, students' adoption of these tools has raised critical questions about academic integrity, ethical usage, and the evolving – and perhaps what many fear, diminishing – role of teacher agency in the language learning process.

In the Japanese university context, where traditional pedagogical approaches often emphasize structured learning along with teacher authority, the introduction of GenAI tools has presented a particularly significant shift (Irwin, 2024). That is, Japanese university students must now navigate their language learning alongside readily available GenAI tools while still maintaining academic integrity and developing autonomous learning skills. This transition presents unique challenges and highlights the need for an empirically informed understanding of how Japanese university students' use of AI tools evolves over time. The impact of GenAI on language education cannot be fully captured through single-point-in-time investigations, as both GenAI technologies and students' AI literacy and ethical awareness continue to develop. Such insights are essential for educators seeking to design instructional practices and learning opportunities that meaningfully integrate GenAI use to support Japanese students' EFL development, including language production, feedback utilization, and learner autonomy.

This prompted the question of when and how students should integrate GenAI tools into their language learning experiences, which potentially impact teaching approaches, materials development, and assessment methods (Neff, et al., 2024a). Our previous work examined several ethical issues inherent in GenAI use in academic settings. While the research revealed Japanese university students' general awareness of GenAI tools, use appeared to be limited to spelling and grammar checking. Teachers and students appeared to hold differing views regarding what constitutes appropriate and ethical use of GenAI in academic contexts. Thus, we decided to replicate this initial study and turn it into a 2-year comparative cross-sectional project, allowing us to compare the results of the first cohort of students against a new cohort.

Taking a long-term perspective on a specific GenAI topic aligns with emerging longitudinal research in Japan. Mok (2024), for example, documented a significant shift in university students' AI adoption over just seven months, with ChatGPT usage increasing from 43.5% to 55.2%, and daily use rising from 6.6% to 18.9%. Similarly, Ohmori et al. (2023) found that while 94.1% of Japanese undergraduates demonstrated high awareness of ChatGPT, actual usage remained selective. Only 32.4% used it for completing homework with results showing notable gender differences (44.8% male vs. 27.1% female usage) in the use of ChatGPT, suggesting that awareness does not automatically equate to adoption of GenAI tools.

Against this backdrop, the present study examined how Japanese university EFL students' GenAI usage may have evolved from the initial adoption period in 2023 to a perhaps more mature phase in 2024. A comparison of quantitative survey data across two student cohorts allowed us to gain insights into students' potentially changing use of GenAI tools, informing both pedagogical practices and institutional policy. The current study addressed the following research question:

- To what extent does Japanese university EFL students' use of GenAI tools change over the course of two academic years?

JALTCALL Trends
Vol. 2 No. 1
2025 JALTCALL Conference Papers Special Issue

Methodology

The study employed a comparative cross-sectional design (Creswell & Creswell 2018), in which data were analyzed from survey results collected in 2023 and 2024. This design allowed for a comparison of students' usage across two similar student cohorts, and thus, provided insight into Japanese university students' use of GenAI tools over a 2-year period.

Participants

Survey responses from 1,034 Japanese university students across two academic years were analyzed. The 2023 cohort included 543 students (Agriculture n = 155, Engineering n = 148, Education n = 96, International Studies n = 66, and Regional Design n = 78), with the 2024 cohort being slightly smaller with 491 students (Agriculture n = 123, Engineering n = 147, Education n = 80, International Studies n = 81, Regional Design n = 60).

Data collection

Both cohorts completed the same survey instrument to ensure comparability across the two-year period. The survey consisted of four sections: (1) demographic questions (e.g., gender, department), (2) GenAI awareness measures (i.e., understanding of GenAI terminology and concepts), (3) GenAI applications (e.g., frequency of GenAI tool use) (see Table 1), and (4) six vignettes to assess students' perceptions of the ethical use of GenAI tools (see Neff et al., 2024b for survey).

Table 1. GenAI applications

GenAI applications included in part 3 of survey	Coding names
I have used GenAI tools to do homework assignments in their entirety, e.g. write an essay or speech.	ai_homework_help
I have used GenAI tools to help me complete my homework assignments.	ai_homework_complete
I have used GenAI tools to translate my teacher's instructions.	ai_translate_instructions
I have used general GenAI tools to translate my responses, e.g. if you wrote your answer in Japanese then used the A.I. to translate it to English.	ai_translate_responses
I have used GenAI tools at school.	ai_at school
I have used GenAI tools to translate my textbook.	ai_translate_textbook
I have used GenAI tools to create images for use in school work, e.g. graphics for a presentation.	ai_create_images
I have used GenAI. tools to create video for use in school work.	ai_create_video
I have used GenAI tools to answer questions during class.	ai_answer_questions

The questions in the third section of the survey (see Table 1) employed a three-point Likert scale: 0 = "No, I have not" (いいえ), 1 = "Yes, with teacher's permission" (はい、先生の許可を得て), and 2 = "Yes, on my own initiative" (はい、自分の意思で、自発的に). The items were presented in both English and Japanese to ensure students' understanding and smooth completion of the survey.

