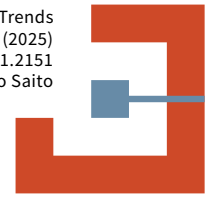
Trends

JALTCALL Trends
ISSN 2982-3749, 1, 2151 (2025)
https://doi.org/10.29140/jct.v1n1.2151



Castledown

Development and evaluation of a course integrating machine translation and generative AI in higher education English learning

Michiyo Oda
Reitaku University
mioda@reitaku-u.ac.jp

Ryo Saito
Tohoku University
ryo@tohoku.ac.jp

This study developed and implemented a course aimed at helping students understand the strengths and limitations of machine translation (MT) and generative artificial intelligence (GenAI), utilize these AI tools to produce higher-quality English translations, and apply them to learning English as a second language. Data from 46 students were analyzed using pre- and post-course surveys, measuring their overall course satisfaction, usefulness, and motivation. The survey results showed that most students were satisfied with the course and found it useful for learning English. Further analysis revealed that the students found that learning about the strengths and limitations of MT, pre- and post-editing, and using ChatGPT useful for learning English. In particular, understanding the strengths and limitations of MT and creating their own prompts contributed to overall satisfaction and usefulness. Additionally, students' self-evaluated achievements in pre- and post-editing contributed to their willingness to continue learning English even as AI develops.

Keywords: machine translation, generative artificial intelligence, university students

Introduction

Recent advancements in natural language processing technologies have led to the development of large language models, which are foundational technologies of generative artificial intelligence (GenAI). The advancement of artificial intelligence (AI) tools, such as machine translation (MT) and GenAI, has generated controversy among

language educators regarding passive learning and academic integrity (Tasmedir et al., 2023; Nakai, 2023), despite students' continued use of AI tools even under explicit prohibition policies (Clifford et al., 2013). The widespread use of these AI tools has raised concerns about students' motivation to learn foreign languages as AI develops (Oda, 2019), especially since MT has become widely used.

Despite educators' concerns, research into AI and language learning has flourished. Law (2024) conducted a literature review on GenAI and language education and suggested that several studies have shown potential benefits, including personalized feedback and improved language learning outcomes, although potential risks such as ethical concerns and over-reliance remain. While recent studies on GenAI and language learning have focused on writing and assessment (Law, 2024), Bowker (2020) emphasized the importance of MT literacy, focusing on cognitive aspects such as critical thinking over technical knowledge. Building on Bowker's (2020) argument, this study emphasizes nurturing learners to not only use AI tools to enhance language learning but also to apply AI literacy, including understanding its strengths, limitations, mechanisms, and ethical considerations. In addition, several studies have shown positive results regarding motivation with GenAI for specific tasks (Guo et al., 2023) or mixed results on motivation as a general perception (Yokono, 2023). However, no studies have assessed learners' motivation in acquiring AI literacy for language-learning.

With this background, this study aimed to develop and evaluate a course to nurture AI literacy for learning English as a second language. This led to the following research questions:

RQ1. What topics covered in the course do students perceive as effective, and what do they identify as their achievements during the course?

RQ2. How do students' perceptions of the effectiveness of course topics and their self-evaluations of achievements influence overall course satisfaction, course usefulness, and motivation?

Method

Course development

This study was conducted at a private university in Chiba, Japan. The course, titled "English Communication: Machine Translation and AI," was developed for the fall semester of the 2023 academic year. This course aimed to help students understand the strengths and limitations of MT and GenAI, utilize these tools to produce higher-quality English translations, and apply them to English learning. The content of the course, shown in Table 1, was developed with reference to studies including Bowker (2020) and UNESCO (2023). The course included 14 classes, each lasting 100 minutes.

