



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

교육학석사학위논문

Effects of Machine Translation Instruction on English Writing Revision of Korean Middle School Students

번역기 활용 수업이 한국 중학생들의
영어 글쓰기 수정에 미치는 영향

2022년 2월

서울대학교 대학원

외국어교육과 영어전공

이 지 현

Effects of Machine Translation Instruction on English Writing Revision of Korean Middle School Students

by
Jihyun Lee

A Thesis Submitted to
the Department of Foreign Language Education
in Partial Fulfillment of the Requirements
for the Degree of Master of Arts in Education

At the
Graduate School of Seoul National University

February 2022

Effects of Machine Translation Instruction on English Writing Revision of Korean Middle School Students

번역기 활용 수업이 한국 중학생들의
영어 글쓰기 수정에 미치는 영향

지도교수 오 선 영

이 논문을 교육학 석사 학위논문으로 제출함
2022년 2월

서울대학교 대학원
외국어교육과 영어전공
이 지 현

이지현의 석사학위논문을 인준함
2022년 2월

위 원 장 _____ (인)

부위원장 _____ (인)

위 원 _____ (인)

Effects of Machine Translation Instruction on English Writing Revision of Korean Middle School Students

APPROVED BY THESIS COMMITTEE:

JIN-WAN KIM, COMMITTEE CHAIR

YOUNGSOON SO

SUN-YOUNG OH

ABSTRACT

Based on the dramatic technological improvement on machine translation (MT), a consensus has been established on the necessity of the research on the educational use of the technology. Still, little is known about its use, especially in a secondary school classroom context. Therefore, this study attempts to expand the findings of the previous studies by investigating Korean middle school English learners' MT use. Moreover, since it is more of importance to provide adequate MT instruction for younger students in order to train them to use this tool pedagogically, this study aims to investigate the effects of MT instruction on Korean middle school students.

This study was conducted at a middle school in Busan, and a total of 42 students from two intact classes participated in a picture-based narrative writing task and a revision task. After the writing task, Class 1 (hereafter, Instruction Group) received two periods of MT instruction before conducting MT-assisted revision. MT instruction includes explanation on MT, the researcher's demonstration on using MT outputs as a reference source to revise texts, and a few focused exercises to practice MT-assisted revision. On the other hand, Class 2 (hereafter, No-instruction Group) was not provided with such instruction, rather only provided with the opportunities to be exposed to use MT freely.

To examine the effects of MT instruction, writing qualities of the

participants' writing products, revision behaviors of the participants, and students' perception toward MT were analyzed and compared via either Wilcoxon signed-rank test or Mann-Whitney U test.

The results showed that spelling and grammar errors significantly decreased in case of Instruction Group, while only grammar errors did so in case of No-instruction Group. Regarding syntactic complexity, Instruction Group showed no change in mean length of T-unit, clauses per T-unit, and mean length of clauses. In contrast, No-instruction Group showed increases in mean length of T-unit and mean length of clauses, but there were differences in quality improvement depending on individuals. Regarding vocabulary use, there was a significant increase in the use of academic words in case of Instruction Group while No-instruction group showed no change in word use. These differences suggest that MT instruction was effective to train middle school learners to use MT to reduce errors and use more academic words. However, it might have an adverse effect that limits students from attempting more sophisticated sentence structures.

Next, the results on revision behavior showed that Instruction Group demonstrated more frequent revisions in all categories compared to No-instruction Group at a significant level. In addition to the high frequencies of revising occurrences, another notable finding is that, unlike No-instruction Group's revision, which was limited to a particular type of size (i.e., clause), purpose (i.e., informational), action (i.e., addition), Instruction Group conducted more various revisions with diverse units (i.e., letter, word,

phrase, and clause), purposes (mechanic, lexical, grammatical, cosmetic, and informational), and actions (addition, deletion, substitution, and permutation). These results showed that MT instruction was successful in training the participants to conduct revisions more actively.

Regarding the users' perceptions, Instruction Group participants responded much more positively to the questions that examine MT's usefulness, satisfaction on MT, and intention to use MT in the future than No-instruction Group participants did. The significantly higher scores given by Instruction Group participants showed that they perceived MT and MT-assisted revision more positively than their counterparts.

Despite several limitations, this study contributes to an understanding of the effects of MT instruction for English writing revision; MT can be used as a revision tool to improve foreign language writing process even of young beginner students as long as appropriate instruction is provided. Pedagogical implications and suggestions for future research are also identified.

Keyword : machine translation, Google Translate, instruction, L2 writing, L2 revision, computer-assisted language learning

Student Number : 2020-28070

TABLE OF CONTENTS

ABSTRACT	i
TABLE OF CONTENTS	iv
LIST OF TABLES.....	vi
LIST OF FIGURES.....	vii
CHAPTER 1. INTRODUCTION	1
1.1. Background of the Study	1
1.2. Purpose of the Study	6
1.3. Research Questions.....	7
1.4. Organization of the Thesis	8
CHAPTER 2. LITERATURE REVIEW	9
2.1. Machine Translation for Foreign Language Writing	9
2.2. Use of Machine Translation for Revision.....	13
2.3. Machine Translation Instruction	16
2.4. Students' Perception toward Machine Translation	19
2.5. Research Gap	20
CHAPTER 3. METHODOLOGY	21
3.1. Participants	21
3.2. Instruments	23
3.2.1. Picture-based narrative task	23
3.2.2. Questionnaires	25
3.2.3. Google Translate.....	26
3.3. Machine Translation Instruction	27
3.3.1. Objective.....	27
3.3.2. Components	28
3.3.3. Sequence	30
3.4. Procedures	35
3.4.1. Pilot Study	35
3.4.2. Main Study.....	39

3.5.	Data Analysis	4	1
3.5.1.	The analysis of the participants' L2 writing qualities	4	2
3.5.2.	The analysis of the participants' revision behavior.....	4	4
3.5.3.	The analysis of the participants' perception toward MT and its use	4	7
3.6.	Preliminary Analysis.....	4	8
CHAPTER 4. RESULTS AND DISCUSSION		5	2
4.1.	Effects of MT Instruction on Revised Text Qualities	5	2
4.1.1.	Accuracy	5	2
4.1.2.	Syntactic Complexity.....	6	2
4.1.3.	Vocabulary Use	6	6
4.2.	Effects of MT Instruction on Revision Behavior	7	0
4.2.1.	Revision Size	7	1
4.2.2.	Revision Purpose	7	2
4.2.3.	Revision Action Type	7	8
4.3.	Effects of MT Instruction on Students' Perception toward MT	8	1
CHAPTER 5. CONCLUSION		8	5
5.1.	Major Findings and Pedagogical Implications	8	5
5.2.	Limitations and Suggestions for Future Research.....	8	9
REFERENCES		9	1
APPENDICES		1	0 0
국 문 초 록		1	1 0

LIST OF TABLES

Table 3.1. Instructional Components and Procedures	3	5
Table 3.2. Summary of Experiment Procedure	4	1
Table 3.3. Revision Taxonomy with Examples	4	5
Table 3.4. Qualities of Original Writings Produced by IG and NG	4	9
Table 3.5. Answers from the Participants on Prior Exposures to MT	5	0
Table 3.6. Answers from the Participants on Prior MT Use	5	0
Table 3.7. Scores on Perception toward MT	5	1
Table 4.1. Comparison of Number of Words in Original and Revised Writings by IG Participants	5	3
Table 4.2. Comparison of Number of Words in Original and Revised Writings by NG Participants.....	5	3
Table 4.3. Comparison of Number of Errors in Original and Revised Writings by IG Participants	5	5
Table 4.4. Comparison of Number of Errors in Original and Revised Writings by NG Participants.....	5	5
Table 4.5. Comparison of Complexity Measures in Original and Revised Writings by IG Participants.....	6	2
Table 4.6. Comparison of Complexity Measures in Original and Revised Writings by NG Participants	6	3
Table 4.7. Comparison of Vocabulary Profiles in Original and Revised Writings by IG Participants.....	6	7
Table 4.8. Comparison of Vocabulary Profiles in Original and Revised Writings by NG Participants	6	7
Table 4.9. Comparison of IG and NG: Frequencies of Revisions According to Linguistic Sizes.....	7	1
Table 4.10. Comparison of IG and NG: Frequencies of Revisions According to Revisional Purposes	7	3
Table 4.11. Comparison of IG and NG: Frequencies of Revisions According to Revisional Actions	7	9
Table 4.12. Post Questionnaire Results	8	1

LIST OF FIGURES

Figure 3.1. An Example of Mistranslation of GT	2	9
Figure 3.2. An Example of Inverted Translation for Comparison	3	1
Figure 3.3. Inverted Translation to Discover Different Expression	3	2
Figure 3.4. An Example of Pre-editing	3	2
Figure 3.5. A Focused Exercise to Practice MT-assisted Revision	3	3
Figure 3.6. Padlet Sharing During the Group Work.....	3	4
Figure 3.7. Student A's MT Use During Revision	3	7
Figure 3.8. Student B's MT Use During Revision.....	3	7
Figure 4.1. An Example of Inverted Translation	7	6

CHAPTER 1. INTRODUCTION

This study aims to explore the effects of instruction on machine translation (MT) for English writing revision of Korean middle school learners. This introductory chapter introduces the background, purpose, and outline of the study. Section 1.1 starts with the background of the study, which is followed by the presentation of purpose of the study in Section 2. Then, Section 1.3 introduces the research questions that the study attempts to investigate. Finally, in Section 1.3, briefly presents the organization of the thesis.

1.1. Background of the Study

Machine translation (MT) technologies have been considerably developed based on the artificial intelligence and machine learning system during the past decade. In relation to this technological development, recent studies have reported that MT is being widely used by foreign language learners and has potentials to present benefits in their foreign language learning (e.g., Alhaisoni & Alhaysony, 2017; Baraniello et al., 2016; Lee, 2020). In the past, MT was not generally perceived as an effective learning tool owing to its translation results with grammatical errors and awkward expressions (Niño, 2008, 2009), so its use in foreign language classroom was greatly limited such as in post-editing tasks (Kliffer, 2005). However, the dramatically increased accuracy of MT (Briggs, 2018; Castelvechi,

2016) is shedding light on the possibility of utilizing MT in diverse ways for foreign language learning.

Most studies on MT as a tool for foreign language learning in the early 21st century attempted to investigate students' post-editing work – editing MT output that contains errors – and presume MT output as a bad model (Niño, 2009) that needs to be corrected by the learners. In this model, researchers focused on lexico-grammatical errors generated by MT, viewing them a source that can trigger students' linguistic improvement. However, as the accuracy of MT increases, its level of accuracy was considered to reach the minimum level required for college admission of EFL (English as a foreign language) students (Groves & Mundt, 2015) or clearly be better than the qualities of novice learners' foreign language writing product (Ahn & Chung, 2020; Somers, 2003). The accuracy improvement in MT output quality implies that MT can be utilized now as a Computer Assisted Language Learning (CALL) tool (Niño, 2009) especially for younger or beginner learners. Some studies also found improvement in syntactic and lexical complexity as well as fluency in MT-assisted foreign language writing. However, the results have not been conclusive, requiring more research.

In addition, one possible way of using MT as a pedagogical tool is to allow it during the revision in the process of foreign language writing, which has been investigated by several previous studies (Lee, 2020; Lee & Briggs, 2020; Tsai, 2019). In these studies, the output of MT was used as a

reference resource or a point of comparison when students revised what they had written. The participants in those studies, who are college students, were able to improve their writings using MT as a revision tool. Comparison with MT output helped the participants detect errors in their own English writings and enable them to correct those errors.

According to the literature, revision is defined as problem-solving process in that it involves detecting discrepancies between intended and generated messages, making changes that might or might not affect meaning of the text, before, while and/or after text is written (Bridwell, 1980; Faigley & Witte, 1981). Revision process of EFL students have been studied fruitfully so far (Min, 2006; Sommers, 1980) and those studies analyzed participants' revision in terms of linguistic sizes (Bridwell, 1980; Hall, 1987), purposes (Falvey, 1993; Min, 2006), or action types (Faigley & Witte, 1981). The problem-solving process of learners' revision can be scaffolded with the help of MT because MT outputs can act as individualized corrective feedback that language learners can refer to when revising their original work (Lee, 2020). In other words, MT outputs can function as similar to peer-editing, which is not perfect but proven to be beneficial in helping L2 writers improve writing qualities by enhancing their metalinguistic awareness (Lee, Wong, Cheung, & Lee, 2009). As Kliffer (2005) mentioned, L2 learners can apply their existing knowledge of both L1 and L2 to make corrections when having access to the MT output during the revision process.

The current study attempts to expand the findings of the previous studies by investigating the influence of MT as a revision tool on Korean middle school English learners. The fact that MT users are not limited to adults or highly proficient learners, and that younger generations are more familiar with using MT or other digital tools adds weight to the need for research to apply MT as a language learning supporting tool in a secondary school context.

Regarding the perceptions of language teachers and learners towards MT, language teachers often have negative views on the use of MT for language learning due to the issues of academic dishonesty or students' overdependence on MT (Ducar & Schocket, 2018; Jolley & Maimone, 2015). Above all, teachers are concerned that the adoption of MT might result in cognitive disengagement from the language learning process (Van Praag & Sanchez, 2015). However, regardless of how their teacher might think of MT or its use for language learning, language learners are already using MT for diverse purposes such as assignments, essays, grammar and vocabulary check in their daily lives (Ahn & Chung, 2020; Alhaisoni & Alhaysony, 2017). This trend is evident in EFL classrooms where students are rarely given enough guidance or feedback from their teachers (Ferris, Liu, & Rabie, 2011). Enforcing regulations to prevent MT has been found to be ineffective since students continue to use the tool regardless of such preventions (Steding, 2009). The current situation can be summarized as when and where it is no use to ban students from using MT, and this forces

teachers to find ways to present MT as a helpful language learning aid in classroom context. Moreover, it is necessary to train those ‘digital natives’ to be able to use the technology efficiently when they encounter linguistic problems (White & Heidrich, 2013).

To maximize the value of MT as an educational tool, it has been suggested that students as well as language teachers should be aware of strengths and limitations of MT (Lee, 2020; Stapleton & Kin, 2019). Some researchers suggest that some types of instruction or training regarding MT use should be provided to students (Cancino Avila & Panes Alvarez, 2021; Haukelid, 2021; Lee, 2018; Mahadi, 2016). Especially, when MT users are young or with insufficient foreign language proficiency, it is more of importance to provide adequate instruction on MT use so that they would not just copy and paste MT output during their L2 writings. Some previous studies have investigated the effects of instruction on MT when students conduct foreign language writing (Cancino Avila & Panes Alvarez, 2021; Jo, 2018; Lee, 2018; O’Neill, 2016). In those studies, the results on differences between the groups with and without instruction were not conclusive, leaving the need for further research. It should also be reminded that the instructions in those studies were designed to provide only the basic features or functions of MT, rather than to demonstrate the ways to utilize MT to improve the writing qualities of students’ L2 writing product. Moreover, none of those studies has targeted beginner learners and their use of MT as a revision tool, or further examined the change of students’ perceptions

towards MT by the existence of instruction. This study, thus, is motivated by the necessity for research into MT instruction as a revision tool to encourage students to review their own sentences and compare them with MT outputs in order to correct lexico-grammatical errors and edit with better expressions by using their linguistic knowledge. To be specific, the focus of this study is on the comparison between the groups with and without instruction on MT while having students utilize MT to review their own writing works.

Given the fact that there has been apparent needs to investigate the pedagogical value of MT in EFL secondary classrooms, this study aims to examine the effects of MT instruction for English writing revision of Korean middle school English learners. Although this research investigates only a small number of students, it may still provide pedagogical implications for the necessity or usefulness of MT instruction in Korean middle school English classroom context.

1.2. Purpose of the Study

The primary purpose of the study is to examine the effect of MT instruction as a revision tool on Korean middle school English learners by comparing the writing qualities and revision performance of the participants. This study also aims to examine how instruction can affect student's perceptions towards MT.

The study is expected to contribute to an understanding of the influence of MT instruction when Korean middle school students use MT as

a revision tool. While some studies related to the current research have been undertaken, they allowed MT from the first draft of students' foreign language writing, not having used MT as a revision tool. Not only are the research results on the effectiveness of instruction inconsistent, but also there is still a need to take a different approach, which seems to be more applicable to middle school English learners. Thus, by allowing MT during the revision process of the participants' foreign language writing, this study focuses on training Korean middle school English learners to revise the first draft of their writing so that this investigation can provide practical pedagogical recommendations for English teachers to utilize MT for writing instruction in their English lessons in Korea.

