論文

Japanese university students' reflections on machine translation used as part of an English presentation activity

Kiyo Sakamoto
全学共通教育推進機構

Abstract

The purpose of this study was to conduct a qualitative analysis of the perceptions of native Japanese-speaking university students of English after they used one of the Neural Machine Translation（NMT）tools, DeepL（https://www.deepl.com/）. This qualitative research was to explore how those students perceived NMT（DeepL）after using it to prepare their English manuscript for an English presentation. University students（CEFR A2~B1 level）who participated in the study wrote their manuscripts in Japanese, received instructions on how to use DeepL in class, and prepared their English manuscripts. At the end of the presentation, a free-response questionnaire was administered to the students to find out 1）how comfortable they were with DeepL, 2）what they thought of the English sentences generated by DeepL, 3）how（and to what extent）DeepL helped them to write their presentation, and 4）how（and to what extent）DeepL has helped them in their English study. A qualitative content analysis of the responses revealed that most of the study participants were satisfied with the ease of use of DeepL and highly regarded the generated English sentences for reasons such as "accurate," "clean," "easy to understand," and because "DeepL gave me the English sentences I was looking for." They also considered it a "magic wand," "time-saving tool," and "reliable helper" based on their experience using DeepL for their assignments. Furthermore, they considered DeepL as a "source of knowledge," "what brings discovery and awareness," and "an effective English writing tool," which helped their own English learning. On the other hand, some students found the generated English sentences too difficult, and others thought it was not a good way to learn the language. Based on these results, this paper discusses the implications of the use of NMT for the participants of this study.

1. Introduction

In September 2016, Google announced the adoption of the Google Neural Machine Translation system（Wu et al., 2016）, which became available to the public in November of the same year. Since then, the accuracy and practicality of machine translation（MT）has improved dramatically, and now various translation services using Neural Machine Translation（NMT）are available easily as long as there is an Internet connection and a terminal. Thus, the evolution of translation technology has dramatically lowered the barrier between languages for the general public, who are not language experts.

At the same time, NMT is having a profound impact on the teaching of English as a foreign language（EFL）. As English learners have come to use NMT on a daily basis, English teachers and English curriculum developers have had some of their assumptions and common sense challenged. Prior to the development of NMT, the translations produced by MT often contained unnatural expressions and obvious mistranslations, and learners did not necessarily have a high degree of trust in MT（e.g., Niño, 2009）. However, as the level of writing produced by NMTs has improved, it is natural that more and more English learners rely on them. English teachers now need to reconsider their English class assignments and activities based on the assumption that learners will be using NMT.

Published research on the relationship between MT and foreign language learning and teaching has been conducted since before NMT, and many literature review articles have been published in recent years（e.g., Deng & Yu, 2022; Jolley & Maimone, 2022; Klimova et al., 2022; Lee 2021; Tamura & Yamada, 2021）. However, because of the emergence of NMT, there is also a need to re-examine the accumulation of previous studies. For example, it has been argued that the sentences produced by MT（prior to NMT）can be used as a bad model, and that activities to correct sentence errors are effective for L2 writing learning. However, as Yamada（2019）

indicated, sentences produced by NMT have dramatically improved in grammatical adequacy and vocabulary richness, and they are no longer suitable for use as a bad model. On the other hand, NMT is, in some ways, far from infallible, and it is necessary for NMT users to be able to detect omissions and mistranslations hidden in seemingly flawless and fluent translations. Thus, research has only just begun on how to use NMT for foreign language learning considering its capabilities and pitfalls. As Gally（2019）suggested, it is desirable that machine translation should be used in real classrooms with diverse learners（with attention to the problems of its use）, so that insights into the learning and attitudes of the learners can be gained and shared among EFL teachers and researchers.

According to the literature review articles mentioned above, one of the demands for future research on MT and foreign language learning is to study the characteristics of learners. There are conditions and restrictions on MT use depending on various learner characteristics, such as learner's first language（L1）, English language proficiency level, age, and purpose of learning English. Hence, the exploration of utilization methods for each of them is required. Furthermore, there is still insufficient analysis of how various learners perceive machine translation. Previous learner perception studies on MT use（e.g., Murtisari et al., 2019）have been mostly based on quantitative analysis of the learners' responses to the options provided by the researchers, and less on qualitative analysis of the free responses. However, a deeper understanding of how learners, the parties involved in the learning process, perceive NMT will be essential when considering learning and teaching using NMT. Previous studies have shown that learners evaluate NMT positively or with varying degrees of positivity and negativity, but it would be meaningful to explore the reasons for such evaluations in their own words.