Table 1. Course content

Week	Topics	Content
Week 1	Orientation	– Course introduction and pre-course survey.
Week 2	Characteristics of MT	– Awareness of different MT systems producing varied results.
Week 3	Strengths and limitations of MT	– Understanding improvements in MT accuracy and fluency, and recognizing common errors.
Week 4	History of MT, data-driven MT systems	– Understanding rule-based, statistical, and neural MT. – MT uses probability calculations using large datasets.
Week 5 – Week 8	Pre-editing and post-editing	– Utilizing pre- and post-editing for English translations. – Understanding the rules and ethical considerations for using MT.
Week 9	Creating MT using Google Colaboratory.	– Programming with a Python library for the Google Translate API.
Week 10	Definition and features of ChatGPT.	– Learning how ChatGPT works, creating a ChatGPT account, and exploring its features.
Week 11	Introduction to prompts for learning English	– Using instructor-introduced prompts in ChatGPT with sample texts.
Week 12	Prompt engineering for learning English	– Understanding tips for creating prompts for ChatGPT. – Creating students' own prompts.
Week 13	Review	– Review course content.
Week 14	Reflection	– Post-course survey

Participants

Participants were third- and fourth-year Japanese students majoring in English who took the course in the fall semester of the 2023 academic year. In the first (pre-course survey) and last (post-course survey) classes, students participated in online surveys (Google Forms). Data from 46 students (40 third-year students and 6 fourth-year students) who completed the pre- and post-course surveys without missing data were used in the analysis.

Survey items

All survey items were created to align with the purpose of this study (Table 2). These items were categorized into four groups and measured using a six-point Likert scale (1: strongly disagree; 6: strongly agree).

Table 2. Survey items

Category	Number of items	Survey question(s)
Course satisfaction/Course's usefulness for learning English	2	– "Were you satisfied with this course?" – "Do you think this course is useful for your English learning?"
Willingness to continue learning English even as AI develops	1	– "Do you think you will continue learning English even as MT and AI develop?"
Perceived effectiveness of course topics	1	– "Do you think the following topics covered in this course will be effective for your English learning?" *(Participants rated the perceived effectiveness of each topic covered in the course for English learning; see Figure 1.)*
Students' self-evaluation of perceived achievements	6	– Three items related to MT – Three items related to GenAI *(see Figure 2)*

Results and discussion

RQ1: What topics covered in the course do students perceive as effective, and what do they identify as their achievements during the course?

Figure 1 shows the perceived effectiveness of the course topics. The first six topics are related to MT, and the next three topics are related to GenAI. More than 80% of students chose "strongly agree" or "agree" for "strengths and limitations of MT," and "pre-editing and post-editing" from the MT topics and all three topics from the GenAI category. Figure 2 illustrates the students' self-evaluation of their perceived achievements. The first three items are related to MT, and the following three items are related to GenAI. No students chose "strongly disagree." "Be able to write effective prompts for English learning" had the lowest rating (54% "strongly agree" or "agree"), while other items ranged from 65% to 70%.

These results indicate that students found it useful to learn about the strengths and limitations of MT, pre-editing and post-editing, and the use of ChatGPT for learning English. Furthermore, while this is speculative, the course alone may not have been sufficient to fully develop the students' confidence in writing effective prompts for learning English.