Also noteworthy is that the survey is part of a larger ongoing research project that explores English teachers and their students' awareness and ethical use of GenAI tools in Japanese tertiary education. However, with the present study focusing on Japanese students' potentially evolving use of GenAI tools, we only draw on section 3 of the survey shown in the left-hand column of Table 1. That section asked students to identify their experiences (i.e., applications) with GenAI tools they were presently using or had used in the past. The section included several examples of GenAI tools for students to indicate current or previous usage (e.g., ChatGPT, Google Translate, Deepl etc.). This initial list was created based on our own experiences with GenAI. Space was also provided for respondents to list other tools not provided in the list.

Data analysis

Statistical analysis was conducted using JASP (Jeffrey's Amazing Statistics Program) software. For each comparable variable (i.e., GenAI applications) between the two cohorts, independent samples *t*-tests using Welch's *t*-test (unequal variances) were performed in order to compare the means between responses provided by both cohorts. Cohen's *d* was also used to assess the practical significance of observed differences.

Results

As shown in Table 2, the independent samples *t*-test results revealed minimal variation in the use of GenAI tools between the 2023 and 2024 cohorts, with only three of the nine GenAI applications showing statistically significant differences. Specifically, the 2024 cohort showed less acceptance than the 2023 cohort towards the use of GenAI to translate classroom instructions ($t = -2.119$, $p = .034$, Cohen's $d = -0.132$) and general AI use in their university studies ($t = -3.148$, $p = .002$, Cohen's $d = -0.195$). By contrast, the 2024 cohort indicated greater acceptance towards the use of GenAI to translate textbook sections ($t = 2.153$, $p = .032$, Cohen's $d = 0.134$). The general "AI use at school" category featured the strongest statistical significance ($p = .002$) and largest effect size (Cohen's $d = -0.195$), showing a substantial decrease of AI use in 2024.

Table 2. Independent samples t-test

GenAI applications	*t*	*df*	*p*	Cohen's *d*	SE Cohen's *d*
ai_homework_help	−1.790	1,029.9	.074	−0.111	0.062
ai_homework_complete	−0.474	1,031.8	.636	−0.029	0.062
ai_translate_instructions	−2.119	1,022.8	.034	−0.132	0.062
ai_translate_responses	0.644	1,027.1	.520	0.040	0.062
ai_at school	−3.148	997.1	.002	−0.195	0.063
ai_translate_textbook	2.153	1,026.5	.032	0.134	0.062
ai_create_images	1.001	1,031.0	.317	0.062	0.062
ai_create_video	1.223	1,029.8	.221	0.076	0.062
ai_answer_questions	1.562	1,031.8	.119	0.097	0.062

Note. Welch's t−test.

JALTCALL Trends
Vol. 2 No. 1
2025 JALTCALL Conference Papers
Special Issue

For the remaining six GenAI applications, homework help ($t = -1.790$, $p = .074$), homework completion ($t = -0.474$, $p = .636$), translating responses to teachers' questions ($t = 0.644$, $p = .520$), creating images ($t = 1.001$, $p = .317$), creating videos ($t = 1.223$, $p = .221$), and answering questions ($t = 1.562$, $p = .119$), no statistically significant differences were identified ($p > .05$).

Overall, effect sizes for all comparisons were small (Cohen's d ranging from −0.195 to 0.134), indicating minor differences in GenAI use between the 2023 and 2024 cohorts. The most significant difference was in the translation-related GenAI application.

Discussion and conclusion

The minor increases in students' use of GenAI use to translate textbook sections (Cohen's $d = 0.134$), alongside a decrease in acceptance for translating classroom instructions (Cohen's $d = -0.132$) and general AI use at university (Cohen's $d = -0.195$), suggest that emerging patterns in student attitudes seem to show significant consideration and discussion to determine why and for what purpose students are using GenAI. As the survey did not directly assess students' underlying motivations the following interpretations are necessarily inferential. The results may indicate that Japanese university students are developing selective preferences for certain Gen AI tools and becoming more cautious about particular applications in formal academic contexts. One possible reason for this selective adoption of specific tools may stem from a lack of clear guidelines (McDonald et al., 2024). As institutions increasingly leave it to teachers on a per class basis as to whether GenAI usage is permissible, students may inevitably be guided towards specific tools or, alternatively, gravitate to ones that have not been specifically prohibited. One could also argue that institutional policies and classroom practices regarding Gen AI use have not substantially evolved between the two academic years.