In Japanese universities, the rapid development of NMT and AI has just begun to trigger a rethinking of English education. Japanese university students' motivation to learn English is usually not so high unless they specialize in English-related fields or aspire to work in an international field. It may be partly because Japanese speakers in Japan can receive their higher education in their native language, and most of them live in an environment where there is very little need to communicate in English in their daily and professional lives. Therefore, it will be worth exploring whether the advent of the NMT will further reduce students' motivation to study English, or it can encourage them to adopt a more positive attitude toward learning the language.

This study aims to explore what NMT is for English learners by qualitatively analyzing their responses after they used one of the NMTs, DeepL, for translation from Japanese to English[1], and to consider the role NMT can have in English learning. The significance of this study is that it conducted a qualitative analysis of the reflections of a group of English language learners who are considered to have certain characteristics（such as their L1, English proficiency level, presumed English learning goals, etc.）on their experience of using NMT. Data were collected through an open-ended questionnaire. This is because a quantitative research design, in which participants are asked to choose from a list of options, would be less likely to represent the varied perceptions of individual learners. The researcher chose DeepL as the NMT because of its ability to provide different options for translated sentences produced by the MT, which will be discussed later.

Research Questions

1. If the instructions on how to use NMT（DeepL）are explained in class, and DeepL is allowed to be used to convert Japanese manuscripts to English ones as part of preparation for an English presentation, how would native Japanese speaking university students（approximately CEFR A2~B1 level）evaluate DeepL?

2. Would the above learners be able to use NMT（DeepL）as part of their preparation for an English presentation? And how would they make use of it?

3. Would the above learners find NMT（DeepL）useful for their English learning? And how would they find it useful?

2. Method

Research design overview

The design of this study can be summarized as follows. 111 native Japanese-speaking university students who agreed in writing to participate in the study were asked to complete a presentation assignment, in which they conducted brief research on a topic of their own choice and gave a 3-minute oral presentation in English about their research. The students were allowed to use NMT（DeepL）and were then asked to complete a questionnaire. The questionnaire was in an open-ended format and was filled out in Japanese, the native language of the participants. The questions were: 1）how comfortable they were using DeepL, 2）what they thought about the English sentences generated by DeepL, 3）how（and to what extent）DeepL was useful for their presentations, 4）how（and to what extent）DeepL was useful for their English study. The Ueno method of qualitative analysis（Ueno, 2018）, a simplified version of the KJ method（Kawakita, 1970）, was used to analyze the questionnaire responses. The Ueno method was chosen because it is suitable for the analysis of small-scale surveys and it can reveal the analyst's thought process by returning to individual data after coding and categorizing,

Study participants and the researcher

Recruitment of research participants was made to students in the four classes in which the researcher taught English in the 2021 academic year. Based on a research proposal that had undergone a research ethics review and obtained approval from the university where the researcher works, a written request for research participation was distributed to the students. The students were informed that their participation in the research was completely voluntary, they would not be disadvantaged by not participating, they could withdraw their consent at any time after giving it, they would be anonymized, and their personal information would be handled in strict confidentiality. After that, the researcher asked those who agreed to participate in the study to submit a consent form.

The classes in which the research participants were recruited were required courses for sophomores, with identical course titles and syllabi. 132 students were registered for the four classes in total, of which 111 students participated in the study. Their average score of TOEIC Listening & Reading Test（https://www.iibc-global.org/english.html）was 410, and most of the students appeared to be equivalent to the A2~B1 level of the CEFR（Tannenbaum & Wylie, 2008）.（See Table 1）

Table 1 Number of students in each class and their average TOEIC score

	Number of students per class	Average score of TOEIC
Class A	30	384
Class B	35	440
Class C	34	411
Class D	33	404
total	132	410

The departments to which the students belong（Ecosystem Studies, Biological Resources Management, Regional Studies, Human Relations Studies）are not so-called English majors, and students are generally more interested in their own specialized subjects than in the language. There seems to be a widely shared awareness that students with English language skills will have an advantage when seeking employment or entering graduate school in the future. At the same time, few students seem to study English independently beyond the assignments given in class partly because they are busy with experiments and assignments in their courses, including the ones in their major. Since the university is a public university, many students take the Common Test for University Admissions in order to enter the university, so there is not much difference in English proficiency among the students in the class. In addition, the required English classes for freshmen and sophomores in this university are divided by department, and each department has a capacity of only about 30 to 60 students per year; therefore, by the time students are in their second year, class members are quite close to each other, and many classes have a friendly atmosphere from the beginning.

In 2021, when the study was conducted, there were repeated waves of COVID-19 epidemics, but the

Figure 1: Questionnaire on the use of DeepL

Questionnaire on Internet Tools

Your Name: _____

What are your thoughts on using "DeepL Translation" introduced in class?

About "DeepL translation

☐ I used it ☐ I didn't use it → (Why didn't you use it?: _____)

What did you think about its ease of use?