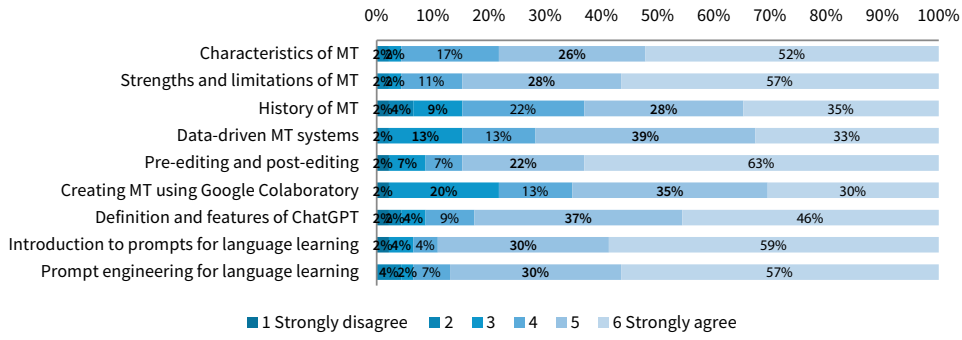


Figure 1. Perceived effectiveness of the course's topics

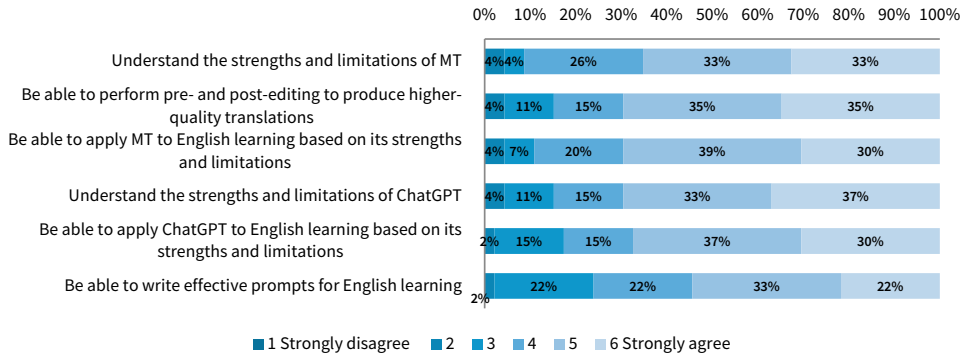


Figure 2. Students' self-evaluation of perceived achievements

RQ2: How do students' perceptions of the effectiveness of course topics and their self-evaluations of achievements influence overall course satisfaction, course usefulness, and motivation?

In total, 94% of students selected "strongly agree" or "agree" for course satisfaction, while 89% did the same for course usefulness. To analyze how perceived effectiveness and students' self-evaluations influenced overall course satisfaction, course usefulness, and motivation, a multiple regression analysis was conducted using the enter method in SPSS ver.29. Course satisfaction, the course's usefulness for learning English, and willingness to continue learning English (changes in average scores from pre- to post-survey) were dependent variables, while the perceived effectiveness of course topics and students' self-evaluations were independent variables. Three variance inflation factors for perceived effectiveness of MT topics exceeded 10, so two were removed and the analysis repeated. As a result, all variance inflation factors were below 10.

Tables 3 and 4 present the correlation analysis results of perceived effectiveness and students' self-evaluations, respectively. Table 5 shows how the perceived effectiveness affects overall course satisfaction, course usefulness, and motivation. For both course satisfaction and usefulness, "strengths and limitations of MT" had significant positive effects, while "introduction to prompts for English learning" had a significant negative effect. "Prompt engineering for English learning" has significant positive effects only for course satisfaction. Table 6 illustrates how students' self-evaluations affect overall satisfaction, course usefulness, and motivation. Only "be able to perform pre- and post-editing to produce higher-quality translations" had a significant positive effect on changes in willingness to continue learning English.

Table 3. Correlational analysis of perceived effectiveness of the course's topics

		1	2	3	4	5	6	7	8	9	10	11	12
1	Characteristics of MT	–											
2	Strengths and limitations of MT	0.92***	–										
3	History of MT	0.72***	0.68***	–									
4	Data-driven MT systems	0.75***	0.73***	0.91***	–								
5	Pre-editing and post-editing	0.88***	0.86***	0.57***	0.65***	–							
6	Creating MT using Google Colaboratory	0.62***	0.64***	0.73***	0.69***	0.53***	–						
7	Definition and features of ChatGPT	0.80***	0.81***	0.71***	0.59***	0.72***	0.66***	–					
8	Introduction to prompts for language learning	0.79***	0.85***	0.62***	0.63***	0.80***	0.63***	0.79***	–				
9	Prompt engineering for language learning	0.78***	0.80***	0.62***	0.61***	0.83***	0.53***	0.80***	0.87***	–			
10	Course satisfaction	0.60***	0.60***	0.47**	0.51***	0.51***	0.38**	0.51***	0.44**	0.57***	–		
11	Course usefulness for English learning	0.68***	0.73***	0.51***	0.50***	0.59***	0.45**	0.68***	0.52***	0.62***	0.63***	–	
12	WCLE	0.14	0.11	0.05	−0.09	0.02	0.10	0.14	0.08	0.00	−0.08	0.06	–

Note. WCLE signifies willingness to continue learning English even as AI develops (changes in average scores from pre- to post-survey). †p< 0.10, *p< 0.05, **p< 0.01, ***p< 0.001.

