English Language Teaching; Vol. 16, No. 11; 2023
ISSN 1916-4742 E-ISSN 1916-4750
Published by Canadian Center of Science and Education

Actual Usage of Machine Translation by Japanese University Students and Verification of Test Results

Chiho Toyoshima[1] & Tsukasa Yamanaka[2]

[1] Research Organization of Science and Technology, Ritsumeikan University, Kusatsu, Japan

[2] College of Life Sciences, Ritsumekan University, Kusatsu, Japan

Correspondence: Chiho Toyoshima, Research Organization of Science and Technology, Ritsumeikan University, Kusatsu, Shiga, 525-8577, Japan. E-mail: c.toyoshima1113@gmail.com

Received: September 30, 2023 Accepted: October 29, 2023 Online Published: October 31, 2023

doi: 10.5539/elt.v16n11p83 URL: https://doi.org/10.5539/elt.v16n11p83

Abstract

The objective of this study is to investigate the actual situation of Japanese university students' use of machine translation (MT). The case study focuses on Japanese university students and not only investigates when students use MT, but also examines how their attitudes change before and after they use MT for their assignments. In this study, Google Translate was used as the MT tool, and Microsoft Excel was used for analysis. By analyzing these results, it was found that when students were allowed to use MT, they themselves decided whether or not to use it depending on their task. Of the skills in writing, reading and listening, it is also found that students tend to use MT the most for writing tasks and the least for listening tasks. In addition, no statistical significance of using MT was found for any of these skills, indicating that the use of MT does not necessarily mean that all language-related questions can be solved. These results could provide valuable data for the future introduction of MT into education. The survey included a diverse range of university students selected through an open application process. However, the sample size was limited, so an extensive survey should be undertaken in the future.

Keywords: Machine translation (MT), Japanese university, the Common University Entrance Examination, ETS Criterion

1. Introduction

In recent years, advances in MT have made it possible for us to use high-quality MT inexpensively and easily. Prior studies have shown how MT has become a tool that learners use on a daily basis. (Clifford et al., 2013; Jolly and Maimore, 2015) While it is possible to prohibit students from using MT in class, it is practically impossible to prohibit students from using MT outside of class. This is compounded that the errors generated by MT systems often closely resemble those made by humans, making it exceedingly challenging for educators to discern whether learners are utilizing MT tools or relying solely on their own language proficiency (Stapleton and Kin, 2019). Moreover, despite institutional policies prohibiting the use of MT in relation to assignments that are graded, it is inevitable that students will use MT to work on assignments. O'Neill (2019), in an extensive survey of second-year university students studying Spanish and French, discovered that more than 80% of them engaged with MT even in the situations when its usage was expressly forbidden. In light of these trends, it is imperative to consider the integration of MT as a constructive learning tool. As long as it is no longer possible for teachers to correctly distinguish between machine-generated and non-machine-generated texts, it is impossible to prohibit the use of MT and grade students fairly. In recent years, many Japanese universities have also announced regulations on the use of ICTs such as MT and ChatGPT. Some universities have banned them on a limited basis, but few have stated that they are banned completely. This prompts us to ponder the harmonious coexistence and the extent to which this technology can bolster foreign language education, ultimately enhancing the efficacy of learning experiences. It is conceivable that the rate of MT utilization varies depending on distinct factors, such as the language of instruction and the nature of the tasks assigned. Paradoxically, despite the increasing prevalence of MT, there is a conspicuous dearth of comprehensive research concerning its integration into English language education, particularly in the context of English and Japanese language learning.

With these considerations in mind, this study endeavors to conduct a comprehensive survey to focus on the current status of MT usage among Japanese university students. There is concern that the increasing number of learners using MT will lead to a decline in learners' language skills and thinking ability, but will MT enable learners to solve all language-related problems? The investigation aims to discern the specific junctures at which English learners opt to use MT when undergoing assessments spanning various language skills, including writing, reading, and listening. It also examines the degree of use of MT by proficiency level, to see at what level it is useful for learners. It further seeks to probe the practical implications of MT adoption, particularly when students solve standardized test questions, while investigating any disparities in performance outcomes under conditions where MT utilization is either allowed or prohibited.

2. Literature Review

Previous studies have investigated the use of MT by language learners; Jolley and Maimore (2015) conducted a survey of 128 learners participating in a Spanish program on their MT use. The findings demonstrated a notably high adoption of MT among learners, with a staggering 97.66% utilizing MT in their language learning. The study also probed learners' ethical awareness, with 86.72% of respondents expressing that the acceptability of MT usage hinged on the nature of the task. For instance, the use of MT in presentations or writing assignments was not widely considered as cheating, but its application in translation assignments raised ethical concerns. Thus, it is evident that a significant portion of language learners incorporate MT into their daily academic routine. While there is still a lack of research between languages with different linguistic structures, such as Japanese to English, various studies have already been conducted on the use of MT in European languages to English.