That participating students were using GenAI tools is not surprising. However, the results provide evidence that student usage may only slowly increase and only in specific areas of students' tertiary studies. At the same time, the decline in the use of six identified applications by the 2024 cohort, including text translation and in-class work, raises the question of why. Students may have adopted newer, more sophisticated alternatives that better meet their specific academic requirements, suggesting students' preference towards more specialized, task-specific AI applications rather than decreased overall GenAI adoption.

Results also imply that university English programs in Japan ought to pay close attention to students' use of various GenAI tools rather than implementing broad policies. In contrast to our expectations at the onset of this study, the marginal overall increase and decrease in GenAI tool use in two thirds of the applications may be a reflection of students' academic integrity concerns about what exactly entails appropriate GenAI use in tertiary education (Cahill & McCabe, 2024).

By documenting different patterns of GenAI use, the present study contributes empirical insight into how Japanese EFL students selectively engage with emerging GenAI tools. As AI technology continues to evolve and more specialized tools become available, the different use across GenAI applications demonstrated in the

present study calls for targeted research examining faculty-specific and discipline-specific AI usage. Future investigations should also explore whether these modest shifts represent the beginning of larger attitudinal and usage changes or reflect a stabilization period as both students and institutions navigate GenAI integration practices and policies, given that students' motivations were discussed as possible explanatory factors.

JALTCALL Trends
Vol. 2 No. 1
2025 JALTCALL Conference Papers Special Issue

References

Cahill, C., & McCabe, K. (2024). Context matters: Understanding student usage, skills, and attitudes toward AI to inform classroom policies. *PS: Political Science & Politics, 57*(4), 594–601. https://doi.org/10.1017/S1049096524000155

Creswell, J. W., & Creswell, J. D. (2018*). Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage Publications.

Huang, J., & Mizumoto, A. (2024). The effects of generative AI usage in EFL classrooms on the L2 motivational self system. *Education and Information Technologies, 30*(11), 6435–6454. https://doi.org/10.1007/s10639-024-13071-6

Irwin, B. (2024). Exploring generative AI tools for English language learning: Student adoption and perceptions. *Sophia University Junior College Division Faculty Journal, 45,* 127–137.

McDonald, N., Johri, A., Ali, A., & Hingle, A. (2024). Generative artificial intelligence in higher education: Evidence from an analysis of institutional policies and guidelines. *arXiv*. https://doi.org/10.48550/arXiv.2402.01659

Mok, J. (2024). Students' use of generative artificial intelligence (ChatGPT) at Rikkyo University. *Foreign Language Education and Research, 5,* 21–39. https://fler.rikkyo.ac.jp/journal/jc0e3e00000000a9-att/JFLER(2024_Vol.5)_2_Mok.pdf

Neff, J., Arciaga, K., & Burri, M. (2024a). EFL students' and teachers' perceptions of the ethical uses of AI tools. *Technology in Language Teaching & Learning, 6*(3), 1714. https://doi.org/10.29140/tltl.v6n3.1714

Neff, J., Arciaga, K., & Burri, M. (2024b). EFL teacher perceptions of AI tools: Questionnaire from "EFL students' and teachers' perceptions of the ethical uses of AI tools" [Research instrument]. IRIS Database, University of York. https://doi.org/10.48316/Dd9qZ-xh3CQ

Ohmori, F., Saito, J., Matsuba, R., & Kita, T. (2023). Undergraduate students' perceptions of using ChatGPT and its influence on learning and academic integrity [Survey report]. Data-Based Education Research Center.

Author biodata

Kasumi Arciaga is a full-time lecturer in the Department of Foreign Languages at Dokkyo University. She holds an MSEd in TESOL from Temple University and a PhD in Applied Linguistics from De La Salle University. Her research interests include AI-integrated language education, EIL, and multilingual children in Japan.

Joel Neff is a Part-Time Lecturer at Utsunomiya University and the International University of Health and Welfare in Ohtawara. He holds an MA in Applied Linguistics

and TESOL from the University of Leicester. His research interests include the ethical use of AI, semantic change, and ludic approaches to language pedagogy.

Michael Burri is an Associate Professor in the Department of Curriculum and Instruction at Wenzhou-Kean University, and an Honorary Fellow in the School of Education at University of Wollongong. His professional interests include pronunciation instruction, teacher education/learning, Mind Brain Education, and the ethical use of GenAI.

JALTCALL Trends
Vol. 2 No. 1
2025 JALTCALL Conference Papers
Special Issue