Table 4. Correlational analysis of students' self-evaluation of perceived achievements

		1	2	3	4	5	6	7	8	9
1	Understand the strengths and limitations of MT.	–								
2	Be able to perform pre- and post-editing to produce higher-quality translations.	0.81***	–							
3	Be able to apply MT to English learning based on its strengths and limitations.	0.85***	0.86***	–						
4	Understand the strengths and limitations of ChatGPT.	0.61***	0.46**	0.57***	–					
5	Be able to apply ChatGPT to English learning based on its strengths and limitations.	0.59***	0.51***	0.59***	0.82***	–				
6	Be able to write effective prompts for English learning.	0.60***	0.49**	0.52***	0.67***	0.78***	–			
7	Course Satisfaction	0.47**	0.44**	0.44**	0.20	0.19	0.30*	–		
8	Course Usefulness for English Learning	0.55***	0.57***	0.55***	0.25 †	0.38*	0.38**	0.63***	–	
9	WCLE	−0.06	0.19	0.00	-0.05	−0.03	−0.02	−0.08	0.06	–

Table 5. Multiple regression analysis of perceived effectiveness of the course's topics

	Course satisfaction		Course usefulness for English learning		WCLE	
	β		β		β	
MT						
Strengths and limitations of MT	0.81 *	[0.13 , 1.08]	1.03 ***	[0.38 , 1.16]	0.46	[−0.41 , 1.44]
Data-driven MT systems	0.13	[−0.16 , 0.33]	−0.06	[−0.24 , 0.16]	−0.42 †	[−0.87 , 0.08]
Pre-editing and post-editing	−0.30	[−0.52 , 0.14]	−0.28	[−0.45 , 0.10]	−0.13	[−0.76 , 0.53]
Creating MT using Google Colaboratory	0.02	[−0.20 , 0.22]	−0.02	[−0.18 , 0.16]	0.14	[−0.28 , 0.52]
GenAI						
Definition and features of ChatGPT	−0.06	[−0.32 , 0.26]	0.30	[−0.05 , 0.42]	0.22	[−0.36 , 0.76]
Introduction to prompts for English learning	−0.72 *	[−0.88 , −0.10]	−0.73 **	[−0.82 , −0.18]	0.04	[−0.71 , 0.80]
Prompt engineering for English learning	0.76 *	[0.11 , 0.92]	0.47 †	[−0.01 , 0.65]	−0.30	[−1.08 , 0.47]
R^2	0.50		0.67		0.14	

Note. β signifies the standardized regression coefficient, [,] indicates the 95% confidence interval, and WCLE represents willingness to continue learning English even as AI develops (change in average scores from pre- to post-survey). † p< 0.10, *p< 0.05, **p< 0.01, ***p< 0.001

Table 6. Multiple regression analysis of students' self-evaluation of perceived achievements

	Course satisfaction		Course usefulness for English learning		WCLE	
	β		β		β	
MT						
Understand the strengths and limitations of MT	0.30	[−0.18 , 0.57]	0.27	[−0.17 , 0.51]	−0.49	[−1.05 , 0.12]
Be able to perform pre- and post-editing to produce higher-quality translations	0.11	[−0.28 , 0.41]	0.23	[−0.18 , 0.45]	0.87 **	[0.24 , 1.31]
Be able to apply MT to English learning based on its strengths and limitations	0.17	[−0.31 , 0.52]	0.15	[−0.28 , 0.47]	−0.36	[−0.99 , 0.30]
GenAI						
Understand the strengths and limitations of ChatGPT	−0.06	[−0.33 , 0.26]	−0.36	[−0.49 , 0.06]	0.09	[−0.39 , 0.54]
Be able to apply ChatGPT to English learning based on its strengths and limitations	−0.28	[−0.53 , 0.19]	0.23	[−0.18 , 0.47]	−0.07	[−0.62 , 0.49]
Be able to write effective prompts for English learning	0.24	[−0.14 , 0.42]	0.10	[−0.20 , 0.32]	0.03	[−0.41 , 0.47]
R^2	0.27		0.40		0.19	