Niño (2020) delved into the utility of MT when students applied it to their assignments. The study engaged 37 Spanish learners, comprising 30 advanced learners, 5 intermediate-advanced learners, 1 intermediate learner, and 1 beginner learner. Participants were tasked with solving reading, listening, writing, and speaking questions using MT, followed by a questionnaire to gauge the efficacy of MT. The results indicated that for reading and writing tasks, the majority of learners (75.6% respectively) found MT to be beneficial, especially in terms of vocabulary, grammar, and structure. However, advanced learners were more discerning, acknowledging the potential for errors introduced by MT and, consequently, deeming it less indispensable. In contrast, MT did not garner significant recognition for its usefulness in listening and speaking tasks, with only 53.1% of respondents acknowledging that MT aided their understanding in listening tasks and 53.8% in speaking tasks. These outcomes suggest a varied perception of MT's utility depending on the task and proficiency level.

Concurrently, there has been extensive research on MT in foreign language learning in Japan. Oda (2019) conducted a survey examining Japanese university students' attitudes toward MT. In 2019, a questionnaire on MT was administered to 90 students majoring in fields other than English, and the results revealed that 96.7% of the respondents had used MT. Remarkably, only 2.2% believed that MT should be restricted in the context of their college assignments. Intriguingly, only 24.1% of students reported that Japanese-to-English translations using MT yielded satisfactory results, indicating a degree of dissatisfaction with MT quality. Given the widespread adoption of MT, it becomes imperative to explore avenues for its incorporation into university education.

From the educators' perspective, Yamada et al. (2021) indicated a willingness among university instructors to leverage MT in their classrooms. However, the challenge lies in determining how to effectively integrate MT into the teaching process. These studies collectively underscore the increasing prevalence of MT among language learners, emphasizing the necessity to identify the specific tasks for which students employ MT and the extent to which MT can enhance their learning experiences.

3. Method

The participants were 28 university students with varying levels of English proficiency. Participants were recruited from the public and the number of participants was closed on a first-come, first-served basis. Since the English level was self-reported at the time of application, its authenticity is not certain, but from the results of the English test, English ability seemed to be scattered. The recruitment process specified that "English proficiency was not required," with the sole condition being that participants must be able to attend the university in Osaka on the day of the experiment. Data analysis was conducted using Microsoft Excel. Participants were tasked with solving reading and listening questions extracted from the Common University Entrance Examination. Additionaly, some writing questions were assigned from ETS Criterion (https://www.ets.org/criterion.html). Participants were required to answer some of the questions solely using their own abilities, without the aid of MT, while they were given the optional permission to use MT for the remaining questions. In total, there were 20

questions, with 12 from the listening section (A-L), 6 from the reading section (M-R), and 2 from the writing section (S-T). Google Translate was specified as the MT tool.

To maintain uniformity and control, participants were randomly divided into two groups, namely Group A and Group B. They then followed the procedural steps outlined in Figure 1 while completing the questions. In Group A, MT was allowed in Test 1 for reading questions and in Test 2 for listening and writing questions. On the other hand, in Group B, MT was allowed for listening and writing questions in Test 1, and for reading questions in Test 2. This procedure was taken in consideration of the order exchange. In order to survey students' attitudes toward MT, a questionnaire was also conducted before and after they worked on the test. Students are monitored as they take the test and they do not receive any assistance other than MT.