1.3. Research Questions

The research questions derived from the background and purpose of the present study are as follows:

- 1) How does MT instruction affect the qualities of L2 writing revised by Korean middle school English learners using MT in terms of accuracy, sentence complexity and lexical features?
- 2) How does MT instruction affect the revision behavior of Korean middle school English learners in terms of the size, the purpose, and the action type of revision?
- 3) How does MT instruction affect Korean middle school English learners' perception on MT use when revising their L2 writing with MT?

1.4. Organization of the Thesis

The present thesis includes five chapters. Following this introductory chapter, Chapter 2 reviews literature relevant to the investigation of machine translation in foreign language learning field. It also presents the gaps derived from the literature introduced. Chapter 3 describes the research methodology of this study including the description of participants, the procedure of experiments, and data analysis methods. Key findings from an analysis of the data and the related discussion are presented in Chapter 4. Finally, Chapter 5 concludes the present study with the summary of major findings, pedagogical implications, and suggestions for future research.

CHAPTER 2. LITERATURE REVIEW

This chapter presents previous literature relevant to the focus of this study and the research gap. Section 2.1 discusses machine translation for foreign language writing, followed by Section 2.2 regarding MT' use for revision. Section 2.3 gives an overview on previous studies on instruction on machine translation, and then students' perception toward MT is discussed in Section 2.4. Last, Section 2.5 presents the research gap that this study aims to fill in.

2.1. Machine Translation for Foreign Language Writing

For EFL learners, writing is not an easy skill to acquire. It is challenging to improve English writing skills in the EFL context not only because there are not enough opportunities to practice English writing but also students tend to be afraid of a writing task itself (Ryu, 2010). Therefore, providing adequate assistance for EFL writing is necessary for the students to try out writing and maintain interest to achieve communicative goals; machine translation (MT) can play that role.

It is not uncommon anymore for EFL students to use MT for their English writing work (Alhaisoni & Alhaysony, 2017; Hamza & Saadalla, 2021), which affirms the necessity of studies investigating the value of MT as a writing aid tool. Recent studies have proved that MT affects the writing qualities of foreign language (FL) writers in terms of fluency (Garcia &

Pena, 2011; O'Neill, 2016), syntactic and lexical complexity (Cancino Avila & Panes Alvarez, 2021; Chon, Shin, & Kim, 2021; Tsai, 2019), and accuracy (Chung & Ahn, 2021; Lee, 2020; O'Neill, 2016; Tsai, 2020). Those studies tried to examine how MT use affects students' written output in terms of syntactic complexity, accuracy, lexical complexity and fluency (CALF). CALF is one of the methodological frameworks that has been proven to be valid indicators of development in L2 writing proficiency (Crossley & McNamara, 2009; McNamara, Graesser, McCarthy, & Cai, 2014; Norris & Ortega, 2009), and previous studies proved the reliability of automated computational tools such as *L2 Syntactic and Lexical Complexity Analyzer* (Lu, 2011), *lChcker* and *VocabProfiler* (Tsai, 2017) to measure such features.

Not a few researchers have indicated that employing MT can have a positive impact on EFL students' FL writings. Specifically, Garcia and Pena (2011) reported that beginner and early intermediate learners, who were learning Spanish as a foreign language, could communicate 'more and better' in their FL writing when assisted with MT. The 16 participants in their study were required to write a short text directly in Spanish without MT and then to perform a similar writing task while being allowed to use MT. The results showed that the participants composed longer texts with MT – showing improved fluency – and the holistic scores on the writing outcomes increased, confirming the improvement of the quality of the participants' writings. Kol, Schcolnik, and Spector-Cohen (2018), who

explored the benefits of MT for English for Academic Purposes (EAP), conducted a case study, which included two writing tasks, one with MT and the other without MT. The results demonstrated that students wrote significantly more words with improved vocabulary profiles when using MT. Similarly, when 124 adult Chinese EFL learners used Google Translate (GT) for their academic writing, they produced higher quality of writing with a greater number of words written, making fewer number of spelling and grammar mistakes, when writing with MT than when writing without MT (Tsai, 2019). More recently, Cancino Avila and Panes Alvarez (2021), who examined Chilean EFL high school learners' writing products, claimed that scores on syntactic complexity as well as accuracy had been improved when the participants were allowed to use MT for English writings. The improvements in writing qualities seem to become more evident with the advent of neural MT system (Lewis-Kraus, 2016).

These positive results that reported improvement in fluency, accuracy and complexity, however, were not conclusive in that some other studies like the following reported different results. Fredholm (2015) examined two groups of Sweden students learning Spanish as a foreign language. One group was allowed to use MT while the other was not when conducting FL writings. The results of the analysis on students' writing showed the improvement in fluency and complexity, but not in accuracy. When comparing the frequencies of morphological and syntactic errors, there was no statistically significant difference in most error categories,

proving that MT does not help learners improve accuracy in their writings. Some students who used MT committed even more errors because they rigorously write in their L1 resulting in automatic translation containing many grammatical errors, unlike the students who wrote directly in L2 without MT composed with more simple sentences, resulting in fewer errors. Accordingly, the author concluded that MT may have adverse effects in helping learners conduct FL writing with accuracy. He attributed these results to the students' low level of foreign language proficiency that could not properly judge or filter incorrect outputs generated by the translators. In Chung and Ahn's study (2021), text analysis of 91 Korean students' MT-assisted English writings showed improvements in accuracy, but failed to gain benefits in syntactic and lexical complexity. The improvement in syntactic complexity was not conclusive since the researchers found that MT was effective in producing sentences with coordination but was not in producing sentences with subordination and phrasal forms. Furthermore, MT helped learners write with increased lexical variation but with decreased lexical sophistication. These results suggested MT's limitation that confines learners to use a wide range of common vocabulary along with relatively simple sentence structures. They also mentioned that the use of MT can provide different benefits or drawbacks depending on the proficiency levels of the students. For example, they argued that the improvement in complexity measures differed according to the participants' proficiency; MT primarily affected syntactic complexity of low proficiency learners while it

affected lexical complexity of high proficiency learners.

Although MT's positive influence on FL writing has been generally acknowledged with more and more empirical studies, more research is still needed to obtain a more complete picture of MT.

2.2. Use of Machine Translation for Revision

Even with the benefits of MT, researchers are not certain whether the improvement gained through MT use can actually lead to benefits in language learning, especially in terms of low-level learners (Alhaisoni & Alhaysony, 2017; Garcia & Pena, 2011). They are concerned of 'copy and paste' strategy of those who lack linguistic knowledge to make proper judgments. Nevertheless, banning MT is not a solution in this era when and where MT is within every learner's fingertip. Ducar and Schocket (2018) argued that students should learn how to use MT as a self-learning tool since it is meaningless to prevent the 'omnipresent GT' in L2 classrooms. Moreover, it is the instructors' responsibility to present MT as a pedagogically helpful tool so that even low-level learners can benefit from this technology.

Niño (2009) distinguished the use of MT for educational purposes into four main uses: MT as a bad model, MT as a good model, MT for vocational use, and MT as a CALL tool. So far, the most actively studied research field has been post-editing – editing MT output as recognizing and

correcting MT-generated errors – utilizing MT as a bad model (Kliffer, 2005; Niño, 2007, 2008). However, as the accuracy of MT increases with the development of technology (Ducar & Schocket, 2018; Groves & Mundt, 2015), MT's value as a CALL tool to support foreign language learning has also been discussed. In this sense, the potential of MT as a CALL tool can be recognized when it is used as a revision tool. In other words, by encouraging students to refer to MT output as a valuable point of comparison or as a corrective feedback, MT can be presented in the classroom pedagogically, through which learners can use MT as a revision tool for FL composition.

Several studies have demonstrated that MT could help students' revision during FL writing. Even during the relative infancy of MT technology, Garcia and Pena (2011) discovered that post-editing MT outputs allowed students to focus more on the process and editing of writing. Kliffer (2005), Niño (2009) and Enkin and Mejías-Bikandi (2016) argued that MT can enhance students' metalinguistic awareness by allowing them to be aware of errors, correct them to produce more accurate and fluent writing. Furthermore, Lee (2020) and Tsai (2019) found that EFL university student learners were able to correct vocabulary and grammar errors by using MT in the process of revision and thereby producing writing products with better qualities. In addition to the linguistic and grammatical aspect, Groves and Mundt (2015) insisted that foreign language learners can pay more attention on global aspects of writing such as organization or content, as they become

relatively free from local aspects of writing such as grammar or spelling errors by receiving help from MT. Considering rapid improvement of MT technology, it is needed to investigate the potentials of MT as a revision tool. Besides, revision is a significant process that EFL writers need to learn or acquire to produce better FL writing products (Chenoweth & Hayes, 2001).

In previous literature, revisions of writing have been classified according to their linguistic sizes: letter, word, phrase, clause, sentence, paragraph, or text (Bridwell, 1980; Hall, 1987), based on discourse units. Revisions have also been classified according to their purposes, and the categories vary depending on researchers. For example, Hall (1990) analyzed revisions with four purposes: mechanical, grammatical, cosmetic, informational, while Falvey (1993) and Min (2006) identified five purposes of revision: grammatical, cosmetic, texture, unnecessary expression, and explicature. Lee (2020) analyzed linguistic sizes and purposes of revision made by Korean university EFL students in her MT study. She analyzed student-made revisions on the basis of their linguistic sizes and purposes: symbol, word, phrase, clause/sentence, and paragraph and mechanic corrections, word use corrections, grammatical corrections, changes to better express the same meaning, and changes in content. The results of the analysis demonstrated that the participants made word revisions and with the purpose of replacing expression the most frequently. Moreover, revisions have been classified according to their action types: addition, deletion, substitution, permutation, distribution, and consolidation as designed in

Faigley and Witte (1981).

Considering FL learners' limited linguistic knowledge to perform their FL writing, they would need some form of external help such as teachers' feedback. However, such external assistance is usually limited due to large class sizes and lack of instructional hours (Ferris et al, 2011). In this context, the presentation of MT as a reference source to which students can refer can be an alternative to scaffold their FL writing. So far, the participants of the previous studies on this issue have been limited to tertiary level students whose language proficiency would be relatively high enough to make a strategic decision on MT output rather than simply adopt it. Therefore, the effectiveness of MT-assisted revision needs to be further investigated with the learners of younger FL learners.

2.3. Machine Translation Instruction

Researchers have argued that instructors should be aware of MT's strengths and limitations in order to provide adequate guidance to students when allowing this tool for FL writing (Bahri & Mahadi, 2016; Giannetti, 2016; Haukelid, 2021; Jo, 2018; Lee, 2021; O'Neill, 2016; Stapleton & Kin, 2019). Bahri and Mahadi (2016) insisted that teachers should prepare students to utilize MT as a supplementary tool for FL learning. This implies the necessity of providing proper instruction or training opportunities prior to presenting MT as a language learning tool in order to help the learners

use it effectively. Several studies attempted to discover the effect of MT instruction, but the results so far have been inconclusive. O'Neill (2016) sought to find whether MT could help university learners' composition in their L2, French. He assigned 32 participants into one of three groups; group A composed without MT, group B composed with MT without instruction, and group C composed with MT after receiving instruction. The results suggested that group C outperformed group A at a level of statistical significance, while group B performed better than group A but not significantly so, suggesting the positive impact of MT as well as prior instruction. Group B and C both received higher scores on comprehensibility, content, grammar, and spelling. In Jo's study (2018), Korean university EFL students were instructed to use MT to engage in free writing activities. Throughout one semester of MT instruction, the researcher dealt with the following features prior to the main writing tasks: 1) basic features of MT 2) major strategies that students could utilize when using MT 3) pre-editing skills to modify L1 input and 4) directions to write reflection papers to recall their own MT use experience. The results of her study demonstrated that both the amount of English composition and students' self-efficacy in English writing increased regardless of students' L2 proficiency. While the studies of O'Neill (2016) and Jo (2018) examined adult language learners, Cancino Avila and Panes Alvarez (2021) contributed to this field by investigating high school EFL learners' L2 writing products. Sixty-one Chilean EFL high school learners were

randomly assigned to one of three groups: MT without instruction, MT with instruction, and a group without MT. Writing qualities written by each group' participants were analyzed in terms of T-unit length, syntactic complexity and accuracy. Findings indicated that both groups that had access to MT showed significantly higher scores in syntactic complexity and accuracy. It was interesting that the group with MT with instruction showed higher scores than the group with MT but without instruction although this difference was not large enough to be statistically significant.

So far, the effects of instruction have not been conclusive, and the participants of the previous studies were older than middle schoolers – university or high school students. Thus, it is necessary to examine this issue further with younger learners. Besides, the instructions on those studies either presented only the basic MT features or allowed MT from the draft writing rather than guiding learners to utilize it for specific educational purposes such as revision. The instruction on MT would be more necessary for younger students who do not have enough experience or opportunities to use this tool during the writing process. Moreover, the instruction regarding specific ways of MT use with the purpose of revising one's original written work might produce different results from the previous studies, leading to valuable discovery on the benefit of MT.

2.4. Students' Perception toward Machine Translation

Previous studies have investigated how FL learners perceive MT and the results show their mixed opinions. Generally, students seemed to hold a positive view when they conduct FL writing with MT. This is because they can express themselves better with MT (Garcia & Pena, 2011) and the cognitive load when producing FL draft can be lessened with the help of MT (Im, 2017; Niño, 2009). The positive view was especially evident in terms of vocabulary search; many of the participants responded that MT allows efficient searching for appropriate vocabulary during writing (Alhaisoni & Alhaysony, 2017; Clifford, Merschel, & Munné, 2013; Kol et al, 2018; Tsai, 2019). In addition, MT can be beneficial for students' affective aspects such as confidence and motivation (Lee, 2020; Niño, 2009; Tsai, 2019) . Overall, the MT users expressed satisfaction with using MT for their FL writing and showed a willingness to use it again in future writing tasks (Ahn & Chung, 2020). On the other hand, learners raised some questions on the accuracy of MT output. Students often discover inaccuracies in MT output, which leads to gradual distrust on the MT system (Clifford et al., 2013; Im, 2017). Kol et al. (2018) reported that students are also concerned that MT might generate a hindrance to practice English. Similarly, the participants in Garcia and Pena's study (2011) responded that MT prevented them from thinking in the target language. Lee (2020) reported that students' proficiency might be related to their perception on MT use. In her study, high proficiency students

mentioned the negative side of MT such as inaccurate grammar more frequently while low proficiency students tended to rely more on MT without such concern. The limitation of those perception studies is that the participants were all tertiary students. It should not be ignored that MT is widely used by younger learners as well. Thus, it is necessary to examine how young learners such as middle schoolers think of MT and whether instruction can affect their perception towards MT.

2.5. Research Gap

Many studies have investigated MT's effectiveness on FL writing by comparing the qualities of writing products written with and without MT. However, the usefulness of MT as a revision tool has been understudied especially in terms of younger or low-level students. Regarding MT instruction, no single study has investigated its usefulness for the revision process. In other words, little is still known about younger students' MT use with revision purpose: how they can utilize MT as a revision tool and how instruction affects this process as well as students' perception on MT use. To promote a 'desirable' use of MT in middle school classroom settings, it is of importance to understand the value of MT as a revision tool and the effects of MT instruction for English writing revision of Korean middle school learners.

CHAPTER 3. METHODOLOGY

This chapter contains the discussion of the methodological approach and research design employed in this study. A mixed method design incorporating both quantitative and qualitative approaches is adopted (Brown & Coombe, 2015; Larson-Hall, 2010) to examine the research questions set in Chapter 1. Section 3.1 describes information with regard to the participants and the setting of the current study. Section 3.2 provides details on the instruments regarding writing prompt, questionnaire, and translator, followed by Section 3.3 regarding the explanation on MT instruction. Then, in Section 3.4, data collection procedure of the pilot and main study is illustrated. Next, Section 3.5 explains data analysis applied for this study. Finally, Section 3.6 presents preliminary analysis on the original writings and the responses from the background questionnaire.