What did you think of the generated English text?

To what extent (and how) do you think it helped your presentation?

To what extent (and how) do you think it has helped you in your own English studies?

Thank you for your cooperation.

university was conducting face-to-face classes. The textbook used in class（Shizuka, 2015）included a plenty of pair activities in which students discussed their own opinions in English related to the content of the textbook. Therefore, the students became somewhat accustomed to coming up with their own opinions and introducing them to each other in English, which prepared them for giving their presentations in English.

The researcher who taught these classes was a native Japanese speaker who had received primary, secondary, and higher education in Japan, and had spent five years in the United States for graduate study. At the time of this study, she had 25 years of experience as an English teacher at Japanese universities and was in her fifth year of employment at the university where she works. For several years, she was incorporating English presentation activities in her classes. This was motivated by the fact that many Japanese university students, based on her experience at other universities as well, have a weak ability to produce spoken and written English. She hoped that by providing opportunities for individual students to share their unique interests and experiences in English, they would realize the value of being able to communicate using English. On the other hand, she realized that the English manuscripts prepared by some of the students were so linguistically inadequate that they could not be understood by the audience even if they made a presentation. With more than 200 students in her charge per semester, it was difficult for her to make careful corrections for all the students. One of the reasons the researcher considered using NMT for students to prepare English manuscripts for presentations was that she wanted them to practice reading manuscripts written in as good English as possible and to have a successful experience of public speaking.

Data collection procedures

Students were asked to choose one of the topics （Infidelity, Work, Lying, Advertising, and Sports） covered in the reading material of the textbook, set their own theme related to it, conduct brief research, add their own opinions, and give an oral presentation in English. The reason for choosing DeepL over Google Translate was that it presents a variety of expressions as candidates, which can then be selected. After inputting Japanese, the learner can place the cursor on any word in the output English text, and then, multiple options are presented for that part of the sentence. By clicking

on one of them, the sentence is rewritten to match the selected word. In this way, learners can easily select and edit words and sentences by themselves. This point was emphasized in the explanation in class. Each student then prepared an English manuscript at home and gave an oral presentation to the class. After all presentations were completed, the study participants were asked to complete a survey about their use of DeepL. The questionnaire was asked and answered in Japanese（see Figure 1 for its English translation）and was filled out by hand. The researcher converted the results of the questionnaire into data by typing them into an Excel file and translated them into English for this paper.

Data analysis

In response to the first question of the survey about whether they used DeepL or not（if not, why not?）, 107 respondents answered that they used it and four answered that they did not. Of the four who did not use it, three respondents answered, "I am used to using Google Translate" as the reason for not using it, and one said "I thought I could write English myself." Subsequent questions were asked only of the 107 respondents who answered that they had used DeepL.

In the Ueno method of qualitative analysis, observational data and summary data from audio recordings are decontextualized and recontextualized through the following process.

1. Data unitization: Observation data and summary data are transcribed as information units（on a discourse basis）to cards. They are called "primary information."
2. Grouping of units: All cards are grouped from the perspective of being "the same as or different from" others.
3. Categorization: The "commonalities" of the cards collected in one group are verbalized and a "nameplate" for each group is given（creating metainformation）.
4. Mapping: Again, "same or different" judgment is made among the nameplates produced above and the "same" ones are mapped near each other and the "different" ones far away from each other. In doing so, all cards of primary information are expanded

around their nameplates to confirm the validity of the relationship.
5. Charting: A chart is created by connecting "commonalities" or nameplates（=meta-information）with each other using either "causality," "conflict," or "correlation". If there is any meta-information that cannot be connected to others, it should be left isolated.
6. Storytelling: A "case report" is created by converting the chart into text（storytelling）. In doing so, all meta-information is used at least once and primary information is quoted when necessary.
7. Discussion: The case report is reviewed and discussed with concepts and contextual information that may not be manifested in the report.

The data obtained in this study were not the content of audio recordings of interviews with a single study participant, but rather the responses to questions in a questionnaire administered to 107 study participants. Separate case reports were created for each of the four questions. In converting the data into units, all the questionnaire responses were intercepted in meaningful sentence units and used as primary information.（For example, "It was easy to use because a translation appeared as soon as I typed in Japanese. Also, when I clicked on a translation, other suggestions appeared" was divided into two parts: "It was easy to use because a translation appeared as soon as I typed in Japanese" and "When I clicked on a translation, other suggestions appeared."）

The current study did not seek feedback from study participants or other researchers on the content of the case reports.

3. Findings

The four case reports, which were produced by analyzing responses to the four questions in the survey, are shown below. Each report is a summary and recontextualization of the responses that emerged within the group of 107 study participants and shows how NMT（DeepL）was perceived by the study participants.