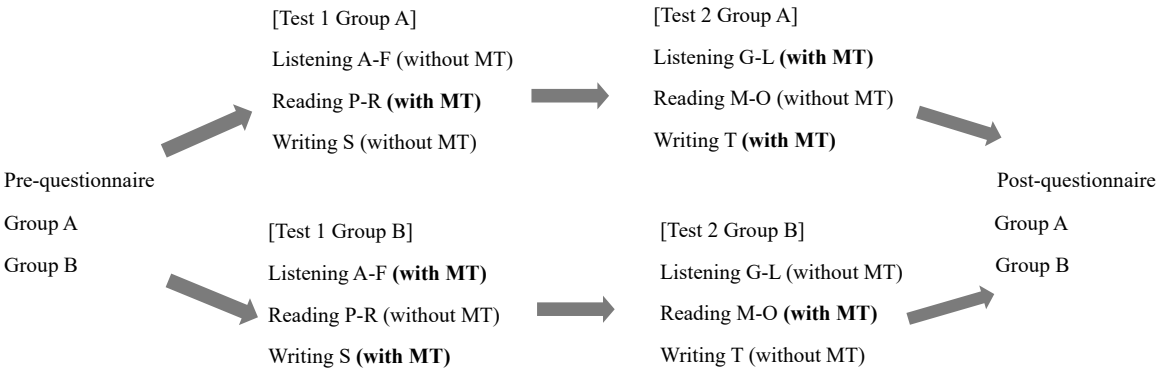

Figure 1. The order of tests

4. Results

4.1 Results of writing test

Table 1, 2, 3, and 4 provides an overview of the average results of the writing test, comprising both with and without MT, for the participants in the current study. Table 1 and 2 compare the differences in results by group. Table 1 shows the results of Group A, which solved the test without MT in Test 1 and used MT in Test 2. Table 3 shows the results of Group B. Group B solved the test with MT in Test 1 and without MT in Test 2. Tables 3 and 4 compare the differences in results by test, which means that scores are compared when participants solve the same problem with and without MT.

Table 1. Writing test results with and without MT in Group A

	Test 1 (without MT)	Test 2 (with MT)
Mean	1.466666667	1.933333333
Variance	1.40952381	2.20952381
Observation	15	15
Pearson Correlation Coefficient	0.180788148	
Hypothesized Mean Difference	0	
Degree of Freedom	14	
t	-1.046826941	
P(T<=t) One-tail	0.156458566	
t Critical One-tail	1.761310136	
P(T<=t) Two-tail	0.312917133	
t Critical Two-tail	2.144786688	

Table 2. Writing test results with and without MT in Group B

	Test 1 (with MT)	Test 2 (without MT)
Mean	2	1.230769231
Variance	2.5	0.525641026
Observation	13	13
Pearson Correlation Coefficient	0.07269493	
Hypothesized Mean Difference	0	
Degree of Freedom	12	
t	1.640299655	
P(T<=t) One-tail	0.063437188	
t Critical One-tail	1.782287556	
P(T<=t) Two-tail	0.126874377	
t Critical Two-tail	2.17881283	

Table 3. Writing test results for Test 1 (Group A: without MT, Group B: with MT)

	Group A (without MT)	Group B (with MT)
Mean	1.466666667	2
Variance	1.40952381	2.5
Observation	15	13
Hypothesized Mean Difference	0	
Degree of Freedom	22	
t	-0.996796027	
P(T<=t) One-tail	0.164850822	
t Critical One-tail	1.717144374	
P(T<=t) Two-tail	0.329701645	
t Critical Two-tail	2.073873068	

Table 4. Writing test results for Test 2 (Group A: with MT, Group B: without MT)

	Group A (with MT)	Group B (without MT)
Mean	1.933333333	1.230769231
Variance	2.20952381	0.525641026
Observation	15	13
Hypothesized Mean Difference	0	
Degree of Freedom	21	
t	1.621484269	
P(T<=t) One-tail	0.059916459	
t Critical One-tail	1.720742903	
P(T<=t) Two-tail	0.119832918	
t Critical Two-tail	2.079613845	

In the writing test, it appears that both Group A and Group B achieved higher scores when MT was not used. Nevertheless, the statistical analysis revealed no significant difference between the two groups.

4.2 Results of Reading Test

Table 5, 6, 7, and 8 provides an overview of the average results of the reading test, comprising both with and without MT, for the participants in the current study. Tables 5 and 6 compare the differences in results by group. Table 5 shows the results of Group A, which used MT in Test 1 and without MT in Test 2. Table 6 shows the results of Group B. Group B solved the test without MT in Test 1 and with MT in Test 2. Tables 7 and 8 compare the differences in results by test, which means that scores are compared when participants solve the same problem with and without MT.

Table 5. Reading test results with and without MT in Group A

	Test 1 (with MT)	Test 2 (without MT)
Mean	27.26666667	23.66666667
Variance	81.92380952	60.52380952
Observation	15	15
Pearson Correlation Coefficient	0.669833691	
Hypothesized Mean Difference	0	
Degree of Freedom	14	
t	2.010069662	
P(T<=t) One-tail	0.032049753	
t Critical One-tail	1.761310136	
P(T<=t) Two-tail	0.064099505	
t Critical Two-tail	2.144786688	

Table 6. Reading test results with and without MT in Group B

	Test 2 (with MT)	Test 1 (without MT)
Mean	22.53846154	20.38461538
Variance	65.76923077	67.58974359
Observation	13	13
Pearson Correlation Coefficient	0.666568553	
Hypothesized Mean Difference	0	
Degree of Freedom	12	
t	1.164478586	
P(T<=t) One-tail	0.133429385	
t Critical One-tail	1.782287556	
P(T<=t) Two-tail	0.26685877	
t Critical Two-tail	2.17881283	

Table 7. Reading test results for Test 1 (Group A: with MT, Group B: without MT)