3.1. Participants

The experiment for this study was conducted at A middle school located in Busan City in Korea. This school was chosen because at the time of the experiment it was possible to conduct in-person instruction in Busan, unlike the Seoul metropolitan area where full-time remote classes were the only mode according to the regional Covid-19 quarantine measures. Participants consisted of a total of forty-seven 9th graders of mixed gender; they are from two intact classes. There were 11 male and 13 female students

in Class 1 and 11 male and 12 female students in Class 2. All students were native speakers of Korean, learning English as their foreign language. English proficiency of the participants varies considerably. There were several students with academic goals to go to foreign language high school with relatively high English proficiency. On the other hand, there were also students whose English proficiency is relatively low. Nevertheless, most of them (except for the two students who will be described in the next paragraph) were able to compose basic English sentences for the target task of this study. In other words, each class was composed of middle school English learners of mixed proficiency, guaranteeing that the participants can be suitable representative for the general Korean EFL students ranging from higher to lower English proficiency. Among the participants (n=47), most of the students (n=34) answered that they have attended a private educational institute to learn English. They mentioned grammar learning (n=9), essay writing (n=5), reading (n=4) and vocabulary learning (n=3) as ways to study English. Some of the students (n=12) reported experiences of learning English abroad, although the majority of them (n=10) stayed less than a year. One student in Class 1 had experience of living in the U.S. for six years and attending a local international school, while another student in Class 2 had stayed in Canada for five years. Both of these students were eliminated from the analysis since the purpose of this study was to examine MT-assisted revision of typical Korean EFL learners who have learned English in a foreign language learning context. Two students from each class were not

equipped with enough basic English proficiency to construct simple sentences in the given writing task. In addition, one student in Class 2 did not fully participate in the study procedures due to absence. Therefore, the data from these five students were removed for the following analysis.

The participants were grouped into two according to their intact classes: Class 1 and Class 2. This study aims to examine the effect of MT instruction, so the researcher provides one group with MT instruction while not providing such instruction to the other group. To confirm the homogeneity of the two groups, preliminary analysis was conducted before the main analysis.

3.2. Instruments

The present study employs a writing task and relevant materials to investigate the effect of MT instruction on Korean middle school English learners' writing revision. Section 3.2.1 presents the rationale for selecting picture-based narrative prompt for the writing task. After this, Section 3.2.2 presents the characteristics of the pre and post questionnaires. Section 3.2.3 explains why Google Translate was chosen as a revision tool in this study.

3.2.1. Picture-based narrative task

A writing prompt for the picture-based narrative task based on six picture strips was chosen for the experiment for the following reasons. First, a narrative writing task is frequently used in the research area of foreign

language learning due to its characteristics of limiting content and reducing cognitive load and thereby requiring extending language use on the part of language learners (Foster & Skehan, 1996). Since the content of the narration is already presented with visual stimuli, it is often easier for the students to focus on language use (Tavakoli & Foster, 2011). Next, this type of task allows the researcher to compare the linguistic features of the same content written by language learners (Kormos, 2011). Because the focus of this study is to compare the participants' revision behavior as well as its products, it was necessary to have them to write similar contents in their own languages by controlling the contents of the writing with visual stimuli. Thus, the picture-based narrative task was assumed to be appropriate for the purpose of the study given that it seeks to examine participants' MT-assisted revision focusing on the participants' language use in particular. The researcher chose a picture prompt that is from one of workbooks developed for National English Ability Test (YBM Si-sa, 2012) (Appendix 1). This task consists of a series of six pictures and each picture describes the story of an abandoned dog and a girl who takes care of it. The connection of the pictures is indicated by horizontal arrows. The participants are required to write a story using the given pictures by encoding visual information into English sentences with some degree of imagination. This picture prompt had been used in a previous study (Lee & Choi, 2019) whose target participants were middle school students. In addition, the appropriateness of the task was examined through the pilot study.

3.2.2. Questionnaires

Two types of questionnaire were developed for this study: a background questionnaire and a post questionnaire. First, the purpose of the background questionnaire is to collect data concerning previous English learning experience and MT-related experience for English writing of the participants prior to the main writing task (see Appendix 2). It firstly asked the participants how they learned English thus far, including whether they have studied English in foreign countries, and if so, for how long they have stayed in the foreign country. Then, it asked the participants' prior MT-related experience. The questionnaire contained five MT-related questions, which were modified from Clifford et al. (2013), Im (2017), and Ahn and Chung (2020). Participants were asked whether they have used MT for English writing before and, if so, how often, and which one, *Google Translate* or *Naver's Papago*, they have used in the past. Lastly, the final two questions in the questionnaire sought how the participants subjectively perceive MT and included an open-ended question to allow the respondents to explain the reasons.

Second, the post questionnaire aimed to examine the participants' perception of MT use for their revision after the experiment was conducted on each group (see Appendix 3). It intended to examine how the participants thought of MT for their writing tasks with and without instructional interventions on MT. Again, questionnaires were developed based on the work of Clifford et al. (2013), Im (2017), and Ahn and Chung (2020). Those

questions asked how the participants perceive MT in terms of its usefulness, satisfaction, and intention to reuse MT for future writing work. The participants were prompted to answer using a 5-point Likert scale ranging from 1 ('strongly disagree') to 5 ('strongly agree'). Two open-ended questions are additionally included to allow the participants to express their opinions freely regarding GT.

To remove a possible misinterpretation of the questionnaire questions in consideration of students' English proficiency, each questionnaire was translated into Korean, and the Korean version was presented to the participants.

3.2.3. Google Translate

This study aims to investigate the value of MT as a revision tool along with the effect of instruction. The researcher has decided to present *Google Translate* (GT) as a revision tool for this study based on the following reasons; first, GT is one of the most widely used MT by language learners (Ahn & Chung, 2020; Clifford et al., 2013) because of its ease of use, free of charge, and quick performance (Google AI Blog, 2016). Second, it does not provide further vocabulary information with a bilingual dictionary – unlike *Naver's Papago*, another popular web-based MT in Korea. Since the main focus of this study is on investigating MT, providing other linguistic source such as bilingual dictionary information may influence the research result. Third, according to Lee (2019), Korean EFL

participants gave GT a higher score than *Papago* when asked which translator is useful for discovering linguistic errors. Since the MT instruction in the current study is supposed to train the participants to detect errors in their original writings and edit them with accurate forms by using MT and its outputs, GT, seen to be better for detecting errors by Korean EFL students, appears to be more appropriate for the research. Finally, the results of the background questionnaire showed that the participants had been more exposed to *Naver's Papago* prior to this study compared to GT. To minimize the exposure effect, GT was chosen as an instrument for this study.

3.3. Machine Translation Instruction

Machine Translation (MT) instruction undertaken throughout the current study is described in the following sections. Section 3.3.1 explains the objective of MT instruction, and Section 3.3.2 gives details regarding five components of MT instruction. Finally, the sequence of MT instruction is explained in Section 3.3.3.

3.3.1. Objective

The objective of MT instruction in this study is to train students learn how to use MT for their revision task by raising their awareness of GT and presenting possible ways to utilize its outputs to edit their self-written texts. To achieve this objective, components and sequences of MT

instruction were designed by the researcher based both on 1) the result and discussion of previous research, (i.e., Burton, 2003; Cancino Avila & Panes Alvarez, 2021; Clifford et al., 2013; Ducar et al., 2018, etc.; Kim, 2019; Lee & Kim, 2018; Luton, 2003) and 2) the result of the pilot study, which will be discussed in Section 3.4.1.

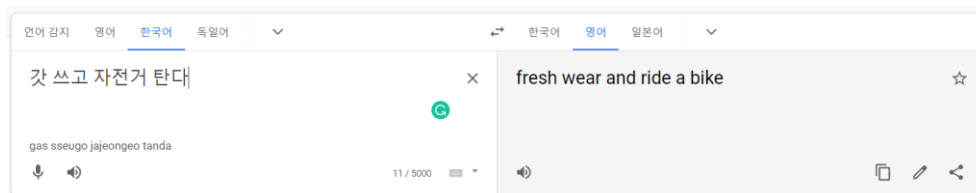
3.3.2. Components

The instruction included five components as follows: 1) features of GT, 2) strengths and weaknesses of GT, 3) GT ‘output’ as a point of comparison and modified ‘input’ with the purpose of revision, 4) revision and GT-assisted revision, 5) the issue of ownership. In the first component, basic features of GT including language selection, inversed translation, provision of synonyms and example sentences, voice translation, audio support for pronouncing unknown words or sentences, were presented with captured images.

In the second component, the strengths and weaknesses of GT were provided, with specific examples adopted from Lee (2020) and Kim (2019). Lee (2020) reported that the accuracy of MT output is higher especially with article and preposition errors, which Korean EFL learners often find difficult to identify (Back, 2011; Bitchener, Young & Cameron, 2005; Lee, 2014). Moreover, MT is good at presenting contextualized words (Bahri & Mahadi, 2016; Chang & Sun, 2009) and collocations (Briggs, 2018). During the instruction, MT’s convenience, immediacy, efficiency, free cost (Niño,

2009; Sukkhwon & Sripetpun, 2014) were cited as advantages of MT. On the other hand, as Kim (2019) cited several crosslinguistic features and culture-specific phrases that are inaccurately translated with GT between Korean and English, students were exposed to such examples during the instruction. For example, in Figure 3.1, *gat* – a Korean traditional hat – was translated incorrectly and the translation result is far from the original proverb’s meaning, which is ‘wearing an outfit or being in state that doesn’t fit a certain situation at all’. This mistranslation was caused by an inadequacy of technology in translating culture-specific words or phrases; GT translated *gat* to *fresh* since Korean word *gat* is a homophone in that it can mean both a traditional Korean hat and an adverb whose meaning is ‘just now’. This was viewed as a good starting point to present MT’s limitation, which would alert the learners to be sensible regarding MT outputs so that they can make proper judgement on MT outputs rather than accepting them without filtering.

Figure 3.1. An Example of Mistranslation of GT



The third component introduced strategies to use MT output as a reference source that participants can refer to and ways to modify Korean

input typed into MT to elicit accurate English translation. In the fourth component of the instruction, the concept of revision and how MT can facilitate this process were introduced mainly with the researcher's demonstration.

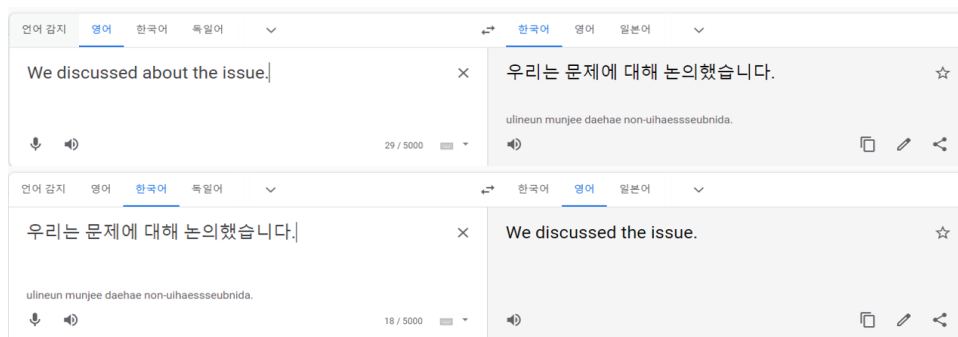
Finally, the issue of ownership, which had been discussed in many of the previous studies related to MT (e.g., Correa, 2014; Ducar et al., 2018; Kazemzadeh & Kashani, 2014; Somers, Gaspari, & Niño, 2006), was dealt with so that the participants could be aware of the authorship or copyright issues of using MT for their writing composition.

3.3.3. Sequence

The researcher provided two periods of MT instruction to the participants in Class 1 (hereafter, Instruction Group) along with the following sequence of steps. Each lesson was carried out for 45 minutes, which is the same as the regular class time for Korean middle school students. In the beginning part of the first instructional session, the researcher's explicit explanation on the characteristics of MT including its main functions, strengths, and weaknesses was provided. After this introductory part, the researcher demonstrated MT-assisted revision with several example sentences by sharing her computer screen using a beam projector in the computer lab. The demonstration process can be divided into three parts as follows: inverted translation, pre-editing, diversifying input sizes. In the first part, English sentences containing grammatical errors,

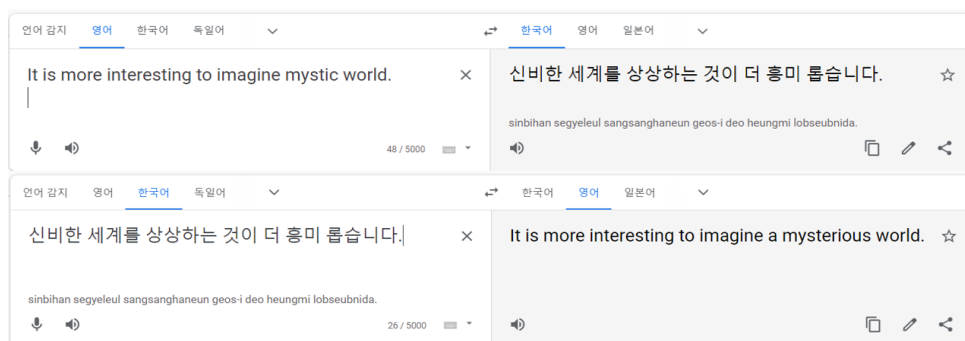
for example, ‘We discussed about the issue’, were entered into MT and the researcher had students check their Korean translation. After confirming Korean translation, the inverted translation was performed. The result of the inverted translation, for example, ‘We discussed the issue’, were used to compare with the original sentences (Figure 3.2). This comparison process allowed the students to notice the differences between the original sentences and the MT outputs, and thereby detect possible grammatical errors in the original sentences.

Figure 3.2. An Example of Inverted Translation for Comparison



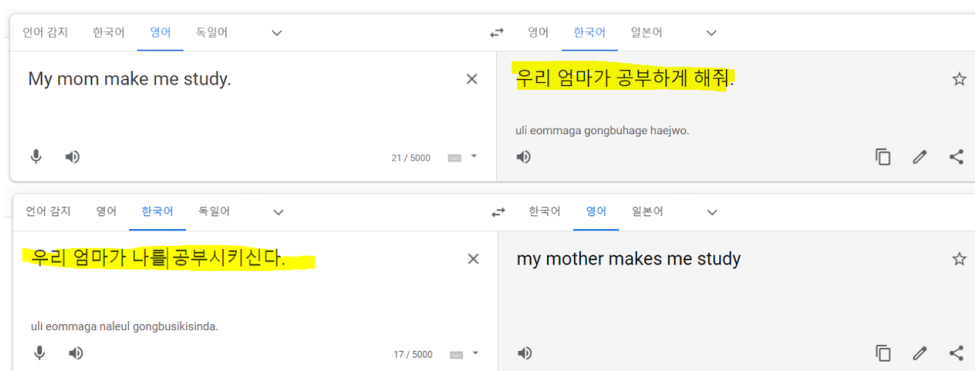
Additionally, it was explained that this comparison process could apply not only to find errors but also to refer to better expressions with similar meaning. For example, the researcher entered the sentence ‘It is more interesting to imagine mystic world’ and performed the inverted translation whose result was ‘It is more interesting to imagine a mysterious world’ (Figure 3.3). It was explained that the results of MT could show different expressions with higher frequency that users can refer to.

Figure 3.3. Inverted Translation to Discover Different Expression



Second, in order to explain ‘pre-editing’, the researcher entered the sentence ‘My mom make me study’ and shared its Korean translation (Figure 3.4). Students recognized that its Korean translation was completely incorrect due to the simple grammatical error – third person singular. The researcher then explained that in case Korean translation of an English sentence was completely inappropriate like the example, it was necessary to either find grammatical errors or if not possible, enter accurate Korean input to elicit proper English translation. The latter process which was called ‘pre-editing’ – refining Korean sentences so that MT’s translation to be more accurate – was briefly introduced.

Figure 3.4. An Example of Pre-editing



Finally, the researcher showed translation with diverse input sizes from word, phrase, and clause and reminded the students that it is important to diversify input sizes according to their linguistic inquiries to get accurate translation results. After this demonstration, the researcher explained what ‘revision’ means and how MT can facilitate this process. Then the researcher wrapped up the ways to utilize MT with the purpose of revision (Appendix 4).

After this, the students were given a few focused exercises to use MT to revise sentences with and without errors. The exercise examples are presented with Figure 3.5.

Figure 3.5. A Focused Exercise to Practice MT-assisted Revision

Let's Practice!

구글 번역기를 활용하여 다음 문장들을 수정해 봅시다.

Our family like horror movie and action movie,	→ _____
I usually play games in the weekend,	→ _____
For example it shows many video,	→ _____
Listening music is make me happy,	→ _____
I saw an English exam,	→ _____
I enjoy drinking red tea,	→ _____

In the second instructional period, a pair work and a group work were presented to have students apply what they had learned in the previous class. A pair work and a group work were designed with sample sentences

and a sample paragraph respectively (see Appendix 5). They were encouraged to use MT during these co-works to revise the texts by correcting errors and changing expressions. After finishing the pair work, possible revisions for each sentence were introduced as for an answer check step. After the group work, the students shared the MT-assisted revision products with classmates via Padlet (<https://padlet.com>) (Figure 3.6).