Case Report 1

Answer to Question 1 "What did you think about DeepL's ease of use?"

DeepL is "easy" and "accurate" and its sentences can be "rephrased" to create "expressions that suit me," but it also has a few "inconvenient" parts.

Percentage of positive and negative evaluations

102 of the 107（95.3%）respondents gave a positive evaluation of its ease of use, stating it was "very good," "very convenient," and "very easy to use". On the other hand, five（4.7%）respondents were negative in their evaluation. The responses were categorized into five（four positive and one negative）commonalities（1-1~1-5）based on the content.

1-1 It was "effortless"

Research participants frankly evaluated the tool as being easy and convenient for them, with expressions such as "I just need to type in," "short waiting time was nice," "can be translated in an instant," and "it was effortless because all I had to do was put in the sentence as it was."

1-2 English sentences generated are "accurate"

They highly evaluated the quality of the English sentences generated by DeepL, saying, "I was impressed that even long sentences did not end up sounding very strange," "the English sentences were natural," and "accurate." In particular, they felt that it was different from machine translation they had used in the past, saying, "unlike google translate, it was much more accurate, which was good," and "unlike other translation software, I was surprised that it was translated exactly as it was meant to be."

1-3 There is an option to "rephrase" the expression in another way.

The ability to re-select words in the generated English sentences was highly appreciated. "The ability to click on a word to change it to another word was very good," "it was good that it gives me a lot of suggestions," and "the function to change a difficult word to another word was very useful."

1-4 As a result, I can choose the "expression that suits me."

They appreciated not only being able to choose, but also to find the expression that "fits" them. "The ability to change the sentence structure and words so that it is closer to what I know is also good," "I was able to fit in phrases and words that did not feel right about the English sentences that came up," "it was easy to use because it came up with many candidates, so I could create sentences that suited my English ability," and "the ability to change the sentence to something I wanted was good."

1-5 There are some "inconveniences".

Some pointed out the inadequacy of the functions, saying, "it was a little inconvenient that a bug occurred when I put in a long sentence all at once," "I felt it necessary to check some of the mistakes," and "it was a little difficult to use the smartphone version because alternative candidates were not available."

Case Report 2

Answers to Question 2 "What did you think of the generated English text?"

Most respondents thought that the English sentences were "accurate, natural, and clean," "easy to understand," "what I was looking for," and "useful for learning." Meanwhile, others said, "it was difficult for me" and "it was inconvenient to have to check."

Percentage of positive and negative evaluations

Of the 107 respondents who used DeepL, 72（67.3%）gave positive evaluations, 16（15.0%）were positive overall while mentioning negative factors, and 19（17.8%）were negative. The responses were categorized into seven（four positive and three negative）commonalities（2-1~2-7）based on the content.

2-1 "Accurate, natural, and clean"

The quality of the English sentences produced by

DeepL were highly evaluated. For example, "I thought they were accurate," "it was good that the sentences came out naturally," "I could produce a sentence in one shot that didn't feel out of place," and "the translation was cleaner [in Japanese *kirei*, which can mean beautiful, neat, refined, etc.] than my own."

2-2 "Easy-to-understand"

The sentences produced first and/or the sentences obtained by rephrasing were also evaluated highly because they were easy for the respondents to understand. "It was good because I could convert to words I could understand," "it was good because there were expressions I have used before," "there were many simple words so I could understand the sentences myself," etc. Some responses also referred to the ease of understanding for the audience, such as, "I think I was able to create English sentences that are easy for the audience to understand" and "I think it was not that difficult for the audience to understand my presentation because I was able to change difficult words for easy ones from the candidate list."

2-3 English sentences "I was looking for"

The respondents appreciated that NMT eliminates the frustration of wanting to say something but not being able to produce it. One wrote: "It was wonderful that the expression I wanted to use came out exactly as I wanted." They also highly evaluated the ability to find the English sentence they were looking for by using the rephrasing function. For example, "it was nice to be able to change the phrases to whatever I wanted," "if there was an expression I didn't like or a word I didn't understand, it was easy to use because I could replace it with another expression."

2-4 English sentences "useful for learning"

On the other hand, some respondents saw it as a "learning experience" even when unknown expressions were shown: "I have never seen this usage before in some sentences, but when I looked it up, there was no mistake and I learned a lot," "I thought it was very good that DeepL presented new expressions such as different subjects and verbs from the English sentence I had

expected," and "there were words and phrases used that I did not know and I was able to use them while thinking about whether to use them as they are or to replace them."

2-5 English sentences "difficult for me"

At the same time, some learners found the generated English sentences difficult, saying, "there were many sentence structures and words that I do not usually use," "I felt there were many difficult phrases," and "I wanted to make the sentences more understandable and shorter."