	Group A (with MT)	Group B (without MT)
Mean	27.26666667	22.53846154
Variance	81.92380952	65.76923077
Observation	15	13
Hypothesized Mean Difference	0	
Degree of Freedom	26	
t	1.457715513	
P(T<=t) One-tail	0.07844642	
t Critical One-tail	1.70561792	
P(T<=t) Two-tail	0.15689284	
t Critical Two-tail	2.055529439	

Table 8. Reading test results for Test 2 (Group A: without MT, Group B: with MT)

	Group A (without MT)	Group B (with MT)
Mean	23.66666667	20.38461538
Variance	60.52380952	67.58974359
Observation	15	13
Hypothesized Mean Difference	0	
Degree of Freedom	25	
t	1.080058629	
P(T<=t) One-tail	0.145213006	
t Critical One-tail	1.708140761	
P(T<=t) Two-tail	0.290426011	
t Critical Two-tail	2.059538553	

In the Reading test, it is noteworthy that the scores in the second test were consistently lower, irrespective of the utilization of MT. Considering this observation, a statistical analysis was conducted to ascertain the significance of the mean difference, with the exclusion of the four participants who did not use MT during the reading test.

Tables 9 and 10 offer a comprehensive breakdown of each participant's scores in both groups, summarizing the variations between scores with and without MT, along with their respective scores for Test 1 and Test 2. Additionally, Tables 11, 12, 13, and 14 delve into a more detailed analysis of the results for each participant.

Table 9. Score results for each participant using MT (Group A)

	Degree of MT use	Test 1 with MT	Test 2 without MT	Test 2 (without MT) - Test 1 (with MT)	with MT - without MT
A1	21	38	36	-2	2
A2	19	37	33	-4	4
A3	4	34	31	-3	3
A4	18	32	28	-4	4
A5	8	23	25	2	-2
A6	14	17	23	6	-6
A7	4	25	21	-4	4
A8	15	20	21	1	-1
A9	8	15	20	5	-5
A10	14	40	18	-22	22
A11	21	26	18	-8	8
A12	9	26	13	-13	13
A13	14	10	8	-2	2

Table 10. Score results for each participant using MT (Group B)

	Degree of MT use	Test 1 without MT	Test 2 with MT	Test 2 (with MT) - Test 1 (without MT)	with MT - without MT
B1	4	32	31	-1	-1
B2	4	32	26	-6	-6
B3	18	29	28	-1	-1
B4	15	28	31	3	3
B5	18	26	17	-9	-9
B6	21	24	9	-15	-15
B7	21	20	28	8	8
B8	10	17	20	3	3
B9	7	13	16	3	3
B10	16	10	5	-5	-5
B11	19	9	14	5	5

Table 11. Reading test results with and without MT in Group A *Exclude those who did not use MT

	Test 1 (with MT)	Test 2 (without MT)
Mean	26.38461538	22.69230769
Variance	88.58974359	62.73076923
Observation	13	13
Pearson Correlation Coefficient	0.640016826	
Hypothesized Mean Difference	0	
Degree of Freedom	12	
t	1.780629721	
P(T<=t) One-tail	0.050140934	
t Critical One-tail	1.782287556	
P(T<=t) Two-tail	0.100281869	
t Critical Two-tail	2.17881283	

Table 12. Reading test results with and without MT in Group B *Exclude those who did not use MT

	Test 2 (with MT)	Test 1 (without MT)
Mean	21.81818182	20.45454545
Variance	72.76363636	81.07272727
Observation	11	11
Pearson Correlation Coefficient	0.699046666	
Hypothesized Mean Difference	0	
Degree of Freedom	10	
t	0.663560933	
P(T<=t) One-tail	0.26098715	
t Critical One-tail	1.812461123	
P(T<=t) Two-tail	0.5219743	
t Critical Two-tail	2.228138852	

Table 13. Reading test results for Test 1 (Group A: with MT, Group B: without MT)

*Exclude those who did not use MT

	Group A (with MT)	Group B (without MT)
Mean	26.38461538	21.81818182
Variance	88.58974359	72.76363636
Observation	13	11
Hypothesized Mean Difference	0	
Degree of Freedom	22	
t	1.246085033	
P(T<=t) One-tail	0.112919912	
t Critical One-tail	1.717144374	
P(T<=t) Two-tail	0.225839824	
t Critical Two-tail	2.073873068	

Table 14. Reading test results for Test 2 (Group A: without MT, Group B: with MT)

*Exclude those who did not use MT

	Group B (with MT)	Group A (without MT)
Mean	22.69230769	20.45454545
Variance	62.73076923	81.07272727
Observation	13	11
Hypothesized Mean Difference	0	
Degree of Freedom	20	
t	0.640782617	
P(T<=t) One-tail	0.264469244	
t Critical One-tail	1.724718243	
P(T<=t) Two-tail	0.528938489	
t Critical Two-tail	2.085963447	

Despite the exclusion of participants who made minimal or no use of MT, the statistical analysis still revealed no significant difference in the mean scores between scenarios with and without MT. Furthermore, participants were requested to specify which reading questions they employed MT for. Table 15 presents the average number of correct answers and the average d egree of MT usage, categorized by each of the reading questions.