Figure 3.6. Padlet Sharing During the Group Work



As the last step of the instruction, the ownership issue was briefly dealt with by the researcher, insisting that MT should be utilized as a revision tool to improve student's writing qualities, not as a copy and paste tool. The key points of instructional materials and the procedures of each instructional period are summarized in Table 3.1.

Table 3.1. Instructional Components and Procedures

Components	Sequence	Periods
1. Feature of GT 2. Strengths and weakness of GT	1. Explanation	1st period (45 min.)
3. GT output and input	2. Demonstration 1) Inverted translation 2) Pre-editing 3) Diversifying inputs	
4. Revision and GT-assisted revision	3. Explanation 4. Focused practice 5. Pair and group works	
5. Ownership issue	6. Explanation	2nd period (45 min.)

3.4. Procedures

3.4.1. Pilot Study

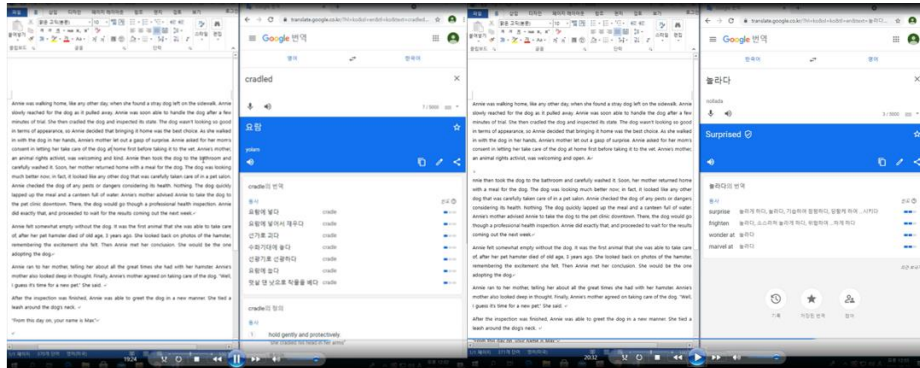
A pilot study was executed to validate the use of the picture-based narrative task (Appendix 1) with 9th graders in A middle school and to develop instructional materials for the main study. The study was conducted in the computer lab of the school and two of the 9th grade students – who would not be included in the main study – were recruited to perform the narrative writing task. Students were asked to write a story based on the given 6 picture prompts using a word processor program, *Microsoft Word*. After finishing writing, the students were asked to revise their original writing by utilizing the web-based version of GT. The grammar-checking function of *Microsoft Word* had been deactivated so that the researcher could

examine how students correct their errors using MT. Their revision performance was recorded with *oCam*, a screen recording software. Both students finished the writing task within 20 minutes and the consecutive revision task within another 20 minutes.

The participants were assumed to have different L2 proficiency based on the different length and quality of the compositions. However, both were able to interpret the given pictures, connect the pictures to make up a story and compose English sentences in order to complete the given writing task without severe difficulty. Moreover, both responded that they were familiar with the topic of the picture prompt, and it was not too difficult for them to perform the writing task. To sum up, the pilot study upheld the validity of the difficulty or the appropriateness of the picture prompt from YBM Si-sa (2012) for the writing task for Korean middle school English learners.

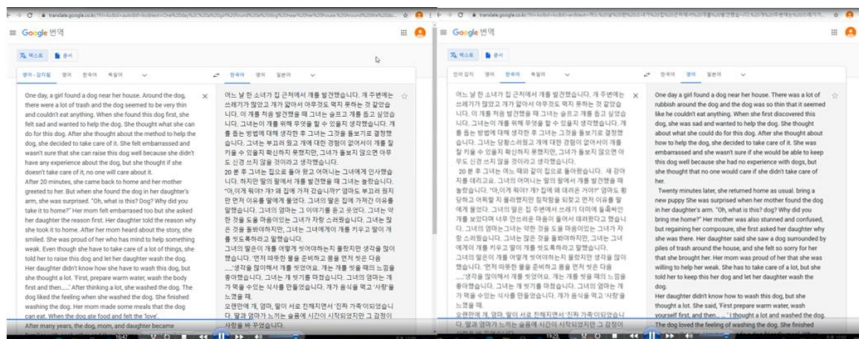
The participants' MT-assisted revision process was examined through the screen recordings and the examination led to the following findings. First, translation direction was limited to only one direction: one student dominantly translated Korean to English, while the other student mainly translated English to Korean. They did not seem to know how to double check the translation results by changing translation directions or using inverted translation function. Second, the sizes of input that each student had typed into GT were also limited. One student entered only 'words' in GT, using MT as if it was a dictionary as in Figure 3.7.

Figure 3.7. Student A's MT Use During Revision



On the other hand, the other student copied and pasted the entire text that she had written in the writing session and then translated it as a whole. After reversing the translation result, she edited her Korean text to elicit English output and tried to revise her English writing based on it as in Figure 3.8. However, since the text was too long, she ended up adding new sentences and paragraphs rather than detecting and revising self-made grammatical errors. This suggests that students should be able to diversify the units of input depending on their linguistic inquiry so that they can elicit diversely segmented language outputs with effective lengths to be used as a reference source.

Figure 3.8. Student B's MT Use During Revision



Third, none of the participants paid attention to synonyms or example sentences that GT provided during the revision. In addition, one of the participants did not use the inversed translation function at all. This signals that students need to be introduced to basic features of GT for its effective use. Finally, they failed to correct grammatical errors in their original writings even with correct forms in MT outputs. This inability to correct errors may be due to a lack of English proficiency. Or it could be because of the absence of MT use experiences for a revision task; it is possible that they were not able to detect the errors since they were not equipped with adequate strategies to utilize abundant amount of language inputs from MT.

These findings comprehensively suggested that the researcher include the following points in her instruction for the main study. First, students should be able to become fully aware of what GT can do: not only generating translation outputs but also providing synonyms as well as example sentences. Second, students should be able to diversify the units of inputs to get more accurate translation depending on their linguistic inquiry. Third, students should be able to use strategies to pay attention to their own errors and correct them through comparison between MT output and their original writings. These points were dealt with instructional sessions in the main study, which will be described in the following section.

3.4.2. Main Study

The main study lasted for three days and 47 students from two intact classes participated. The whole procedures were conducted in the school computer lab, where students can use desktop computers individually as well as listen to the instructor's lectures and perform group and pair works. The researcher took on the role of an instructor. On the first day, a background questionnaire was given to all participants to collect data concerning previous English learning experiences and prior MT-related experience for English writing. Then, a picture-based narrative task was given to the participants. The participants were required to write a story in English based on the given picture prompt – the same picture strip used in the pilot study. The students in the pilot study had completed the writing and revision tasks within 20 minutes respectively. Based on this, it was decided that the participants in the main study were given 30 minutes for each task completion. They wrote without consulting any external resources in this session.

On the second day, the researcher provided two periods of instruction to the participants in Class 1 (hereafter, Instruction Group). In the first instructional period, the researcher's explicit explanation, demonstration and some revision exercises were provided. In the second instructional period, pair work and group work were presented, and the researcher explained the issue of ownership at last.

On the other hand, the participants in Class 2 (hereafter, No-

instruction Group) were only provided with an explanation of the characteristics of MT and revision respectively, but not any further. The instructor had the students find grammatical errors in the given example sentences and correct them without MT. No additional instruction between the MT use and revision components was presented to this group. But to allow learners to become used to the revision itself, several independent revision practices were provided in the form of pair and group work. In addition, to equalize the amount of exposure to GT, these learners were given additional time (20 minutes) to use GT freely.

Next, on the last day of the study, the participants in each group conducted MT-assisted revision: they were supposed to revise what they had originally written by utilizing MT. They were allowed to use only GT as a revision supporting tool. To be aligned with this purpose, the grammar checking function of *MS word* – including spell check – had been deactivated during the revision task. This task was conducted for 30 minutes. Lastly, a post questionnaire was administered at the end of the procedure to examine each group participants' perceptions on MT use for their revision. The experiment procedure of the main study for each group is summarized in Table 3.2.

Table 3.2. Summary of Experiment Procedure

	Instruction group	No-instruction group
Day 1	A background questionnaire (10 min.)	
	↓	
	A picture-based narrative task (30 min.)	
Day 2	1st instructional session (45 min.)	1st instructional session (45 min.)
	<ul style="list-style-type: none"> • Explanation on MT • Teacher's demonstration how to use MT to revise sentences • Explanation on revision and MT-assisted revision • A few focused revision exercises using MT 	<ul style="list-style-type: none"> • Explanation on MT • Explanation on revision • Correcting errors in sample sentences • A few focused revision exercises
Day 2	2nd instructional session (45 min.)	2nd instructional session (45 min.)
	<ul style="list-style-type: none"> • Pair work for revision practices with sample sentences by using MT • Group work for revision practices with sample writings • Sharing MT-assisted revision products via Padlet 	<ul style="list-style-type: none"> • Pair work for revision practices with sample sentences • Group work for revision practices with sample writings • Time allotment for free MT use
Day 3	MT-assisted revision (30 min.)	
	↓	
	A post questionnaire (10 min.)	

3.5. Data Analysis

This study implements a mixed method to resolve the research questions. This section introduces statistical and qualitative methods in accordance with the three research questions.

3.5.1. The analysis of the participants' L2 writing qualities

Research question 1 was 'How does MT instruction affect the qualities of L2 writing revised by Korean middle school English learners using MT in terms of accuracy, sentence complexity and lexical features?' To answer this question, the original and the revised writing products of Instruction Group and No-instruction Group were analyzed with the following analysis tools. Three types of online freeware, *1Checker* (www.1checker.com), *Web-based L2 Syntactical Complexity Analyzer* (<https://aihaiyang.com/software/l2sca/>), and *VocabProfiler* (www.lexutor.ca/vp/eng) were used to assess the following aspects of students' writings respectively: accuracy, syntactic complexity and lexical usages. First, *1Checker* provides automated writing evaluation by counting errors in spelling and grammar. The number of spelling errors and grammatical errors were respectively counted by the tool. Plus, after counting the total number of words written in original and revised versions of students' writings, the researcher calculated the ratio of total errors in each version. Second, *Web-based L2 Syntactical Complexity Analyzer* analyzes given texts with 14 different measures using length of production units, amounts of coordination and subordination, and degree of phrasal sophistication (Lu, 2010). According to Norris and Ortega (2009), general complexity, complexity via subordination, and subclausal complexity via phrasal complexity should be comprehensively considered when examining EFL learners' writing products. This study chose mean length of T-unit

(MLT) as a measurement for general complexity, clauses per T-unit (C/T) as a measurement for complexity via subordination, and mean length of clause (MLC) as a measurement for phrasal complexity. MLT and MLC were chosen as in Kang and Lee (2019), which measured syntactic complexity of Korean middle school EFL learners' writings. They did not analyze C/T because, due to the participants' weak English competence, the measure was not discriminative. However, this study included C/T since not only students' writing products but also their MT assisted revision products would be analyzed. Lastly, *VocabProfiler* analyzes the features of vocabulary that writers use in their writings, classifying the words of the texts into different categories: K1 (the most frequent 1000 words), K2 (the second most frequent thousand words), AWL (academic word list), and off-list words.

Each group participants' original writings and their revised writings were analyzed and compared according to the statistics in order to see if the MT instruction has influenced the qualities of revision outcomes. Since the number of participants ($n=22$ for IG, $n=20$ for NG) did not meet the conditions for using a parametric analysis, a non-parametric statistics for paired samples t -test, Wilcoxon signed-rank test, was employed. In addition, to examine the intervention effect of MT instruction, a measure of effect size, r , was calculated by dividing the absolute Standardized text statistic Z by the square root of the number of pairs ($r=Z/\sqrt{n}$). Plonsky et al. (2014) reported field-specific classification of effect sizes, which is .25 (small

effect), .40 (medium effect), .60 (large effect), which this study refers to. The statistical analyses were carried out using the statistics software *SPSS ver. 25* with the level of significance set at .05. Additionally, qualitative examination on writing outcomes would supplement the statistical findings.

3.5.2. The analysis of the participants' revision behavior

Research question 2 was 'How does MT instruction affect the revision behavior of Korean middle school English learners in terms of the size, the purpose, and the action type of revision?' To find out the answer to the question, the changes appearing in the revised version of each group of students' writings were identified on the basis of their size, purpose, and action type and accordingly coded.

Size of revision refers to the linguistic unit of change: letter, word, phrase, and clause. Purpose of revision refers to the reason for the change: mechanic corrections, lexical exchange, grammatical correction, cosmetic change and informational change. Cosmetic change refers to changes enacted to better expressions that preserves the original meaning, and informational change refers to the changes in contents. Action type of revision refers to addition, deletion, substitution, permutation, distribution, and consolidation. Distribution, in this study, refers to the changes in which a reviser rewrites same information in larger chunks. Consolidation refers to the opposite revisional behavior: putting separate information together. Permutation refers to the revisional act in which a reviser rephrases

information including reordering it. The taxonomy had been established based on the literature (Choi, 2007; Hall, 1990; Lee, 2020; New, 1999; Min, 2006; Stevenson, Schoonen, & De Glopper, 2006), and were finalized through the sample analysis. The taxonomy with the examples is presented in Table 3.3.

Table 3.3. Revision Taxonomy with Examples

		Examples (Changes in bold)	
		Original	Revised
Size	Letter	Sally always liiked to play with Daniel	Sally always likes to play with Daniel
	word	but the dog looks bad .	but the dog looks dirty .
	phrase	Amy's mom prepared some hot milk for the dog. The puppy was on the heap of trash .	Amy's mom prepared a bowl of hot milk for the dog. The puppy was on the pile of garbage .
	clause	Then she wash a dog. The dog has brown dot.	She washed it as soon as she was allowed. There were brown spots around the dog's eyes.
Purpose	Mechanic	She named the dog 'belatrixy'. We went outside and have a walk.	She named the dog ' Belatrixy '. We went outside and have a walk.
		Lexical	She took the dog . she took the puppy .
	Grammatical	she turned back to find out what was it .	she turned back to find out what it was .
	Cosmetic	They decide the dog's name Charlie .	They named the dog Charlie .

	Information	The girl washed the puppy and gave little milk to the puppy.	The girl gave the puppy a new name, washed it and her mother gave the puppy milk.
Action type	Addition	Sujin worried the dog and thought someone lost the dog.	Sujin worried about the dog and thought someone had lost the dog.
	Deletion	So she took the dog and went to her home.	So she took the dog home.
	Substitution	Shelly go to the house quickly, with the dog.	Shelly goes home quickly with her dog.
	Permutation	As shelly likes animals, she decided to take care of the poor puppy.	Shelly, who loves animals, decided to take care of the poor puppy.
	Distribution	When she found a dog, she worries about it, so she takes it and go home with a dog.	When she found the dog, she worried about it. So, she picked up the dog and went home with it.
	Consolidation	Next day, she washes the dog. Also, her mom prepares some milk and foods for the small dog.	The next day she washed the puppy and mother prepared milk and food for her little dog.

The revisions that the participants made during the MT-assisted revision task were coded into different categories according to the developed taxonomy. The revision frequencies coded with each category were compared with statistics in order to see whether MT instruction had an impact on the participants' revision behaviors. Again, because of the small sample size, a non-parametric statistics for two independent samples, Mann-

Whitney U test, was employed with the level of significance of .05. Again, a measure of effect size, r , was calculated by dividing the absolute Standardized text statistic Z by the square root of the number of participants ($r=Z/\sqrt{n}$), which will be classified either as small, medium, or large. Examples of revisions performed by the participants were investigated qualitatively as well to support the statistical analysis.

3.5.3. The analysis of the participants' perception toward MT and its use

For the third research question, 'How does MT instruction affect Korean middle school English learners' perception on MT use when revising their L2 writing with MT?', each group of participants' responses collected via the post questionnaire were examined and compared. The seven Likert scale-based questions and two open-ended questions in the post questionnaire aimed to elicit students' perceptions on MT use depending on the existence of MT instruction. Those Likert scale-based questions asked how the participants recognize MT in terms of its usefulness, satisfaction, and intention to reuse MT for future English writing tasks. The open-ended questions asked the participants to express their opinions freely regarding MT. Cronbach α 's reliability of the questionnaire was checked to guarantee its reliability before converting the Likert scale points into the mean scores for each question ($\alpha=0.955$). The mean scores of each question were compared between the groups through the Mann-Whitney U test. In addition,

students' answers to the open-ended questions were examined to search for emerging themes. This qualitative analysis step is expected to shed light on broadening the understanding on the effects of MT instruction on Korean middle school learners' perception towards MT by supporting the results obtained from the quantitative analysis.