2-6 English sentences "need to be confirmed"

Some respondents warned against using the translation as it is because of the problems with the generated English text. "Once, the subject was translated incorrectly," "it is necessary to check the full text by yourself," "sometimes it looked like a bug," and "there were some strange sentences or omissions in places, so I thought it was necessary to check the text."

2-7 English sentences "hard to remember"

Regarding the impact on presentation, one respondent said, "It was hard to remember the generated sentences since I didn't make them from scratch by myself."

Case Report 3
Answers to Question 3 "To what extent (and how) do you think DeepL helped your presentation?"

Respondents regarded DeepL as a "magic wand," "time-saving tool," and "reliable helper," but some people wanted more.

Percentage of positive and negative evaluations:

Of the 107 respondents who used DeepL, 101 (94.4%) found it "useful". Meanwhile, six respondents (5.6%) answered either "not useful," "useful but with reservations," or no response. The answers that described the reasons for each evaluation were categorized into four (three positive, one negative) commonalities (3-

1~3-4）based on the content.

3-1: DeepL as a "magic wand"

The respondents viewed DeepL as a "magic wand" that took care of the translation work as represented by the responses such as "I relied on DeepL for almost all of the translation," "I owe［my presentation］almost entirely to DeepL," and " I couldn't have done it without it（=DeepL）."

3-2: DeepL as a "time-saving tool"

Many of the respondents said, "it saved me a lot of time" or "it saved me a lot of trouble." One respondent wrote, "I found it useful when I wanted to write my own sentences but didn't have the time.（It helped me a lot）." Some respondents also indicated that they were able to spend more time on other aspects of their studies thanks to the shorter time necessary for translation. One respondent said, "It was very helpful because I spent less time drafting and more time practicing for my presentation, such as slides and practice."

3-3 DeepL as a "reliable helper"

Some of the responses indicated that NMT helped them with things that would have been difficult for them to do on their own in preparing their presentation drafts. For example, "it helped me to write grammatically correct sentences because I wasn't confident in my grammar," "it helped me to see alternative translations that I could have written myself but couldn't think of, " "I was able to come up with expressions that I wouldn't have thought of on my own, which helped me to communicate more clearly," and "it helped me to rephrase words that were difficult to understand."

3-4 "This is not enough."

Negative responses, where the reason was clearly stated, were : "I felt it was not very useful because I had to rework the generated text later" and "it helped me to write literary expressions that were difficult for me to conceive of on my own. However, the presentation is spoken, so I felt the need to re-translate it myself". They both seemed to have the ability to further modify the text generated by DeepL by themselves to get closer to what

they wanted to express.

Case Report 4
Answers to Question 4 "To what extent（and how）do you think DeepL has helped you in your own English studies?"

DeepL has helped their English learning by providing "knowledge," "discovery," and "learning" and also as a "usable tool"; but some think it also requires effort on their part to make it useful for their study.

Percentage of positive and negative evaluations

Of the 107 respondents who used DeepL, 96（89.7%）said it helped them improve their English in some way, and 11（10.3%）said it did not help them. The responses to the reasons for each evaluation were categorized into seven（five positive, two negative）commonalities（4-1~4-7）based on the content.

4-1 I have gained "knowledge"

Some respondents said they gained new knowledge on English expressions: "It was a good opportunity to learn new expressions I did not know," "I learned how to form English sentences correctly," "I learned various expressions because of the synonyms," "it was very useful to learn phrases," "using new words is important, but I now understand how I can use the words I already knew," and so on.

4-2 It has given me "noticing," "discovery," and "surprise"

In addition, some participants were not merely acquiring knowledge but were pleasantly surprised to discover expressions that were different from what they had expected: "I felt that there is more than one way to express in English," "I discovered a new way to use English grammar," "I got often surprised, like, 'you use this expression?'," and "I learned and felt like, 'Wow, that's how I can express it!'"

4-3 I have learned from "checking"

Some respondents felt that checking the NMT-produced English sentences against their own Japanese

sentences was instructive. For example, "While checking, I was able to review things like, 'Surely I learned this kind of grammar before,'" and " I did not trust the English text produced by DeepL and took a moment to look it over myself. Then, I was able to find inappropriate parts in the Japanese text I typed in," "I think I became able to find strange sentences because I had to review them with DeepL," and "I don't think I gained the ability to translate Japanese into English, but I realized that the ability to translate English into Japanese is necessary in order to check the text produced by DeepL."

4-4 I have learned from "rephrasing"

Other respondents said that the rephrasing function of DeepL was useful for studying English. For example, "when I felt uncomfortable with a word, it can be changed into another word I knew, so I could use my own knowledge, which was quite useful," and "it was useful in that I could think about which syntax I should use for a sentence that appeared in English translation while looking at my own Japanese sentence."