Table 15. Average number of correct answers and average frequency of MT use for each reading test (Group A)

	Average number of correct answers for multiple questions	Percentage of correct answers out of 15	Average frequency of MT use
Question P	11.8	79%	3.2
Question Q	10.6	71%	4.2
Question R	8.8	59%	4.3

Table 16. Average number of correct answers and average frequency of MT use for each reading test (Group B)

	Average number of correct answers for multiple questions	Percentage of correct answers out of 13	Average frequency of MT use
Question M	9.4	72%	3.7
Question N	6.6	51%	4
Question O	4.8	37%	4.5

In the case of the final questions (Question R and Question O), the sentences are longer, which naturally increases the level of difficulty. Notably, when we examined MT usage for each question, it became apparent that more MT was used for the more challenging questions.

4.3 Results of Listening Test

Table 17, 18, 19, and 20 provides an overview of the average results of the listening test, comprising both with and without MT, for the participants in the current study. Tables 17 and 18 compare the differences in results by group. Table 17 shows the results of Group A, which solved the test without MT in Test 1 and used MT in Test 2. Table 18 shows the results of Group B. Group B used MT in Test 1 and without MT in Test 2. Tables 19 and 20 compare the differences in results by test, which means that scores are compared when participants solve the same problem with and without MT.

Table 17. Listening test results with and without MT in Group A

	Test 1 (without MT)	Test 2 (with MT)
Mean	65.33333333	68.8
Variance	278.0952381	411.8857143
Observation	15	15
Pearson Correlation Coefficient	0.884513538	
Hypothesized Mean Difference	0	
Degree of Freedom	14	
t	-1.405405405	
P(T<=t) One-tail	0.090851234	
t Critical One-tail	1.761310136	
P(T<=t) Two-tail	0.181702468	
t Critical Two-tail	2.144786688	

Table 18. Listening test results with and without MT in Group B

	Test 1 (with MT)	Test 2 (without MT)
Mean	60.07692308	65.76923077
Variance	132.9102564	416.025641
Observation	13	13
Pearson Correlation Coefficient	0.720553331	
Hypothesized Mean Difference	0	
Degree of Freedom	12	
t	-1.416068043	
P(T<=t) One-tail	0.091092669	
t Critical One-tail	1.782287556	
P(T<=t) Two-tail	0.182185338	
t Critical Two-tail	2.17881283	

Table 19. Listening test results for Test 1 (Group A: without MT, Group B: with MT)

	Group A (without MT)	Group B (with MT)
Mean	65.33333333	60.07692308
Variance	278.0952381	132.9102564
Observation	15	13
Hypothesized Mean Difference	0	
Degree of Freedom	25	
t	0.98009466	
P(T<=t) One-tail	0.168214737	
t Critical One-tail	1.708140761	
P(T<=t) Two-tail	0.336429475	
t Critical Two-tail	2.059538553	

Table 20. Listening test results for Test 2 (Group A: with MT, Group B: without MT)

	Group A (with MT)	Group B (without MT)
Mean	68.8	65.76923077
Variance	411.8857143	416.025641
Observation	15	13
Hypothesized Mean Difference	0	
Degree of Freedom	25	
t	0.393039947	
P(T<=t) One-tail	0.348810791	
t Critical One-tail	1.708140761	
P(T<=t) Two-tail	0.697621581	
t Critical Two-tail	2.059538553	

In the Listening test, participants achieved higher scores in the second test, regardless of whether they utilized MT or not.

4.4 Types of Exams in Which Students Used MT for the Most

Additionally, apart from the aforementioned findings, the study also investigated in which skills - writing, reading and listening - participants would use MT the most if they were permitted to use it. The results were counted in a post-test questionnaire that asked, "To what extent did you use MT in solving this problem?"

Table 21. Mean degree of MT used

	Group A	Group B
Writing	6.1	6.2
Reading	3.9	4.1
Listening	2.2	3

The survey results revelaed that MT is primarily utilized in writing tasks. Conversely, fewer individuals resorted to MT for listening and reading tasks compared to what was initially anticipated.

4.5 Differences in the Degree of Use of MT and Scores by Examinee's English Level

The study also analyzed how the degree of use of MT varies by English proficiency level. In order to group the students by proficiency level, they were divided into the following categories: 80% or higher, 60-79%, 40-59%, 20-39%, and 0-19% for the percentage of correct answers when they answered the questions on their own without using MT. After that, the average efficiency score and average degree of MT use for each group were calculated and are summarized in Table 22. The value in (D) is the sum of the values of the three items in the questionnaire that asked how much MT was used in the reading test.