3.6. Preliminary Analysis

Two preliminary analyses were conducted prior to the main data analysis with the purpose of confirming the homogeneity on 1) the qualities of original writings and 2) prior MT-related experiences and perception on MT between Instruction Group and No-instruction group.

Mann-Whitney U test was used to compare the difference between the original writing qualities of the two groups. As presented in Table 3.4, there is no statistically significant difference in the original writing products of Instruction Group and No-instruction Group in terms of number of words written ($Z=-.264, p>.05$), number of spelling errors ($Z=-.370, p>.05$), number of grammar errors ($Z=-.785, p>.05$), total errors ratio ($Z=-.607, p>.05$), mean length of T-unit ($Z=-.151, p>.05$), clauses per T-unit ($Z=-.176, p>.05$), mean length of clauses ($Z=-.554, p>.05$), and the proportions of vocabulary use from K1 ($Z=-.793, p>.05$), K2 ($Z=-.239, p>.05$), AWL ($Z=-.966, p>.05$), OFF ($Z=-1.90, p>.05$) categories.

Table 3.4. Qualities of Original Writings Produced by IG and NG

	IG (n=22)		NG (n=20)		<i>Z</i>	<i>p</i>
	M	SD	M	SD		
Words	170.64	77.26	166.30	86.11	-.264	.791
Spelling errors	2.68	2.77	3.15	3.65	-.370	.711
Grammar errors	5.68	5.53	6.25	4.60	-.785	.432
Total errors ratios	0.05	0.04	0.06	0.03	-.607	.356
MLT	10.41	2.16	11.28	5.20	-.151	.880
C/T	1.46	0.21	1.53	0.53	-.176	.860
MLC	7.14	1.07	7.25	1.80	-.554	.579
K1	91.97	3.16	91.13	3.54	-.793	.428
K2	4.81	1.82	4.73	2.31	-.239	.811
AWL	0.18	0.06	0.28	0.08	-.966	.334
OFF	3.04	0.44	3.86	0.51	-1.90	.057

As the second preliminary analysis, the participants' MT related experiences and perception toward it were measured and compared each other using Chi-Square test and Mann-Whitney U test. To the question of the background questionnaire, 'How many times have you used machine translation for your English writing?', the participants were supposed to choose their answers among 1) every day, 2) less than once a week, 3) less than once a month. This question aims to examine the participants' prior exposure to MT. Table 3.5 shows the distribution of students' responses and the result of Chi-Square Test reveals that there is no statistical difference between the groups ($X^2(2, N = 42) = 0.545, p=0.913$).

Table 3.5. Answers from the Participants on Prior Exposures to MT

		Every day	Less than once a week	Less than once a month	Total
IG	Count	3	11	8	22
	Expected Count	3.1	10	8.9	22.0
	% within Group	13.6%	50.0%	36.4%	100.0%
NG	Count	3	8	9	20
	Expected Count	2.9	9.0	8.1	20.0
	% within Group	15.0%	40.0%	45.0%	100.0%
Total	Count	6	19	17	42
	Expected Count	6.0	19.0	17.0	42.0
	% within Group	14.3%	45.2%	40.5%	100.0%

Next, the participants were asked to report which MT they have used so far between *Google Translate* and *Naver's Papago*. Again, Table 3.6 shows what kinds of MT have the students used so far and the result of Chi-Square Test reveals that there is no statistical difference between the groups ($X^2(2, N = 42) = 2.518, p=0.270$).

Table 3.6. Answers from the Participants on Prior MT Use

		Google Translate	Papago	Both	Total
IG	Count	1	11	10	22
	Expected Count	2.1	12.0	7.9	22.0
	% within Group	4.5%	50.0%	45.5%	100.0%
NG	Count	3	12	5	20
	Expected Count	1.9	11.0	7.1	20.0
	% within Group	15.0%	60.0%	25.0%	100.0%
Total	Count	4	23	15	42
	Expected Count	4.0	23.0	15.0	42.0
	% within Group	9.5%	54.8%	35.7%	100.0%

Finally, the participants' responses to the question 'how much do you agree with the following statement: MT helps to write in English,' were compared via Mann-Whitney U test. The result in Table 3.7 presents that the average score from IG was 2.68 and that NG was 2.65, which were not statistically different ($Z=-.133, p>.05$).

Table 3.7. Scores on Perception toward MT

Group	Mean	SD	<i>Z</i>	<i>p</i>
IG	2.68	0.89	-.133	0.905
NG	2.65	0.93		

Based on the results of the above preliminary analysis, Instruction Group and No-instruction Group are likely to possess the homogeneity regarding original writing qualities and MT-related experience. This enables the researcher to compare the differences depending on the existence of MT instruction with the analysis in the following result and discussion section.

CHAPTER 4. RESULTS AND DISCUSSION

This chapter reports the results of the study and discusses the findings regarding the research questions raised in Chapter 1. Section 4.1 addresses the first research question on whether MT instruction affects Korean L2 learners' writing qualities in terms of accuracy, sentence complexity, and lexical features. Next, Section 4.2 answers the second research question on whether MT instruction affects Korean L2 learners' revision behavior in terms of revision size, purpose, and action type. Lastly, in Section 4.3, students' perception towards MT depending on instructional treatment is examined to answer the third research question. Wilcoxon signed-rank test and Mann-Whitney U test are used for the examination.

4.1. Effects of MT Instruction on Revised Text Qualities

The first research question of the current study is 'How does MT instruction affect the qualities of L2 writing revised by Korean middle school English learners using MT in terms of accuracy, sentence complexity, and lexical features?' To answer the question, each group participants' original and revised writings were compared via Wilcoxon signed-ranks test.

4.1.1. Accuracy

As presented in Table 4.1, there is no statistically significant increase or decrease in terms of the number of words written in the revised

version compared with the original version in case of Instruction Group (IG) ($Z=-.374, p>.05$). In contrast, as the Table 4.2 shows, the number of words written by No-instruction Group (NG) in the revised writings was greater significantly than the number of words written in the original writings ($Z=-3.245, p<.05$). The effect size was .73, which is a large effect according to Plonsky et al. (2014). This shows that the participants in NG wrote significantly more with MT revision while the participants in IG tended to maintain the length of their original writings.

Table 4.1. Comparison of Number of Words in Original and Revised Writings by IG Participants

	Original Writing				Revised Writing				<i>Z</i>	<i>p</i>	<i>r</i>
	<i>M</i>	<i>SD</i>	<i>Min.</i>	<i>Max.</i>	<i>M</i>	<i>SD</i>	<i>Min.</i>	<i>Max.</i>			
Number of Words	170.64	77.26	65	332	169.95	66.63	79	320	-.374	.722	.08

Small effect, $.25 \leq r \leq .39$, medium effect, $.40 \leq r \leq .59$, large effect, $r \geq .60$

Table 4.2. Comparison of Number of Words in Original and Revised Writings by NG Participants

	Original Writing				Revised Writing				<i>Z</i>	<i>p</i>	<i>r</i>
	<i>M</i>	<i>SD</i>	<i>Min.</i>	<i>Max.</i>	<i>M</i>	<i>SD</i>	<i>Min.</i>	<i>Max.</i>			
Number of Words	166.30	86.11	42	416	182.25	88.44	65	468	-3.245	.000	.73

Small effect, $.25 \leq r \leq .39$, medium effect, $.40 \leq r \leq .59$, large effect, $r \geq .60$

The results of NG participants are consistent with the results of Garcia and Pena's research (2011) in that beginners of foreign language

wrote longer texts with the help of MT compared to the participants who wrote without the help of MT. They viewed this as a sign of ‘more’ communication for those who ‘could work outside limited vocabulary constraints’ (p.478) by using MT. In a similar way, NG participants in this study managed to describe what they had wanted to express in longer and more detailed texts with MT. However, unlike Garcia and Pena’s study (2011), which allowed the participants to use MT to write L2 draft, MT was utilized as a revision tool in the current study, requiring the participants to perform with more diverse capacities beyond writing longer. More features other than the number of words should be considered to examine the effect of MT instruction on revision process, which will be discussed in the following.

When it comes to errors, as presented in Table 4.3, both spelling errors and grammar errors decreased significantly in the revised version of IG ($Z=-3.750$, $p<.05$ for spelling errors, $Z=-4.022$, $p<.05$ for grammar errors), yielding large effect sizes respectively ($r=.80$ for spelling errors and $r=.86$ for grammar errors). On the other hand, only the number of grammar errors decreased at a significant level ($Z=-3.146$, $p<.05$) with a large effect size, $r=.70$, while the number of spelling errors did not ($Z=-.637$, $p>.05$) in case of NG as presented in Table 4.4. Still, both groups reduced total errors compared to the number of words they made ($Z=-4.028$, $p<.05$ for IG, $Z=-3.522$, $p<.05$ for NG). The effect sizes ($r=.86$ for IG and $r=.79$ for NG) were both large. This can lead to false reasoning such as the participants in

NG had produced longer texts with fewer grammar errors compared to the participants in IG. However, the fact is that IG participants homogeneously improved the accuracy of revised writings by using MT whereas NG participants showed performances that differed greatly from individual to individual.

Table 4.3. Comparison of Number of Errors in Original and Revised Writings by IG Participants

	Original Writing				Revised Writing				<i>Z</i>	<i>p</i>	<i>r</i>
	<i>M</i>	<i>SD</i>	<i>Min.</i>	<i>Max.</i>	<i>M</i>	<i>SD</i>	<i>Min.</i>	<i>Max.</i>			
Spelling errors	2.68	2.77	0	11	0.09	0.43	0	2	-3.750	.000	.80
Grammar errors	5.68	5.53	0	26	0.77	2.45	0	11	-4.022	.000	.86
Total errors per words	0.05	0.04	0.00	0.15	0.00	0.01	0.00	0.03	-4.028	.000	.86

Small effect, $.25 \leq r \leq .39$, medium effect, $.40 \leq r \leq .59$, large effect, $r \geq .60$

Table 4.4. Comparison of Number of Errors in Original and Revised Writings by NG Participants

	Original Writing				Revised Writing				<i>Z</i>	<i>p</i>	<i>r</i>
	<i>M</i>	<i>SD</i>	<i>Min.</i>	<i>Max.</i>	<i>M</i>	<i>SD</i>	<i>Min.</i>	<i>Max.</i>			
Spelling errors	3.15	3.65	0	14	2.85	3.67	0	16	-.637	.564	.14
Grammar errors	6.25	4.60	0	19	3.95	4.73	0	20	-3.146	.001	.70
Total errors per words	0.06	0.03	0.00	0.16	0.03	0.02	0.00	0.08	-3.522	.000	.79

Small effect, $.25 \leq r \leq .39$, medium effect, $.40 \leq r \leq .59$, large effect, $r \geq .60$

IG participants' ability to check and correct spelling errors appears to be consistent among the participants. The average number of spelling errors in the original versions written by IG participants was 2.68, which was reduced to 0.09 in the revised versions. All participants except one were able to correct every spelling error they committed, yielding '0' for spelling errors. As shown in Excerpt 4.1, the participants changed the incorrectly spelled words (*abundanted*, *shower*, *trew*) either to the correct forms (*abandoned*, *shower*) or to the different expression (*take out*) through MT-assisted revision. Changes made during the MT-assisted revision are displayed in bold.

Excerpt 4.1. IG Participants' Spelling Error Correction

- 1.a) One day, the girl went to **trew** some waste. (From Original Writing, hereafter, marked as a)
- b) One day, the girl went to **take out** the trash. (From Revised Writing, hereafter, marked as b)
- 2.a) Wendy thought the dog was so poor and someone has **abundanted** him,
- b) Wendy felt the dog was so pitiful and someone **abandoned** him,
- 3.a) The dog need to take a **shower**.
- b) I think he needs to take a **shower** right away.

On the other hand, NG participants showed a large variations depending on individuals. Only three of the participants corrected every spelling error whereas most of the participants were only partially able to correct the spelling errors they committed. In addition, there were four

participants who could not reduce any spelling errors and left them as in the original versions. The examples 1.a) and 1.b) in Excerpt 4.2 show that the participant corrected the spelling error (*decided*) in his original writing with the correct form (*decided*) in the revised writing like IG participants. However, another student failed to correct misspelled words (*darty*, *deside*) as in the examples 2.a) and 2.b). The examples 3.a) and 3.b) show that the other participant changed ‘*convieant*’ to ‘*convenience*’ correctly while leaving the incorrectly spelled word ‘*snaks*’.

Excerpt 4.2. NG Participants’ Spelling Error Correction

- 1.a) She was so kind that **decideded** to help it by raising it.
- b) She was so kind that **decided** to help it by raising it.
- 2.a) The dog was very small and **darty**. She **deside** to keep the dog and wash it and feed it.
- b) The dog was very small and **darty**. So she **deside** to keep the dog.
- 3.a) Grace was going to the **convieant** store to buy some **snaks**.
- b) Grace was going to the **convenience** store to buy some **snaks**.

Both groups significantly reduced the number of grammar errors through MT-assisted revision, but the patterns are different. First, in case of IG participants, the average number of grammar errors in their original version was 5.68, and it decreased to 0.77. Most of IG participants (n=19) corrected all grammar errors, leaving no errors in the revised versions of their writings. As shown in Excerpt 4.3., they tended to apply partial modifications while maintaining basic sentence structures in order to correct grammar errors.

Excerpt 4.3. IG participants' Grammar Error Correction

- 1.a) She first got permission **to** her mom.
b) She first got permission **from** her mom.
- 2.a) Lucy was thrilled that her mom **actually said she can** keep Sam.
b) Lucy was thrilled that her mom **said she could actually** keep Sam.
- 3.a) **A** lesson that this story gives us is, a little attention can **effect** one's life and we must buy a pet in **a proper situation**.
b) **The** lesson this story gives us is that a little attention can **affect** one's life and we should buy a pet in **the right circumstances**.

In contrast, the average number of grammar errors in NG participants' original writings was 6.25 and it decreased to 3.95 in their revised writings. Although the decrease was statistically significant, the figure still means many grammar errors remained even after the revision process. Several NG participants could reduce grammar errors by correcting them as IG participants did as in the examples 1.a) and 1.b) in Excerpt 4.4. However, overall, NG participants' revision results differed greatly from participant to participant. One of the students reduced 14 grammar errors to 1 through MT-assisted revision, while some of the students reduced only less than three grammar errors ($n=6$), and a few of them did not reduce any ($n=2$). Moreover, three of the participants committed more errors in their revised writings. This could result from NG participants adding new sentences while using MT during the revision, creating new grammar errors, as shown in the examples in 2.a) and 2.b) in Excerpt 4.4. These NG

participants committed more errors and thereby judged to perform worse in terms of accuracy. However, they showed improved fluency with the increased lengths by elaborating their original texts. This pattern of these three students was similar to that of the participants in Fredholm's study (2015) in that they typed longer and more complex L1 sentences in MT and used the automatically generated MT outputs containing more grammar errors without proper editing. The improved fluency of these students will be discussed at the end of this chapter.

Excerpt 4.4. NG Participants' Grammar Error Correction

- 1.a) So she **take** the dog to her house.
- b) So she **decided to take** her dog to her house.
- 2.a) For good news he was healty.
- b) The dog was angry that he went to hopital and got a shot but the girl Alice gave him some delicious snacks. So after visiting the vet there is two good news. First, dog became full and happy because of snacks. Next, Alice and her mom became nice because dog is still healty though he had dehydration symptoms.

Through the instruction provided before MT-assisted revision, IG participants had learned that input sizes into MT can vary from words, phrases, sentences, and MT outputs can be used as a reference source for checking errors and comparing expressions. They appeared to apply what they had learned, and as a result, became able to notice obvious spelling and grammar errors. By doing so, they succeeded in improving the accuracy of their writings by reducing spelling and grammar errors.

On the other hand, NG participants who were not provided with such instruction on MT-assisted revision behaved differently depending on the individuals. Some participants tended to copy and paste MT outputs in their revised writings rather than editing their self-written sentences. This tendency produced lots of sentences without writers' efforts and ownership, as many of the instructors worried about MT use for foreign language writings with the issue of (dis)honesty (Case, 2015; Correa, 2011; Ducar & Schocket, 2018; Somers et al., 2006). Because of this dominance of copy and paste strategy of NG participants, spelling and grammar errors did not seem to be effectively corrected and the success or failure of correcting errors differed greatly depending on individuals.