4-5 It's useful as an "English writing tool"

Some respondents indicated that the NMT was useful for their studies in the sense that they obtained a tool that they could use in producing English. Some of the responses included, "it helped me to use expressions that I wanted to use but did not know how to use," "when writing English, I was able to make sentences more communicative," "it was great to know that I could see various ways of saying things more accurately!," "When giving a presentation, it was useful to be able to adapt my writing to my level of ability."

4-6 It's "not useful" for my study

On the other hand, some responded negatively about the NMT's contribution to English learning. "To be honest, I don't think it has helped me much in my English study," "I don't think it has helped me much, because the translation is done automatically after input," "DeepL is better than me at expressing sentences, so it did not help me much," and "I have been doing English translations on my own, but I don't know which is better for my English study." Some responses mentioned the danger of

stopping writing in English, such as "I'm afraid I would rely on DeepL translation before I think up my own English text from now on," and "it is a little too helpful because I depend on it."

4-7 I "didn't make the effort" to study

Some reflected that they were aware of DeepL's potential to contribute to English learning, but failed to take advantage of it. One said, "I was able to learn words and grammar I didn't know, but I didn't check them properly, so I didn't fully master them." Another wrote, "it helped me a little.（I picked easy-to-understand English sentences, so I didn't study much.）"

4. Discussion
Answers to Research Questions

Based on the four case reports above, answers to Research Questions are summarized as follows:

1. Most of the learners who participated in this study gave a positive evaluation of DeepL's ease of use, when they used it to convert Japanese manuscripts to English ones as part of their preparation for English presentation. They also highly regarded the English text generated by DeepL, since it was "good quality," "easy to understand," "what I was looking for," and so on.

2. Many of the above learners used DeepL as a "magic wand," "time-saving tool," and "reliable helper" as part of their preparation for English presentations.

3. Many of the above learners found DeepL useful for their English learning as a "source of knowledge," "something to provide discovery or notice in the process of checking and rephrasing," and "an effective English writing tool."

From here, this section will discuss what NMT means to the participants in this study from six perspectives and consider the possible uses of NMT for their English learning.

⑴ Learners' high appreciation of NMT and risk of dependence

First, for most of the EFL learners who participated in this study, DeepL was a "helpful" presence. This can

be confirmed by categories such as "it was effortless" （case report section 1-1）, "a magic wand" （3-1）, "a time-saving tool" （3-2）, and "a reliable helper" （3-3）. The NMT was seen, above all, as a useful and appreciated resource. Of course, if the use of a convenient tool becomes blind dependance, it can be a double-edged sword; and some students seemed to be aware of it.

⑵ Learners' wishes to find an expression that is satisfactory to themselves

At the same time, the study participants are not evaluating DeepL simply because they want to make things easier. This is because they are describing the English sentences produced with the help of DeepL using the words and phrases such as "accurate" （1-2 and 2-1）, "natural," （2-1）, "clean" （2-1）, "easy to understand" （2-2）, and "what I wanted" （2-3）. In other words, they have their own values and aesthetic sense of desirable English expressions, and they seem to aspire to create an English text that reflects such sense of worth and beauty and is satisfactory to themselves. Needless to say, learners should strive and teachers should support them so that they can produce such sentences on their own. However, as Briggs （2018） and Rushton （2022） suggest, the experience of being able to create a satisfactory English text and communicate using it, even if it is through the scaffold of MT, may bring about the possibility of a more positive willingness to communicate in EFL learners.

⑶ Learners' wishes to find an expression that suits them

Research participants also highly evaluated the function to rephrase expressions in DeepL （1-3）, citing, in part, the ability to choose the "expressions that suit me" （1-4） and to produce English sentences that "I was looking for" （2-3）. This suggests that learners have a desire to express themselves in a way that suits them even when they are using a foreign language. It may be important for EFL teaching and learning reseach to explore in more detail what they mean by the expressions such as "suiting me," why it is important, and what kind of learning and experience will enable them to create such expressions for themselves.

⑷ Possibilities of NMT contributing to better English learning for EFL learners

Nearly 90% of the study participants thought they were able to "learn English" from the use of DeepL. When asked about their reflections on the English sentences generated, they already stated that the English sentences that came out were "useful for learning" （2-4）. Then, when asked "to what extent （how） do you think DeepL was useful for your English study?", they answered, "I gained knowledge" （4-1）, "it has given me 'noticing,' 'discovery' and 'surprise'" （4-2）, "I have learned from checking" （4-3）, and "I have learned from rephrasing" （4-4）.