Note. β signifies the standardized regression coefficient, [,] indicates the 95% confidence interval, and WCLE represents willingness to continue learning English even as AI develops (change in average scores from pre- to post-survey). **p< .01

These results suggest that students who value understanding MT's strengths and limitations, and value creating prompts are more satisfied with the course. However, simply using GenAI prompts created by others may negatively affect course satisfaction and usefulness. Furthermore, confidence in performing pre- and post-editing with MT increases students' willingness to continue learning English as AI develops.

Conclusion

This study developed and implemented a course focused on MT and GenAI for learning English, demonstrating its contribution to student satisfaction and usefulness. This study also revealed that if students feel confident in performing pre- and post-editing with MT, their willingness to continue learning English as AI develops will increase. Although this study investigated satisfaction, usefulness, and motivation in the course, it did not assess how MT and ChatGPT content contributed to these aspects separately. Future studies will explore MT and GenAI separately to gain deeper insights into students' perceptions and motivation regarding these tools.

References

Bowker, L. (2020). Machine translation literacy instruction for international business students and business English instructors. *Journal of Business & Finance Librarianship, 25,* 1–19. https://doi.org/10.1080/08963568.2020.1794739

Clifford, J., Merschel, L., & Munné, J. (2013). Surveying the landscape: What is the role of machine translation in language learning? *@tic Revista d'innovació Educativa (Research and Education and Learning Innovation Archives), 10,* 108–121. https://doi.org/10.7203/attic.10.2228

Guo, K., Zhong, Y., Li, D., & Chu, S. K. W. (2023). Effects of chatbot-assisted in-class debates on students' argumentation skills and task motivation. *Computers & Education, 203,* 104862. https://doi.org/10.1016/j.compedu.2023.104862

Law, L. (2024). Application of generative artificial intelligence (GenAI) in language teaching and learning: A scoping literature review. *Computers and Education Open, 6,* 100174. https://doi.org/10.1016/j.caeo.2024.100174

Nakai, H. (2023). 生成AIの中等英語教育における可能性と課題 [Prospects and challenges of using generative AI in secondary English education]. *Journal of Osaka Jogakuin University, 20,* 145–164.

Oda, T. (2019). 機械翻訳と共存する外国語学習活動とは [Machine translation and foreign language education in Japan]. *Tokyo Keizai University The Journal of Humanities and Natural Sciences, 145,* 3–27.

Tasmedir, S., Lopez, E., Satar, M., & Riches, N. (2023). Teachers' perceptions of machine translation as a pedagogical tool. *The JALT CALL Journal, 19,* 92–112. https://doi.org/10.29140/jaltcall.v19n1.24

UNESCO. (2023). *Guidance for generative AI in education and research.* UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000386693

Yokono, N. (2023). AI通訳・翻訳の使用が学生の英語学習意欲に与える影響 [Effects of the use of AI interpretation and translation on students' motivation to learn English]. *Memoirs of Kanazawa Seiryo University Women's Junior College,* 52, 15–26.

Author biodata

Michiyo Oda received a PhD in Information Sciences from Tohoku University. Her research interests include machine translation, AI and VR-supported language learning, and instructional design.

Ryo Saito received a PhD in Information Sciences from Tohoku university. His research interests include cognitive science and psychology in educational practice.

JALTCALL
Trends
Vol. 1
2024 JALTCALL
Conference Papers
Special Issue