Despite the different patterns, both groups reduced number of errors and thereby improved text accuracy and this is consistent with the previous studies (Tsai, 2019; Lee, 2020) in which college level students also improved grammatical accuracy throughout MT assisted revision. On the other hand, some NG participants had committed more grammar errors during the revision since they added more contents in their writings. These participants were able to write 'more' thanks to MT, and this was similar to the participants in Garcia and Pena's study (2011), in which the participants were allowed to use MT from drafting. Garcia and Pena (2011) asserted that MT was beneficial especially for beginner students in that they could express themselves better in longer sentences with a larger number of words. These patterns of MT use – using MT from the draft stage or presenting MT

without particular instruction – can be considered more ‘natural’ in that it is more consistent with the original development intention of translation technology. Beginner learners could show improved fluency with more sophisticated languages through MT, which would have been impossible due to their insufficient foreign language proficiency. In contrast, allowing MT as a revision tool or limiting students’ MT use with the form of instruction would not be completely ‘natural.’ However, using MT for revision in EFL classroom and providing related instruction have a different pedagogical value: making learners to be involved in cognitive engagement from the initial composition stage and then encouraging them to rework on those self-made sentences by using their linguistic knowledge to improve accuracy. While accuracy is not a single factor for high English proficiency, it can certainly contribute to linguistic development, especially in case of learners with low proficiency. In this sense, it was meaningful to discover that MT instruction had been positive to improve IG participants’ accuracy. This showed that MT can be a meaningful revision aid tool in secondary EFL writing classrooms when it is provided with proper instruction. Thus, this study result is significant in that it has expanded the previous studies’ findings to middle schoolers and further proved the positive effect of MT instruction that resulted in accuracy improvement of middle school students’ revisional products.

4.1.2. Syntactic Complexity

To compare syntactic complexity of Korean middle school English learners' writings before and after MT-assisted revision with a variable of MT instruction, this study analyzes mean length of T-unit (MLT), clauses per T-unit (C/T), and mean length of clauses (MLC) using *web-based syntactic complexity analyzer*. Table 4.5 demonstrates the results of Wilcoxon signed-ranks test, and there is no statistical difference in all categories between IG participants' original and revised writings ($Z=-.520$, $p>.05$ for MLT, $Z=-1.195$, $p>.05$ for C/T, $Z=1.445$, $p>.05$ for MLC). This suggests that IG participants revised their original sentences without significant changes in syntactic complexity.

Table 4.5. Comparison of Complexity Measures in Original and Revised Writings by IG Participants

	Original Writing				Revised Writing				Z	p	r
	M	SD	Min.	Max.	M	SD	Min.	Max.			
MLT	10.41	2.16	7.16	16.73	10.30	1.74	7.56	14.09	-.520	.616	.11
C/T	1.46	0.21	1.04	1.88	1.41	0.18	1.07	1.78	-1.195	.241	.25
MLC	7.15	1.07	5.27	9.68	7.34	1.13	5.64	10.33	1.445	.154	.31

Small effect, $.25 \leq r \leq .39$, medium effect, $.40 \leq r \leq .59$, large effect, $r \geq .60$

Clearly, IG participants revised their writings by correcting errors or partially changing sentences while maintaining overall sentence structures, rather than rewriting completely new sentences after deleting the whole

sentences. Thus this does not bring any change in syntactic complexity of their writings. An example can be seen in Excerpt 4.5.

Excerpt 4.5. IG Participants' Revision Maintaining Overall Sentence Structures

- 1.a) A girl named Lucy found an abandoned puppy **around** the corner of a wall. The puppy was **on the heap of trash**. (...) She was **first** shocked but she decided to bring the dirty and **wounded** dog.
- b) A girl named Lucy found an abandoned puppy **in** the corner of a wall. The puppy was **on the pile of garbage**. (...) She was shocked **at first**, but she decided to bring the dirty and **injured** dog.

On the other hand, MLT and MLC of the revised version of NG participants increased at a significant level ($Z=-2.461$, $p<.05$ for MLT, $Z=-3.018$, $p<.05$ for MLC) as in Table 4.6. The effect size for MLT was medium ($r=.55$) and the effect size for MLC was large ($r=.67$).

Table 4.6. Comparison of Complexity Measures in Original and Revised Writings by NG Participants

	Original Writing				Revised Writing				<i>Z</i>	<i>p</i>	<i>r</i>
	M	SD	Min.	Max.	M	SD	Min.	Max.			
MLT	11.26	5.23	4.20	28.10	12.28	4.60	6.50	28.10	-2.461	.012	.55
C/T	1.53	0.53	1.00	3.40	1.55	0.50	1.10	3.40	-.521	.619	.12
MLC	7.23	1.84	4.20	13.82	7.90	1.52	5.90	12.82	-3.018	.001	.67

Small effect, $.25 \leq r \leq .39$, medium effect, $.40 \leq r \leq .59$, large effect, $r \geq .60$

Although MLT and MLC increased significantly in case of NG participants, only some of the participants' increases led to an improvement in writing quality. For example, the examples 1.a) and 1.b) in Excerpt 4.6 show that the student revised most of the sentences in his original writing throughout MT-assisted revision process. MLT and MLC in his original writing were 8.25 and 7.07 respectively, and they all increased to 11.17 and 8.93 in the revised writing. Along with those increases in syntactic complexity measures, the overall improvement in the revised writing was observed: sentences became longer with fewer grammar errors, with more detailed explanations. In another example set, 2.a) and 2.b), on the other hand, MLT and MLC were 6.92 and 6.38, and they increased to 8.75 and 6.56 in his revised writing. However, he changed only some of the self-written sentences using MT, leaving other sentences as originals. The bolded sentences in 2.b) were assumed to be produced by MT. Those MT-generated sentences were written in the past while the other self-written sentences were written in the present, resulting in an inconsistent text. Besides, the MT-generated sentences were grammatically error-free while there remained the awkward expressions like 'She doesn't let him' in the student's self-made sentences. Plus, the MT-generated sentences tended to be much longer compared to his original sentences. As a result, there was less consistency among the sentences that made up the entire text; some sentences had been written by the student by himself while others had been generated by MT and copied without proper editing.

Excerpt 4.6. NG Participants' Increased Syntactic Complexity

- 1.a) She heard the sound that a dog is barking, so she walked to there. The girl saw the dog which was thrown away. The looked so messy. So she take the dog to her house.
- b) She heard a dog barking in the distance and the girl walked to it. She felt something moving in her pile of garbage, and when she looked closely, a dog had been left behind. The dog had been in the trash for a very long time and was very dirty. So she decided to take her dog to her house and look after her for a while.
- 2.a) She doesn't let him. She brings the dog in her house. In a house, her mom welcomes them. She loves him. The girl washes the dog in the bathroom. And the girl's mom feed him with milk.
- b) She doesn't let him. She brings the dog in her house. **At house, her mom welcomed them when she saw them coming.** She loves him. The girl washes the dog in the bathroom. **The girl's mom wanted to feed him, thus she gave him some milk which she had bought from a nearby supermarket.**

Moreover, a large discrepancy was observed among NG participants regarding syntactic complexity: some students improved overall syntactic complexity of their writings while others hardly make changes in terms of sentence structures. Like error correction in the previous section, NG participants showed performances that greatly differed from individual to individual.

In contrast, in case of IG, most of the participants managed to undergo sentence revision and did not make noticeable changes in terms of sentence structures. Their final products were composed of the sentences that had been modified to the extent that IG participants can utilize their

own linguistic – usually lexical – knowledge by using MT outputs. This revisional methods or behaviors can be considered ‘safe’ in that students only revise to the extent that they would not commit additional new errors by partially correcting the original sentences, rather than rewriting sentences or trying out new sentence structures. In other words, IG participants were too careful to try out new sentence structures or to elaborate what they had written in longer sentences. Considering that learners need to take a certain amount of risk in order to test their linguistic hypotheses and confirm them to achieve a higher level of proficiency, it is necessary to encourage beginner students to try out more sophisticated sentences with higher syntactic complexity despite minor grammatical errors. In this sense, instruction in this study may have an adverse effect that limits students’ creativity, autonomy, or courage to try out more diverse forms of sentence structures throughout the revision process. Since the focus of this study is limited to compare two groups of participants with and without MT instruction, follow-up research is needed to obtain a more detailed picture of MT instruction.

4.1.3. Vocabulary Use

When comparing vocabulary use of the categories of K1, K2, AWL and OFF via Wilcoxon signed-ranks test, the participants in IG used significantly more AWL words in their revised versions ($Z=3.059$, $p<.05$). Other than that, there was no statistically significant difference ($Z=-1.218$,

$p > .05$, $Z = .504$, $p > .05$ for K2, $Z = -.560$, $p > .05$ for OFF). In contrast, the proportions of words that belonged to each category did not reach to significant difference in all categories in case of NG participants ($Z = -.784$, $p > .05$ for K1, $Z = .282$, $p > .05$ for K2, $Z = .864$, $p > .05$ for AWL, $Z = .187$, $p > .05$ for OFF). Table 4.8 and 4.9 displays the above.

Table 4.7. Comparison of Vocabulary Profiles in Original and Revised Writings by IG Participants

	Original Writing				Revised Writing				Z	p	r
	M	SD	Min.	Max.	M	SD	Min.	Max.			
K1	91.97	3.16	86.09	96.97	91.29	3.62	81.48	96.90	-1.218	.223	.26
K2	4.81	1.82	1.52	9.27	4.89	1.82	0.69	8.86	.504	.614	.11
AWL	0.18	0.30	0.00	0.93	0.69	0.81	0.00	2.53	3.059	.002	.65
OFF	3.04	2.06	0.58	9.52	3.13	3.19	0.00	12.96	-.560	.575	.12

Small effect, $.25 \leq r \leq .39$, medium effect, $.40 \leq r \leq .59$, large effect, $r \geq .60$

Table 4.8. Comparison of Vocabulary Profiles in Original and Revised Writings by NG Participants

	Original Writing				Revised Writing				Z	p	r
	M	SD	Min.	Max.	M	SD	Min.	Max.			
K1	91.13	3.54	83.44	95.96	90.82	3.44	80.85	97.30	-.784	.433	.18
K2	4.73	2.31	1.40	10.60	4.84	2.15	1.35	11.35	.282	.778	.06
AWL	0.28	0.34	0.00	0.81	0.44	0.44	0.00	1.25	.864	.388	.19
OFF	3.86	2.27	2.02	12.50	3.90	1.82	1.35	8.33	.187	.852	.04

Small effect, $.25 \leq r \leq .39$, medium effect, $.40 \leq r \leq .59$, large effect, $r \geq .60$

Among IG participants, the proportion of AWL words increased in 13 participants' writing products out of 22. As shown in Excerpt 4.7., the proportion of AWL words increased by 1) changing a word in a sentence, 2) correcting the spelling errors, and 3) adding a new sentence containing AWL words. However, exchange to AWL words was not always positive as in 6.a) and 6.b); the sentence modification using AWL words like '*obtain*', '*consent*' generated awkward expressions.

Excerpt 4.7. IG Participants' Vocabulary Change

- 1.a) She was first shocked but she decided to bring the dirty and **wounded** dog.
- b) She was shocked at first, but she decided to bring the dirty and **injured** dog.
- 2.a) She went out the bakery, and find the sound coming spot.
- b) She went out the bakery, and find the **source** of the sound.
- 3.a) She realized that someone **threw away** the dog.
- b) She realized that someone had **abandoned** the dog.
- 4.a) The dog looked like it was **abanded** for a long time and it was hungrily looking for something to eat.
- b) The dog looked **abandoned** for a long time and was looking for something to eat with a hungry expression.
- 5.a) (*There is no original sentence.*)
- b) On weekends, she used to **volunteer** for **abandoned** dogs.
- 6.a) After getting agree from her mother, the girl tried to take care of the dog.
- b) The girl tried to take care of the dog after **obtaining** her mother's **consent**.

It was also noteworthy that all but one of IG participants whose AWL words use increased had used the word '*abandoned*' in their revised writings. Considering that the word '*abandoned*' belongs to the basic

vocabulary list for 2015 Revised National English Curriculum (Ministry Education, 2015), it is highly plausible for the participants to have been exposed to this word. However, they failed to use the word in their original writings; instead, they used a more colloquial expression (*throw away*) or used the word in incorrect spelling (*abanded*). Then through MT-assisted revision, many of them became able to use the word correctly. This represents that MT-assisted revision helped the participants to elicit vocabulary items that are beyond a certain level by triggering their receptive vocabulary knowledge. MT served a scaffolding role for the middle school participants to use more difficult words in their writings.

In case of NG, since the total number of words increased, the proportion of AWL words of each participant remained the same or became lower. Five participants had written more AWL words as they revised with MT and three of them used the word '*abandoned*' in their revised writings.

To sum up, IG and NG, who went through MT-assisted revision with and without instruction, could commonly improve accuracy to some degree, but they quite differed in the following aspects. IG participants were able to effectively use MT and its outputs to detect and correct self-made errors as applying what they had learned through instruction, and thereby most of them succeeded in eliminating spelling and grammar errors in the revised versions of their writings. In contrast, the success of NG participants' revision in terms of correcting spelling and grammar errors varies greatly depending on individuals, so there were quite a few participants who could

not correct errors during MT-assisted revision. Regarding syntactic complexity, there was no statistical difference in MLT, C/T, and MLC in case of IG, but MLT and MLC increased significantly in case of NG. This is because IG participants tended to revise the texts while maintaining sentence structures they originally had used, whereas some NG participants rather copied and pasted whole MT outputs including more complex sentence structures. Again, there was a large discrepancy among NG participants' performance; some of them could improve the overall quality of their writings with syntactically more elaborated sentences while others deteriorated the consistency of their texts. In addition, the unchanged syntactic complexity measures in IG's products may also imply that instruction could have prevented students from trying out more complex sentence structures and thereby limited their autonomy. Lastly, only IG participants used more AWL words in the revised writings at a significant level after performing MT-assisted revision. This suggests that MT can scaffold students' vocabulary uses in FL writings with proper instruction. Samples of IG and NG participants' original and revised writings are included in Appendix 6 and 7.

4.2. Effects of MT Instruction on Revision Behavior

The second research question of the current study is 'How does MT instruction affect the revision behavior of Korean middle school English

learners in terms of the size, the purpose, and the action type of revision?’

To answer this question, the coding results of each group were compared through Mann-Whitney U test.

4.2.1. Revision Size

Table 4.9 illustrates the frequencies of revisions classified according to their linguistic sizes.

Table 4.9. Comparison of IG and NG: Frequencies of Revisions According to Linguistic Sizes

	IG		NG		<i>Z</i>	<i>p</i>	<i>r</i>
	M	SD	M	SD			
Total	23.56	8.44	10.25	5.69	-4.589	.000	.71
Letter	3.77	2.45	1.85	1.81	-2.510	.011	.39
Word	8.45	4.90	2.35	2.16	-4.554	.000	.70
Phrase	7.55	2.67	2.70	2.30	-4.521	.000	.70
Clause	3.77	3.48	3.35	3.28	-.267	.797	.04

Small effect, $.25 \leq r \leq .39$, medium effect, $.40 \leq r \leq .59$, large effect, $r \geq .60$

The total number of revision was different between the groups at a significant level ($U=38.0$, $Z=-4.589$, $p<.05$) and the effect size was $r=.71$, which is considered large. Moreover, IG participants made more revisions than NG participants in all categories, and the differences were statistically significant ($U=121.5$, $Z=-2.510$, $p<.05$ for Letter, $U=40.0$, $Z=-4.554$, $p<.05$ for Word, $U=41.5$, $Z=-4.521$, $p<.05$ for Phrase) except for clause level ($U=209.5$, $Z=-.267$, $p>.05$). The effect sizes for Word and Phrase revision

were large ($r=.70$ for both). This indicates that the students who had participated in MT instruction prior to using MT performed revisions with higher frequency at various linguistic levels.

Specifically, IG participants made revisions on word level (Mean=8.45) and phrase level (Mean=7.55) a lot more than letter level (Mean=3.77) and clause level (Mean=3.77). This is quite different from NG participants who made clause-level revisions the most (Mean=3.35). This shows that IG participants seemed to be able to revise more on the smaller sizes than clauses unlike NG participants, proving the effects of instruction that taught students the strategies to diversify input sizes depending on their linguistic inquiry. In contrast, the participants without MT instruction edited or changed their original writings mostly in clause level since they tended to copy and paste whole MT outputs rather than using them for comparison to revise letters, words, phrases in their self-written sentences.

4.2.2. Revision Purpose

Table 4.10 presents the frequencies of revisions classified into one of revisional purposes from mechanic correction, lexical exchange, grammatical correction, cosmetic change to informational change.