What they have gained may be summarized as "new knowledge," "review opportunity and extension of knowledge they already had," and "awareness, discovery, and surprise." The third one seems particularly insightful. Perhaps the reason why the learning brought such exciting emotions to the learners was because the English expressions were derived from the learners' input of the Japanese manuscripts they wrote, and, in such a sense, their "own" words.

Also, NMT may be able to provide a good model of what the learners want to express without pointing out their mistakes or reminding them of their inadequacies. As Ene & Yao （2021） discussed, the affective aspect of the learner in receiving feedback on L2 writing is of great importance. It may be worth exploring the effectiveness of NMT for foreign language writing study from the learners' affective aspects as well.

⑸ Considerations from learners' negative opinions on the introduction of NMT into language teaching

The participants that had negative reactions to DeepL can be divided into three groups. The first one felt that the English sentences generated by DeepL were "difficult for me" （2-5）. Perhaps their command of English was not sufficient to use machine translation, or they did not fully understand how to use DeepL. The second group was able to use DeepL effectively and pointed out the limitations of this tool. In addition to the technical problems of NMT （1-5） and the need to check the generated English sentences （1-5 and 2-6）, they also

noticed that there are things that DeepL alone cannot do. They pointed out "it was hard to remember the generated sentences" (2-7) and DeepL was "not enough" for presentation preparation (3-4). As Klimova et al. (2022) noted, advanced learners tend to consider the output of NMT texts more critically than beginner and intermediate learners; such differences in proficiency may bring out the difference between the first group and the second group. A third group of study participants, about 10% of the total, responded that DeepL was not useful for their own English study. They seem to understand that being able to use DeepL is different from being able to write by themselves (4-6), and that it takes effort on their part to study the language (4-7). It would be essential for teachers with students of varying levels of English language proficiency and motivation to understand the readiness of each individual learner and to continue careful orientation if they choose to introduce NMT in their classes.

(6) Exploration of appropriate learning activities

Finally, it may be worth mentioning the significance of the in-class English presentation activity that served as the background for this study. Oda (2021), in advocating collaboration between English language education and machine translation, cited English presentations as an example of a worthwhile activity while using machine translation. The fact that the learners in this study would eventually share their presentation with their classmates may have influenced their approach to this translation activity. In the category of 2-2 in Case Report 2, two respondents appreciated the rephrasing function of DeepL because they could produce English sentences that were easy to understand for the audience. In other words, they tried to revise their manuscripts while keeping their audience in mind. Also, the learners' interest in each other's presentations may have inspired them to make better presentations with better English writing.

After analyzing various syntactic and semantic problems that may arise in texts produced by NMT, Yanase and Lees (2022) warned that NMT is by no means a "magic wand." As they emphasized, careful re-reading of the original and translated texts side by side is essential when using NMT. It will be even more important, then, for the learner to be motivated to write as good a text as possible. In this sense, writing education using machine translation will need to further explore activities and conditions that make learners want to write better.

One of the limitations of this study is that it ended when the researcher analyzed the responses to the questionnaire, and as a result, there is not enough scrutiny of the analysis through feedback from the study participants. This is an issue to be improved in the future research.

5. Conclusion

This study qualitatively analyzed the reflections of university EFL leaners (Japanese native speakers with CEFR A2~B1 level command of English) on their use of DeepL for the preparation of manuscripts for their English presentations in class. The results provide a multifaceted view of how NMT was perceived by the participants in this study and suggest what possible roles NMT could have in their future English language learning.

Research has only just begun to investigate what kinds of learners, in what kinds of settings, and in what ways shoud use NMT for language learning. At the same time, exploring how NMT can be used in English language learning will reveal what NMT cannot do. For example, in her study of L2 writing instruction using MT, Lee (2020) pointed out that although MT can help learners correct vocabulary and grammar errors much better than human classmates can, peer feedback is superior to MT for corrections beyond a single sentence. When it comes to writing, NMT alone cannot foster learners' engagement in thought and discussion, scrutiny of logic and organization to be better understood by the reader, or enhancement of motivation to express oneself and communicate with others. It is hoped that proper use of NMT can reduce some burden of learning a foreign language and will lead learners to enjoy and engage more in human activities made possible by that foreign language.

A part of this paper was orally presented at the 50th National Convention of The Japan Society for English Usage and Style（held online on December 18, 2021）.

Notes

1）As Yamada（2019）mentioned, in using translation technology for foreign language theaching/learning, it is important to understand the differences between receptive language use and productive language use. This study is concerned with productive language use, that is, the use of NMT to generate L2 texts, while receptive language use（e. g., translation from L2 to L1）may involve completely different issues.