Table 22. Degree of use and effectiveness of MT by English proficiency level

Percentage of correct answers without MT	Score	**(A)** Number of Applicants	**(B)** with MT- without MT	**(C)=(B)/(A)** Average efficiency score	**(D)** Degree of MT use	**(E)=(D)/(A)** Average Machine Usage
80-100%	32-40 points	4	-1	-0.3	48	12
60-79%	24-31 points	10	-21	-2.1	111	11
40-59%	16-23 points	9	30	3.3	110	12
20-39%	8-15 points	5	18	3.6	65	13
0-19%	0-7 points	0	0	0	0	0

As depicted by the results in (E), there is no noteworthy difference in self-reported MT usage across varying English proficiency levels. In contrast, the value in (C) shows that the average efficiency score, signifying the effectiveness of MT usage, tended to be higher as English proficiency levels decreased.

4.6 Comparison of participants' attitudes toward MT before and after the test

To investigate students' perceptions of MT, a pre-test and post-test questionnaire was administered. The top 10 and bottom 10 scoring items have been extracted and are summarized in Tables 23, 24, 25, and 26. Among these, the following five questions display statistically significant (*) differences in means before and after the responses:

- Using MT eliminates English spelling errors.
- Using MT (from English to Japanese) increases the number of unnatural sentences with direct translations.
- Using MT eliminates English spelling errors.
- Using MT makes it easier to write English.
- If MT becomes widespread, there will be no need for English dictionaries.

Table 23. Top 10 with the highest score (Pre-test)

Group A + B	Pre		Post		Different between pre and post		Test
Questionnaire entries	Mean	Variance	Mean	Variance	Mean	Variance	
(When translating from Japanese to English,) some documents are suitable for machine translation and some are not.	6.1	0.7	6.0	1.0	-0.07	0.31	0.502
(When translationg from English to Japanese,) some documents are suitable for machine translation and some are not.	6.1	0.7	5.9	1.3	-0.18	0.55	0.362
Before using machine translation, one should learn how to use machine translation properly.	5.9	0.8	5.9	1.8	0.07	1.05	0.611
Machine translation helps to get the gist.	5.8	1.3	5.6	1.6	-0.18	0.38	0.443
Using machine translation eliminates English spelling errors.	5.8	1.6	6.6	0.5	0.82	-1.13	**0.021***
Using machine translation makes English assignments easier than when machine translation is not used.	5.7	2.1	5.5	1.5	-0.25	-0.62	0.238
Using machine translation (from English to Japanese) increases the number of unnatural sentences with direct translations.	5.7	1.0	4.9	1.8	-0.75	0.81	**0.023***
Using machine translation helps to read difficult texts, such as English papers.	5.6	2.1	5.8	1.8	0.14	-0.29	0.913
Using machine translation, users can write sentences in English faster than without machine translation.	5.5	2.0	5.6	3.7	0.11	1.70	0.804
In college classes, students become lazy if they are allowed to use machine translation.	5.4	1.2	5.3	1.4	-0.11	0.19	0.638

Table 24. Top 10 with the highest score (Post-test)

Group A + B	Pre		Post		Different between pre and post		Test
Questionnaire entries	Mean	Variance	Mean	Variance	Mean	Variance	
Using machine translation eliminates English spelling errors.	5.8	1.6	6.6	0.5	0.82	-1.13	**0.021***
(When translating from Japanese to English,) some documents are suitable for machine translation and some are not.	6.1	0.7	6.0	1.0	-0.07	0.31	0.502
Before using machine translation, one should learn how to use machine translation properly.	5.9	0.8	5.9	1.8	0.07	1.05	0.611
(When translationg from English to Japanese,) some documents are suitable for machine translation and some are not.	6.1	0.7	5.9	1.3	-0.18	0.55	0.362
Using machine translation makes it easier to write English.	4.8	2.8	5.9	1.6	1.07	-1.24	**0.017***
Using machine translation helps to read difficult texts, such as English papers.	5.6	2.1	5.8	1.8	0.14	-0.29	0.913
Machine translation helps to get the gist.	5.8	1.3	5.6	1.6	-0.18	0.38	0.443
Using machine translation, users can write sentences in English faster than without machine translation.	5.5	2.0	5.6	3.7	0.11	1.70	0.804
Using machine translation helps to read simple English sentences.	4.7	2.6	5.5	2.1	0.82	-0.47	0.089
Using machine translation makes English assignments easier than when machine translation is not used.	5.7	2.1	5.5	1.5	-0.25	-0.62	0.238

Table 25. Worst 10 with lowest scores (Pre-test)