As in the results from the analysis of revisional size, the frequency of total revisions that IG participants made was significantly higher than that of total revisions that NG participants made ($U=40.0$, $Z=-4.540$, $p<.05$). This shows again that participants who had been provided with MT

instruction performed a greater number of revisions. Regarding the subcategories, IG participants performed more revisions with mechanic, lexical, grammatical and cosmetic purposes than NG participants at a significant level ($U=116.0$, $Z=-2.677$, $p<.05$ for Mechanic, $U=79.0$, $Z=-3.622$, $p<.05$ for Lexical, $U=42.0$, $Z=-4.508$, $p<.05$ for Grammatical, $U=47.0$, $Z=-4.378$, $p<.05$ for Cosmetic). In contrast, NG participants revised texts with informational purpose significantly more than IG participants did ($U=333.0$, $Z=-2.957$, $p<.05$). The effect sizes in each category were all above medium – more than $r=.40$ – according to Plonsky's (2014) field-specific categorization of effect size.

Table 4.10. Comparison of IG and NG: Frequencies of Revisions According to Revisional Purposes

	IG		NG		<i>Z</i>	<i>p</i>	<i>r</i>
	M	SD	M	SD			
Total	23.90	8.97	10.3	5.96	-4.540	.000	.70
Mechanic	2.73	2.21	1.20	1.44	-2.677	.007	.41
Lexical	3.73	3.50	1.05	1.10	-3.622	.000	.56
Grammatical	8.18	3.98	2.15	2.85	-4.508	.000	.70
Cosmetic	8.27	3.33	2.90	2.69	-4.378	.000	.68
Informational	0.95	1.53	3.00	2.92	-2.951	.003	.46

Small effect, $.25 \leq r \leq .39$, medium effect, $.40 \leq r \leq .59$, large effect, $r \geq .60$

Specifically, in case of IG participants, cosmetic change – to change expressions while maintaining the meanings – appeared the most frequently (Mean=8.27), followed by grammatical correction (Mean=8.18), lexical

exchange (Mean=3.73), and mechanic corrections (Mean=2.73). Revision with informational purpose – to add new contents – was performed the least by the participants of this group (Mean=0.95). Examples of the revisions made with different purposes are presented in Excerpt 4.8. Some revisions were made with the purpose of lexical exchange as shown in 1.a) to 2.b). Example 3.a) to 4.b) shows revisions with cosmetic change that the participants changed their expressions for better expressions of the same meaning. The participants corrected grammatical errors especially in verbs, prepositions, and articles as the examples in 5.a) to 8.b) present. Given the fact that EFL learners made errors most frequently on those areas (Lee, 2014), MT-assisted revision with instruction could serve as a helpful feedback tool that can lead middle school learners to correct their grammar errors. Furthermore, some participants changed a common noun ‘*the dog*’ to the pronoun ‘*it*’ or added new pronoun ‘*its*’ and improved the quality of the text by not using the same words redundantly and improving cohesion as in 9.a) to 10.b).

Excerpt 4.8. IG Participants’ Revision with Different Purposes

- 1.a) She said, “Oh, what a cute **dog**!
- b) She said, "Oh, what a cute **puppy**!
- 2.a) The dirty dog looks like very **poor**.
- b) The dirty dog looked very **pitiful**.
- 3.a) When she **came back to her house**, her mom was ready to wash the dog.
- b) When she **returned home**, her mother was ready to wash the dog.

- 4.a) Suji and her mom **decided the dog's name**.
- b) Susie and her mom **came up with a name for the dog**.
- 5.a) The mom **was never thought about live** with a dog ...
- b) The mother **had never thought of living** with the dog ...
- 6.a) The dog **drink** milk fast.
- b) The puppy **drank** milk quickly.
- 7.a) ... so she **decided live** with a dog.
- b) ... so she **decided to live** with **the** dog.
- 8.a) ... and her mother **allowed raising puppy**.
- b) ... and her mother **allowed her to keep it**.
- 9.a) And then, she washed the dog because **the dog** was so muddy.
- b) Then she washed the puppy because **it** was muddy.
- 10.a) The dog seems to be happy to meet new owner.
- b) The dog looks happy to have met **its** new owner.

On the other hand, the frequency of informational change was the highest in case of NG participants (Mean=3.00), followed by cosmetic change (Mean=2.90), and grammatical correction (Mean=2.15). They revised with the purpose of adding new contents the most frequently as in Excerpt 4.9.

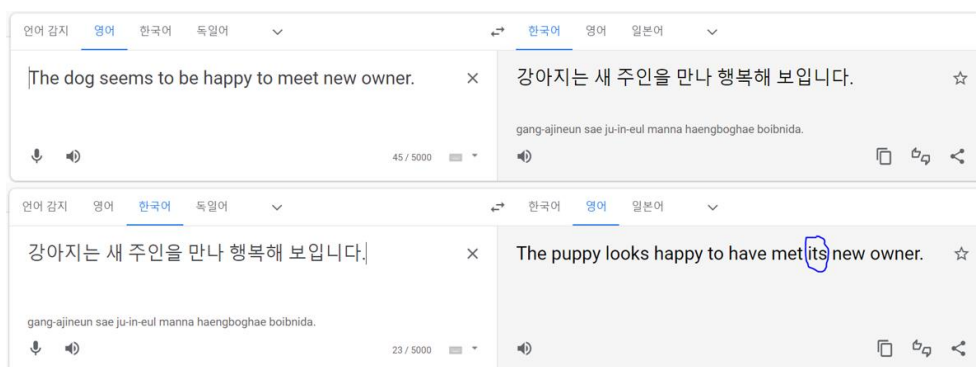
Excerpt 4.9. NG Participants' Revision with Informational Purpose

- 1.a) The girl is so excited that she washes the dog.
- b) Worried that the dog is too dirty and the wound will get infected, the girl hastily washes the dog.
- 2.a) She brings the dog to her house carefully.
- b) Excited at the thought of eating bosintang, the girl carefully takes the dog and brings it home.

Regarding other revisional purposes, patterns of lexical change and cosmetic change of NG were similar to those of IG except for the fact that

the frequencies were much lower. However, regarding grammatical change, interestingly, the use of cohesive markers such as pronouns was not observed in NG participants' revision. Related to this difference, it was assumed that the inverted translation function of MT – one of the key features in the given MT instruction – automatically generated pronouns, stimulating IG participants to use such cohesive markers. Figure 4.1 shows GT output of the sentence 3.k), which produces the pronoun '*its*' automatically as going through the inverted translation.

Figure 4.1. An Example of Inverted Translation



The fact that NG participants revised with informational purpose the most confirmed the researcher's observation that they typed in Korean into MT to obtain its English version and then copied and pasted the whole outputs into their writings rather than reviewing their original sentences by utilizing MT as a revision tool. In other words, they used MT with its original purpose: translation. In contrast, IG participants utilized MT in order to improve their original products by changing to better expressions or

correcting grammatical and mechanic errors. This seems possible by using MT outputs as a point of comparison or individualized corrective feedback. The comparison process can raise the participants' metalinguistic awareness of the target language because they need to interpret linguistic inputs and make related linguistic judgements. The fact that IG participants revised more frequently with more diverse purposes implies that they participated in the revision process more actively, which can positively correlate with their foreign language development (Enkin et al., 2016). Plus, from the perspective of *writing-to-learn* (Cumming, 1990), students' production of written output can promote language learning by utilizing and consolidating their linguistic knowledge despite imperfection of writing outcomes. Although it cannot be said that IG participants had 'noticed' and thereby 'learned' all the grammar points including cohesive markers through MT-assisted revision, the process of revising and rewriting may have potentials that can possibly lead to future learning just as *writing-to-learn* framework believes that learning can occur through the writing process itself even with insufficient linguistic knowledge. In this sense, MT instruction was meaningful in that it promoted students' participation in rewriting self-made sentences with diverse purposes. Still, since this study did not examine the process of students' MT assisted revision, MT's capacity related to the L2 development as well as learning need to be investigated further with future research.

The aforementioned results of revision size and purpose had an

implication when compared with Lee's study (2020), whose participants were college level students. In her study, the participants revised at the word level (referred as 'lexical level' in the study) the most, and with the cosmetic purpose (referred as 'replacing expressions' in the study) the most. This is similar to the revision pattern of IG participants in this study; their revision at word level was the most frequent and the frequency of their revision with cosmetic purpose was the highest. In contrast, the frequency of clause/sentence level revision in Lee's study (2020) was a lot lower and there was no content change made during the revision, meaning that the college students in the study did not revise that much at clause level, nor did they revise with informational purpose, as opposed to NG participants in this study. Considering the participants in her study tended to have higher English proficiency than the participants in this research, instruction played a critical role in scaffolding beginner English learners to perform revision focusing on smaller linguistic sizes with the purpose of changing to better expressions rather than relying on copy and paste strategy in their MT-assisted revision.

4.2.3. Revision Action Type

Table 4.11 illustrates that the frequencies of revisions that were classified according to the following action types that the participants conducted: Addition, Deletion, Substitution, Permutation, Distribution, and Consolidation.

Table 4.11. Comparison of IG and NG: Frequencies of Revisions According to Revisional Actions

	IG		NG		<i>Z</i>	<i>p</i>	<i>r</i>
	M	SD	M	SD			
Total	17.95	5.39	8.75	4.27	-4.461	.000	.69
Addition	2.59	1.40	2.65	2.64	-.562	.582	.09
Deletion	2.68	1.52	0.60	0.94	-4.308	.000	.66
Substitution	7.77	3.45	3.35	1.81	-4.163	.000	.64
Permutation	4.41	2.38	2.00	2.43	-3.383	.000	.52
Distribution	0.23	0.53	0.10	0.31	-.748	.665	.12
Consolidation	0.27	0.63	0.05	0.22	-.508	.611	.08

Small effect, $.25 \leq r \leq .39$, medium effect, $.40 \leq r \leq .59$, large effect, $r \geq .60$

When comparing the frequencies of total actions, IG participants took significantly more revision actions than NG ($U=43.5$, $Z=-4.461$, $p<.05$) and the effect size was large ($r=.69$). IG participants conducted Deletion, Substitution, and Permutation actions more than NG participants at a significant level ($U=53.5$, $Z=-4.308$, $p<.05$ for Deletion, $U=56.0$, $Z=-4.163$, $p<.05$ for Substitution, $U=87.0$, $Z=-3.383$, $p<.05$ for Permutation), yielding more than medium effect sizes ($r=.66$ for Deletion, $r=.64$ for Substitution, $r=.52$ for Permutation), but not the actions of Addition, Distribution, and Consolidation ($U=198.0$, $Z=-.562$, $p>.05$ for Addition, $U=202.0$, $Z=-.748$, $p>.05$ for Distribution, $U=211.0$, $Z=-.508$, $p>.05$ for Consolidation). Overall, in all categories except for addition, IG showed higher frequencies. The most frequently taken action was Substitution for both groups, but IG participants took the actions significantly more than NG participants

($U=56.0$, $Z=-4.163$, $p<.05$). IG participants took the action of Permutation the second most, and Deletion the third most. On the other hand, NG participants took the action of Addition the second most, and Permutation the third most.

The above results demonstrate that the participants given MT instruction were more able to or willing to substitute expressions, permute them to rephrase, or delete unnecessary parts in order to revise their writings compared to the participants who were not provided with such instructions. However, in case of Addition, there were no statistically significant difference between the groups, which implies that both groups were able to utilize MT to add missing parts regardless of instruction. Lastly, the frequencies of Distribution and Consolidation were low for both groups, suggesting none of the groups edited their original texts by distributing or consolidating the sentences that they had written. This tendency may come from the absence of teaching on how to distribute or consolidate sentences effectively during the MT instruction. Or it may also be due to the characteristics of the given task, picture-based writing task, which does not require the participants to compose complex sentences that may need distribution or consolidation.

4.3. Effects of MT Instruction on Students' Perception toward MT

Last research question of this study is ‘How does MT instruction affect Korean middle school English learners’ perception on MT use when revising their L2 writing with MT?’ To find answers for this question, each group students’ response answered based on the 5-point Likert scale to the post questionnaire was analyzed and compared via Mann-Whitney U test. The questions attempted to examine the participants’ perception on MT in terms of usefulness, satisfaction, and intention for future use.

Table 4.12. Post Questionnaire Results

	Groups		<i>Z</i>	<i>p</i>	<i>r</i>
	IG	NG			
Q1. MT was useful for revising my English writing.	3.45	2.15	-3.147	.001	.49
Q1-1. MT is useful for improving the content of English writing.	3.45	1.90	-3.827	.000	.59
Q1-2. MT is useful for using vocabulary of English writing.	3.50	2.55	-2.394	.016	.37
Q1-3. MT is useful for sentence patterns of English writing.	3.45	2.05	-3.743	.000	.58
Q1-4. MT is useful for grammar of English writing.	3.55	1.75	-4.512	.000	.70
Q2. I’m satisfied with the use of MT for this task.	3.59	2.05	-3.846	.000	.59
Q3. I will use MT for the next English writing tasks.	3.55	2.05	-3.622	.000	.56
Small effect, $.25 \leq r \leq .39$, medium effect, $.40 \leq r \leq .59$, large effect, $r \geq .60$					

The questionnaire results are shown in Table 4.12 and notable differences were found. IG participants’ average scores for all questions in the post questionnaire were significantly higher than those of NG. Question

1 asked the participants whether MT had helped their revision process and the results show that the average score from IG participants was 3.45 and that from NG participants was 2.15, which were significantly different ($U=101.0$, $Z=-3.147$, $p<.05$). Its effect size was $r=.49$, which is considered as a medium effect.

Question 1-1 to 1-4 asked for what specific purpose MT had been useful to the participants. Overall, the significantly higher scores in all subcategories of purpose from IG participants showed that they perceive MT's usefulness more positively than NG participants. Specifically, in case of IG participants, the score of Question 1-4 was the highest (Mean=3.55), which indicates IG students thought that MT was the most helpful when checking grammar. It produces the largest effect size, $r=.70$. This result is incongruent with the results of previous studies, in that the participants in Tsai's study (2019) gave the lowest scores on the grammar issue with using MT, and the participants in Lee's study (2020) questioned the grammatical accuracy of MT-generated sentences. When considering NG participants' lowest score on the same question (Mean=1.75), the instruction is considered to be a factor to have brought such a difference. The instruction aimed to train the participants to utilize MT as a revision tool by using MT outputs as a reference source that they can refer to. Provided with such instruction, IG participants were able to use MT output to check grammar in their own sentences, leading to more positive perception on MT's grammatical help. The scores on Question 1-2, asking MT's usefulness on

vocabulary, were the second highest in case of IG participants (Mean=3.50) and the highest in case of NG participants (Mean=2.55). Though the difference between the groups was statistically significant, its effect size was relatively small ($r=.37$). These relatively high scores in vocabulary-related question were consistent with previous studies that reported positive views of study participants on vocabulary search with MT (Bahri & Mahadi, 2016; Correa, 2014; Lee, 2020; Jeong, 2021; Tsai, 2019).

Next, Questions 2 and 3 asked about general satisfaction and intention to reuse MT for their future English writing tasks. Similarly, the scores of IG participants were significantly higher than that of NG participants ($U=78.0$, $Z=-3.846$, $p<.05$ for Question 2 and $U=84.0$, $Z=-3.622$, $p<.05$ for Question 3), yielding medium effect sizes ($r=.59$ for Q2 and $r=.56$ for Q3).

For the two open-ended questions, some common responses emerged. Some IG participants ($n=9$) expressed that they are satisfied especially with grammar checking by using MT for the revision and some ($n=7$) reported that they could search for better expressions or vocabulary with MT. It was impressive that none of IG participants' responses to the open-ended questions reported negative views on MT use. During the MT-assisted revision, they were able to "revise their writings by searching unknown words or correcting errors to improve contents" as one student remarked. On the other hand, only 40% of the participants in NG ($n=8$) answered positively regarding MT use for the revision, and their positive

remarks are limited within word search (n=6). Some of NG participants (n=7) were outspoken critics of MT, whose answers are as follows: ‘using MT is not satisfying because of its frequent errors’, ‘not helpful except for vocabulary searches’, or ‘I could not revise efficiently with MT.’

In conclusion, the effect of instruction on the participants’ perception toward MT was obvious in that all the scores with IG participants were significantly higher than NG participants. This shows MT instruction play a decisive role in making Korean middle school English learners recognize MT as useful for their English revision work.