References

Briggs, N.（2018）. Neural machine translation tools in the language learning classroom: students' use, perceptions, and analyses. *The JALT CALL Journal, 14*（1）, 3-24. http://dx.doi.org/10.29140/jaltcall.v14n1.221

Deng, X., & Yu, Z.（2022）. A systematic review of machine-translation-assisted language learning for sustainable education. *Sustainability: Science Practice and Policy, 14*（13）, 7598. http://dx.doi.org/10.3390/su14137598

Ene, E., & Yao, J.（2021）. How does that make you feel: Students' affective engagement with feedback. *Language Teaching Research Quarterly, 25,* 66-83. http://dx.doi.org/10.32038/ltrq.2021.25.04

Gally, T.（2019）. The implication of machine translation for English education in Japan. *Language Teacher Education, 6*（2）, 1-14. http://www.waseda.jp/assoc-jacetenedu/VOL6NO2.pdf

Jolley, J. R., & Maimone, L.（2022）. Thirty years of machine translation in language teaching and learning: A review of the literature. *L2 Journal, 14*（1）. https://doi.org/10.5070/L214151760

Kawakita, J.（1970）. *Zoku hassou hou - KJ-hou no tenkai to ohyo* [The KJ method: Its development and application]. Chuokoron-sha.

Klimova, B., Pikhart, M., Benites, A.D. et al.（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

Lee, S.-M.（2020）. The impact of using machine translation on EFL students' writing. *Computer Assisted Language Learning, 33*（3）, 157-175. https://doi.org/10.1080/09588221.2018.1553186

Lee, S-M.（2021）. The effectiveness of machine translation in foreign language education: A systematic review and meta-analysis. *Computer Assisted Language Learning,* http://dx.doi.org/10.1080/09588221.2021.1901745

Murtisari, E. T., Widiningrum, R., Branata, J., & Susanto, R. D.（2019）. Google Translate in language learning: Indonesian EFL students' attitudes. *The Journal of Asia TEFL. 16*（3）, 978-986. https://doi.org/10.18823/asiatefl.2019.16.3.14.978

Murtisari, Widiningrum, & Branata.（2019）. Google Translate in language learning: Indonesian EFL students' attitudes. *The Journal of Asia TEFL., 16*（3）, 978-986. DOI:10.18823/asiatefl.2019.16.3.14.978

Niño, A.（2009）. Machine translation in foreign language learning: language learners' and tutors' perceptions of its advantages and disadvantages. *ReCALL, 21*（2）, 241-258. http://dx.doi.org/10.1017/S0958344009000172

Oda, T.（2021）. The influence of machine translation on general education English in Japan. *The Journal of Humanities and Natural Sciences, 149,* 3-27. https://repository.tku.ac.jp/dspace/handle/11150/11672

Rushton, A.（2022）. Motivating Japanese university EFL learners to produce longer speaking turns with web based machine translation: A pilot study. *Kobe Kaisei Jogakuin University Research Bulletin, 66,* 73-81. https://kaisei.repo.nii.ac.jp/?action=repository_action_common_download&item_id=249&item_no=1&attribute_id=22&file_no=1

Shizuka, T.（2015）. *Burning Issues - Intermediate Level.* Shohakusha.

Tamura, H. & Yamada, M.（2021）. A survey on the use of machine translation in foreign language education: A review of previous studies. *MITIS Journal, 2*（1）, 55-66. http://jaits.web.fc2.com/Yamada_Tamura_3.pdf

Tannenbaum, R. J., & Wylie, E. C.（2008）. Linking

English-language test scores onto the Common European Framework of Reference: An application of the standard-setting methodology. ETS, Princeton, NJ. http://dx.doi.org/10.1002/j.2333-8504.2008.tb02120.x

Ueno, C.（2018）. *Joho seisansh ni naru.* [To Become an Information Producer]. Chikuma Shobo.

Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K., Klingner, J., Shah, A., Johnson, M., Liu, X., Kaiser, Ł., Gouws, S., Kato, Y., Kudo, T., Kazawa, H., ... Dean, J.（2016）. Google's neural machine translation system: Bridging the gap between human and machine translation. In arXiv [cs.CL]. https://doi.org/10.48550/arXiv.1609.08144

Yamada, M.（2019）. Language learners and non-professional translators as users. in M. O'Hagan（Ed.）, *The Routledge Handbook of Translation and Technology*（pp. 183-199）. Routledge. http://dx.doi.org/10.4324/9781315311258-11/language-learners-non-professional-translators-users-masaru-yamada

Yanase, Y. & Lees, D.（2022）. Categorizing errors in machine-translated academic essays from Japanese（L1）to English（L2）: Some specific findings and general implications from Kyoto University EGAP writing classes. *Bulletin of the Faculty of International Advanced Studies,* Kyoto University, *5,* 59-79. http://dx.doi.org/10.14989/ILAS_5_59