Group A + B	Pre		Post		Different between pre and post		Test
Questionnaire entries	Mean	Variance	Mean	Variance	Mean	Variance	
If machine translation becomes widespread, there will be no need to leam English.	2.2	2.4	2.0	2.0	-0.18	-0.44	0.22
If machine translation becomes widespread, there will be no need for translators.	2.4	2.5	2.2	2.5	-0.14	0.08	0.355
Machine translation should be used from elementary school.	2.5	2.6	2.4	2.0	-0.11	-0.52	0.67
If machine translation becomes widespread, there will be no need for interpreters.	2.7	2.4	2.4	2.9	-0.36	0.47	0.251
Using machine translation makes it easier to remember English grammar.	3.0	2.9	2.9	2.4	-0.07	-0.52	0.798
Using machine translation should be used from junior high school.	3.0	3.1	2.9	2.4	-0.11	-0.71	0.577
Using machine translation makes it easier to remember English words.	3.4	3.4	2.9	2.9	-0.46	-0.55	0.242
Using machine translation makes it easier to listen and comprehend English.	3.4	2.7	3.3	3.0	-0.14	0.25	0.707
If machine translation becomes widespread, there will be no need for English dictionaries.	3.5	4.3	2.8	3.0	-0.71	-1.27	**0.035***
Machine translation should be used from high school.	3.6	2.6	3.9	2.2	0.32	-0.30	0.332

Table 26. Worst 10 with lowest scores (Post-test)

Group A + B	Pre		Post		Different between pre and post		Test
Questionnaire entries	Mean	Variance	Mean	Variance	Mean	Variance	
If machine translation becomes widespread, there will be no need to learn English.	2.2	2.4	2.0	2.0	-0.18	-0.44	0.22
If machine translation becomes widespread, there will be no need for translators.	2.4	2.5	2.2	2.5	-0.14	0.08	0.355
If machine translation becomes widespread, there will be no need for interpreters.	2.7	2.4	2.4	2.9	-0.36	0.47	0.251
Machine translation should be used from elementary school.	2.5	2.6	2.4	2.0	-0.11	-0.52	0.67
If machine translation becomes widespread, there will be no need for English dictionaries.	3.5	4.3	2.8	3.0	-0.71	-1.27	**0.035***
Using machine translation makes it easier to remember English words.	3.4	3.4	2.9	2.9	-0.46	-0.55	0.242
Using machine translation should be used from junior high school.	3.0	3.1	2.9	2.4	-0.11	-0.71	0.577
Using machine translation makes it easier to remember English grammar.	3.0	2.9	2.9	2.4	-0.07	-0.52	0.798
Using machine translation makes it easier to listen and comprehend English.	3.4	2.7	3.3	3.0	-0.14	0.25	0.707
Allowing the use of machine translation in university classes will improve students' English language skills	3.6	3.1	3.3	3.0	-0.29	-0.10	0.507

For the following two items, there is a correlation coefficient in the range of -0.4, indicating a "negative correlation" level with the results of solving Reading without MT:

・Using MT makes it easier to listen and comprehend English.

・If MT becomes widespread, there will be no need to learn English.

5. Discussion and Conclusion

It has become evident that not all students utilize MT even when it is permitted, and their usage of MT varies depending on the task. Furthermore, this study has revealed that there is no statistical advantage to using MT across any of the language skills, whether in writing, reading, or listening. These findings underscore that the utilization of MT does not guarantee a solution to all language-related questions. The survey results indicate that, while most students used MT for writing tasks, it is not as effective for listening. This outcome aligns with Niño's (2020) research, which suggested that MT is less beneficial in language teaching for speaking tasks. MT does not resolve all issues, and the fact that students use MT daily did not mean that they would no longer have problems to solve on their own in learning English. Moreover, it is surprising to note that there was less usage of MT for reading questions than initially anticipated, and the test score results were equally unexpected. Additionally, the study revealed an increase in MT usage as the test questions grew more challenging. This suggests that learners may have perceived it as more efficient to address easier questions on their own, rather than relying on MT. Further detailed analysis unveiled those individuals with higher English proficiency often achieved negative scores when using MT, while those with lower proficiency levels yielded positive scores. From this, we can infer that MT may be more helpful for learners in the intermediate and lower proficiency ranges.

As far as monitoring them, participants often used MT to help them understand in their native language. Even when MT was used for English-to-Japanese translations, it could not always provide the answer. While it may have assisted in the process of converting English into Japanese, this did not necessarily result in correct answers. Errors made when using MT to solve the questions should be compared to those made when people solved the questions on their own, but no comparison was made in the present analysis. The participants in this study were college students who were gathered through an open application process, and while the participants were diverse, the sample size was not large enough. In future analysis, we would like to increase the sample size. In adittion, to isolate the effects of the practice, a control group that would take the test twice without MT should have been included, but could not be done in this study. We would like to focus on qualitative research in addition to quantitative research by collecting qualitative feedback from participants on the use of MT for each question.

References

Cancino, M., & Panes, J. (2021). The impact of Google Translate on L2 writing quality measures: Evidence from Chilean EFL high school learners. *System, 98*. https://doi.org/10.1016/j.system.2021.102464.

Clifford, J., Merschel, L., & Munné, J. (2013). Surveying the landscape: what is the role of MT in language learning?. *@tic. revista d'innovació educativa, 10*, 108–121. https://doi.org/10.7203/attic.10.2228.

Crossley, S, A. (2018). Technological disruption in foreign language teaching: The rise of simultaneous machine translation. *Language Teaching, 51*(4), 541–552. https://doi.org/10.1017/S0261444818000253.

Gally,T., Osaki, S. & Hisamura, K. (2020). *MT ga nihon no eigo kyouiku ni ataeru eikyou* (The impact of MT on English language teaching in Japan) [in Japanese]. *Language Teacher Education: JACETSIG-ELE journal, 7*(1), 1-12. https://www.waseda.jp/assoc-jacetenedu/VOL7NO1.pdf

Jolley, J. R., & Maimone, L. (2015). Free online machine translation: Use and perceptions by Spanish students and instructors.In A. J. Moeller (Ed.), Learn languages, explore cultures, transform lives. Central States Conference on the Teaching of Foreign Languages 2015, 181-200.

Klimova, B., Pikhart, M., Benites, A. D., Lehr, C., & Sanchez-Stockhammer, C. (2022). Neural machine translation in foreign language teaching and learning: a systematic review. *Education and Information Technologies*. https://doi.org/10.1007/s10639-022-11194-2

Kondo, Y., Kimiura, S., Sakaba, H., Toyoshima, C., Nakanan, M., Yamashita, M., & Yamanaka, T. (2023). *Kikai honyaku no eigo seikajyugyou heno daikibo dounyuu to sono kadai*. (Large-scale introduction of machine translation into regular English language classes and its challenges) [in Japanese]. Proceedings of the CIEC Spring Conference 2023, 14, 41-44. https://www.ciec.or.jp/archives/002/202303/CIEC%E6%98%A5%E5%AD%A3%E3%82%AB%E3%83%B3%E3%83%95%E3%82%A1%E3%83%AC%E3%83%B3%E3%82%B92023%E8%AB%96%E6%96%87%E9%9B%86.pdf

Lee, S. M. (2020). The impact of using machine translation on EFL students' writing. *Computer Assisted Language Learning, 3*(3 3), 157-175. https://doi.org/10.1080/09588221.2018.1553186

Niño, A. (2008). Evaluating the use of MT post-editing in the foreign language class. *Computer Assisted Language Learning, 21*(1), 29–49. https://doi.org/10.1080/09588220701865482

Oda, T. (2019). Kikai honyaku to kyouzon suru gaikokugo gakusyuu toha (MT and Foreign Language Education in Japan) [in Japanese]. *The Journal of Humanities and Natural Sciences, 145*, 3-27. https://repository.tku.ac.jp/dspace/bitstream/11150/11398/1/jinbun145-03.pdf

O'Neill, Errol M. (2019). Online translator, dictionary, and search engine use among L2 students. *CALL- EJ: Computer-Assisted Language Learning–Elec- tronic Journal, 20*(1), 154–177.

Stapleton, P., & Kin, B. L. K. (2019). Assessing the accuracy and teachers' impressions of Google Translate: A study of primary L2 writers in Hong Kong. *English for Specific Purposes, 56*, 18–34. https://doi.org/10.1016/j.esp.2019.07.001.

Tsai, S. C. (2020). Chinese students' perceptions of using Google Translate as a translingual CALL tool in EFL writing. *Computer Assisted Language Learning*, 1–23. https://doi.org/10.1080/09588221.2020.1799412.

Yamada, M. (2019). The impact of Google neural machine translation on post-editing by student translators. *The Journal of Specialised Translation, 31*, 84-105. Retrieved August 20, 2021 from https://www.researchgate.net/publication/330831614_The_impact_of_Google_Neural_Machine_Translation_on_Post-editing_by_student_translators

Yamada, M., Langlitz, H., Oda, T., ANGLITZ., Mochida, T., Tamura, H., Hiraoka, Y. & Irie, T. (2021). *Nihon no daigaku niokeru kyouyou eigo kyouiku to kikaihonyaku ni kansuru yobiteki chyousa* (Preliminary study on English language teaching and MT in Japanese universities) [in Japanese]. JAITS (The Japan Association for Interpreting and Translation Studies), 23, 139-156. http://honyakukenkyu.sakura.ne.jp/shotai_vol23/No_23_007-Yamada_et-al.pdf