CHAPTER 5. CONCLUSION

The present study seeks to examine the effects of MT instruction for English writing revisions of Korean middle school English learners. This chapter summarizes the major findings of the present study. First, a summary of major findings and pedagogical implications are presented in Section 5.1. Next, Section 5.2 concludes with reporting limitations and suggestions for future research.

5.1. Major Findings and Pedagogical Implications

The main purpose of the study was to investigate the effects of MT instruction on Korean middle school students when having them use MT as a revision tool for foreign language writing particularly in terms of writing quality, revision behavior and perceptual change. The study was conducted in a middle school in Busan, and 42 students participated either in Instruction Group or No-instruction Group. A mixed-method research design was adopted to examine the gathered data quantitatively and qualitatively.

Regarding the first research question that examined the effects of MT instruction on the qualities of L2 writing revised by the participants, the results of the study revealed that instructional intervention is significantly effective in improving accuracy by reducing spelling errors and grammar errors and allowing the participants to use more academic words. However,

instruction did not increase syntactic complexity since the participants focused on editing their original sentences with partial corrections. On the other hand, without such instruction, NG participants could not detect or correct spelling errors efficiently or even create more grammar errors. Their use of words did not change, and two measurements of syntactic complexity, MLT and MLC, increased significantly because the participants frequently copied and pasted MT outputs to elaborate their original writings rather than to revise them, using MT for literal translation not for revision.

As for the second research question about revision behaviors of Korean middle school students, the results of coding based on revision size, purpose, and action showed that MT instruction is beneficial for encouraging the participants to conduct revision with more various linguistic sizes, purposes and action types. The frequencies of total revisions that the participants made were significantly higher regardless of analysis criteria when they had been provided with MT instruction for English writing revision. This adds evidence on the benefits of MT instruction in that the participants showed more active revisional performances when they learned how to utilize MT to revise their writings through MT instruction. In other words, they showed more engagement during the revision process. This may contribute to students' cognitive engagement by reviewing self-written sentences. Specifically, they revised more on smaller sizes such as letter, words, phrases, unlike the participants who had not been provided with MT instruction and revised on clause level the most. MT instruction

also led the participants to revise their writings with the purposes like correcting grammatical or mechanic errors, changing to better expressions. In contrast, the students who did not participate in MT instruction used MT with the purpose of content change the most. When students were treated with instructional intervention, they performed the actions of substitution, permutation, and deletion significantly more whereas the intervention did not bring difference on the actions of addition, distribution, and consolidation.

In addition, a post questionnaire was performed to find answers for the third research question to examine the participants' perceptual changes on MT use as a revision tool with a variable of MT instruction. The effect of MT instruction was obvious in that the notably high scores were gained for all questions in the questionnaire from the students who participated in MT instruction. They answered more positively to the questions related to MT's usefulness, satisfaction, and intention for future use. Furthermore, the answers for the open-ended questions proved that the students in Instruction Group perceived MT use positively in terms of vocabulary, grammar, or content elaborations. In contrast, not only the scores from the students who did not participate in MT instruction were low but also their answers to the open-ended questions reported negative views on MT use for the revision task.

In sum, this study provides evidence to support the positive effects of MT instruction when Korean middle school students conducted revision

work with the help of MT. Therefore, it can guide teachers and school administrators to decide a design for curriculum and the ways to implement MT instruction in middle school classrooms as Ata and Debreli (2021) asserted in their research. In addition, the results of this study discovered MT's potentials – when it is provided with proper instruction – in a classroom setting especially for middle school students' English writing teaching. Considering it is not easy for EFL teachers to provide appropriate and sufficient feedback during writing lessons, MT can be used as a supplementary tool that provides corrective feedback or linguistic input to middle school learners. Although MT output would not be perfect, it can be useful in some ways as the participants in this study improved their text accuracy and vocabulary use and showed active engagement in the process of revision.

Due to the convenience and availability of MT, it is useless to prevent foreign language learners from using it (Jolley & Maimone, 2015). Rather, it is teachers' responsibility to find a way to help or train students to use it pedagogically. Without such educational intervention, students would merely rely on such technology, which may not lead to language learning or acquisition. In this sense, this study is significant in that it empirically examined the necessity of providing MT instruction as well as presenting MT as a practical educational tool in middle school classroom context.

5.2. Limitations and Suggestions for Future Research

Since this study is not without limitations, several recommendations for future research are discussed. First, due to the small sample size and regional limitation, the results were insufficient to generalize. Therefore, future research with a greater number of participants having diverse regional backgrounds is needed. Second, the participants' foreign language proficiency was not considered. Since one's foreign language proficiency is a critical factor not only for revision performance but also MT use, it is necessary to examine the participants with different English proficiencies in future research. Third, this study investigated only the writing results, not the process. So, future studies need to examine the MT-assisted revision process performed by students. Fourth, among various translators, only GT was used in this study. Although GT was chosen to minimize the prior exposure effect, when translating culturally specific content between Korean and English, *Naver's Papago* is known to be more accurate. So different types of translators may produce different results. Therefore, it is necessary to expand the findings of this study with follow-up studies using different translators. Finally, only two sessions of instruction were designed and implemented in this study, which was a relatively short instructional intervention. Besides, the validity of instruction matters; different formats, procedures, or contents of instruction will bring different results. Therefore, the effects of instruction should be further studied from various angles along

with different forms of instruction.

Despite these limitations, the study proves the effectiveness of MT instruction and the value of MT itself as a pedagogical tool when given proper instruction in Korean middle school classrooms.

REFERENCES

- Ahn, S., & Chung, E. S. (2020). Students' perceptions of the use of online machine translation in L2 writing. *Multimedia-Assisted Language Learning*, 23(2), 10-35.
- Alhaisoni, E., & Alhaysony, M. (2017). An investigation of Saudi EFL university students' attitudes towards the use of Google Translate. *International Journal of English Language Education*, 5(1), 72-82.
- Ata, M., & Debreli, E. (2021). Machine translation in the language classroom: Turkish EFL learners' and instructors' perceptions and use. *IAFOR Journal of Education*, 9(4), 103-122.
- Back, J. (2011). Preposition errors in writing and speaking by Korean EFL learners: A corpus-based approach. *Studies in British and American Language and Literature*, 99, 227-247.
- Bahri, H., & Mahadi, T. S. T. (2016). Google Translate as a supplementary tool for learning Malay: A case study at Universiti Sains Malaysia. *Advances in Language and Literary Studies*, 7(3), 161-167.
- Baraniello, V., Degano, C., Laura, L., Zahonero, M. L., Naldi, M. & Petroni, S. (2016). A wiki-based approach to computer-assisted translation for collaborative language learning. In Li, Y., Chang, M., Kravcik, M., Popescu, E., Huang, R., Kinshuk & Chen, N.-S. (Eds.), *State-of-the-art and future directions of smart learning* (pp. 369-379). Singapore: Springer.
- Bitchener, J., Young, S. & Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing*, 14(3), 191–205.
- Bridwell, L. S. (1980). Revising strategies in twelfth grade students' transactional writing. *Research in the Teaching of English*, 14(3), 197-222.
- Briggs, N. (2018). Neural machine translation tools in the language learning

- classroom: students' use, perceptions, and analyses. *Jalt Call Journal*, 14(1), 2-24.
- Brown, J. D., & Coombe, C. (2015). *The Cambridge Guide to Research in Language Teaching and Learning*. Cambridge: Cambridge University Press.
- Burton, C. (2003). The online translator: Implementing national standard 4.1. *Hispania*, 86(2), 320-321.
- Cancino, M., & Panes, J. (2021). The impact of Google Translate on L2 writing quality measures: Evidence from Chilean EFL high school learners. *System*, 98, 102464. doi:<https://doi.org/10.1016/j.system.2021.102464>
- Case, M. (2015). Machine translation and the disruption of foreign language learning activities. *eLearning Papers*, 45, 4-16.
- Castelvecchi, D. (2016). Deep learning boosts Google Translate tool. *Nature*. doi:10.1038/nature.2016.20696
- Chang, W. L., & Sun, Y. C. (2009). Scaffolding and web concordancers as support for language learning. *Computer Assisted Language Learning*, 22(4), 283-302.
- Chenoweth, N. A., & Hayes, J. R. (2001). Fluency in writing: Generating text in L1 and L2. *Written Communication*, 18(1), 80-98.
- Chung, E. S., & Ahn, S. (2021). The effect of using machine translation on linguistic features in L2 writing across proficiency levels and text genres. *Computer Assisted Language Learning*, 1-26. doi:10.1080/09588221.2020.1871029
- Choi, Y. H. (2007). On-line revision behaviors in EFL writing process. *English Teaching*, 62(4), 69-93.
- Chon, Y. V., Shin, D., & Kim, G. E. (2021). Comparing L2 learners' writing against parallel machine-translated texts: Raters' assessment, linguistic complexity and errors. *System*, 96, 102408. doi:<https://doi.org/10.1016/j.system.2020.102408>
- Clifford, J., Merschel, L., & Munné, J. (2013). Surveying the landscape: What is the role of machine translation in language learning. *Research*

in Education and Learning Innovation Archives, 10, 108-121.

- Correa, M. (2014). Leaving the “peer” out of peer-editing: Online translators as a pedagogical tool in the Spanish as a second language classroom. *Latin American Journal of Content and Language Integrated Learning, 7*(1), 1-20.
- Crossley, S. A., & McNamara, D. S. (2009). Computational assessment of lexical differences in L1 and L2 writing. *Journal of Second Language Writing, 18*(2), 119-135.
- Cumming, A. (1990). Metalinguistic and ideational thinking in second language composing. *Written Communication, 7*(4), 482-511.
- Ducar, C., & Schocket, D. H. (2018). Machine translation and the L2 classroom: Pedagogical solutions for making peace with Google Translate. *Foreign Language Annals, 51*(4), 779-795.
- Enkin, E., & Mejías-Bikandi, E. (2016). Using online translators in the second language classroom: Ideas for advanced-level Spanish. *Latin American Journal of Content and Language Integrated Learning, 9*(1), 138-158.
- Falvey, P. (1993). *Towards a description of corporate text revision*. (Unpublished doctoral dissertation). University of Birmingham, Birmingham, UK.
- Faigley, L., & Witte, S. (1981). Analyzing revision. *College Composition and Communication, 32*(4), 400-414.
- Ferris, D., Liu, H., & Rabie, B. (2011). The job of teaching writing: Teacher views of responding to student writing. *Writing and Pedagogy, 3*(1), 39-77.
- Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition, 18*(3), 299-323.
- Fredholm, K. (2015). Online translation use in Spanish as a foreign language essay writing: Effects on fluency, complexity and accuracy.

- Revista Nebrija de Lingüística Aplicada a la Enseñanza de Lenguas*, 18, 7-24.
- Fritz, C., Morris, P., & Richler, J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology General*, 141(1), 2-18.
- Garcia, I., & Pena, M. I. (2011). Machine translation-assisted language learning: writing for beginners. *Computer Assisted Language Learning*, 24(5), 471-487.
- Giannetti, T. R. (2016). *Google translate as a resource for writing: A study of error production in seventh grade Spanish*. (Master's thesis). St. John Fisher College, Rochester, U.S.A.
- Google AI Blog. (2016). *A neural network for machine translation, at production scale*. Retrieved from <https://ai.googleblog.com/2016/09/a-neural-network-for-machine.html>.
- Groves, M., & Mundt, K. (2015). Friend or foe? Google Translate in language for academic purposes. *English for Specific Purposes*, 37, 112-121. doi:<https://doi.org/10.1016/j.esp.2014.09.001>
- Hamza, H. A., & Saadalla, F. J. (2021). Using English language learning mobile applications among Iraqi EFL university learners. *Review of International Geographical Education Online*, 11(5), 3507-3516.
- Hall, C. (1987). *Revision strategies in L1 and L2 writing tasks: a case study*. (Unpublished doctoral dissertation). The University of New Mexico, Albuquerque, U.S.A.
- Hall, C. (1990). Managing the complexity of revising across languages. *TESOL Quarterly*, 24(1), 43-60.
- Haukelid, K. F. (2021). *Google Translate: A helpful aide? A mixed method sequential explanatory study on the usage and effects of Google Translate in three Norwegian EFL classes*. (Master's thesis). University of Stavanger, Stavanger, Norway.
- Im, H.-J. (2017). The university students' perceptions or attitudes on the use

- of the English automatic translation in a general English class – Based on English writing lessons. *Korean Journal of General Education*, 11(6), 727-751.
- Jo, I. (2018). The effect of machine translation on English writing instruction: Focusing on the writing self-efficacy and the writing quantity. *The Jungang Journal of English Language and Literature*, 60(2), 253-279.
- Jolley, J. R., & Maimone, L. (2015). Free online machine translation: Use and perceptions by Spanish students and instructors. In A.J. Moeller (Ed.), *Learn language, explore cultures, transform lives* (pp.181-200). Egg Harbor, WI: Central States Conference on the Teaching of Foreign Languages.
- Jeong, N.-S. (2021). A study on the effects of machine translators on college students' writing proficiency and affective attitude. *Multimedia-Assisted Language Learning*, 24(1), 134-157.
- Kang, S., & Lee, J.-H. (2019). Are two heads always better than one? The effects of collaborative planning on L2 writing in relation to task complexity. *Journal of Second Language Writing*, 45, 61-72. doi:<https://doi.org/10.1016/j.jslw.2019.08.001>
- Kim, S. (2019). Playing with machine translation in language classroom: Affordances and constraints. *Multimedia-Assisted Language Learning*, 22(2), 9-28.
- Kliffer, M. (2005). An experiment in MT post-editing by a class of intermediate/advanced French majors. In *EAMT 10th Annual Conference* (pp. 160-165).
- Kol, S., Schcolnik, M., & Spector-Cohen, E. (2018). Google Translate in Academic Writing Courses? *The EuroCALL Review*, 26(2), 50-57.
- Kazemzadeh, A., & Kashani, A. (2014). The effect of computer-assisted translation on L2 learners' mastery of writing. *International Journal of Research Studies in Language Learning*, 3, 29-44

- Kormos, J. (2011). Task complexity and linguistic and discourse features of narrative writing performance. *Journal of Second Language Writing*, 20(2), 148-161.
- Larson-Hall, J. (2010). *A Guide to Doing Statistics in Second Language Research Using SPSS*. New York, NY: Routledge.
- Lee, C., Wong, K. C., Cheung, W. K., & Lee, F. S. (2009). Web-based essay critiquing system and EFL students' writing: A quantitative and qualitative investigation. *Computer Assisted Language Learning*, 22(1), 57-72.
- Lee, H., & Choi, Y. H. (2019). Effects of task complexity and cooperative learning on Korean middle school students' English writing based on model texts and activity preference. *Journal of Research in Curriculum & Instruction*, 23(5), 487-503.
- Lee, S. (2014). Grammatical error patterns in EFL students' writing across different language proficiency levels. *Foreign Languages Education*, 21(1), 1-29.
- Lee, S.-B. (2018). Process research into post-editing: How do undergraduate students post-edit the output of Google Translate? *The Journal of Translation Studies*, 19(3), 259-286.
- Lee, S., & Kim, S. (2018). Pre-editing rules to Enhance output quality of machine translation: English-Korean and Korean-English. *The Journal of Translation Studies*, 19(5), 121-154.
- Lee, S. M. (2020). The impact of using machine translation on EFL students' writing. *Computer Assisted Language Learning*, 33(3), 157-175.
- Lee, S. M., & Briggs, N. (2020). Effects of using machine translation to mediate the revision process of Korean university students' academic writing. *ReCALL*, 33(1), 18-33.
- Lee, Y.-J. (2021). Still taboo? Using machine translation for low-level EFL writers. *ELT Journal*, 75(4), 432-441.

- Lewis-Kraus, G. (2016). The great A.I. awakening. *The New York Times Magazine*, 14(12), 2016.
- Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly*, 45(1), 36-62.
- Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15(4), 474-496.
- Luton, L. (2003). If the computer did my homework, how come I didn't get an "A"? *The French Review*, 76, 766-770.
- McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated Evaluation of Text and Discourse with Coh-Metrix*. Cambridge: Cambridge University Press.
- Min, H.-T. (2006). The effects of trained peer review on EFL students' revision types and writing quality. *Journal of Second Language Writing*, 15(2), 118-141.
- Ministry of Education. (2015). *2015 revised national English curriculum* (2015-74-14). Sejong, Korea: Ministry of Education.
- NEAT Formula Speaking Level 2 (2012). Seoul: YBM.
- New, E. (1999). Computer-aided writing in French as a foreign language: A qualitative and quantitative look at the process of revision. *The Modern Language Journal*, 83(1), 81-97.
- Niño, A. (2007). MT post-editing: A text repair experience for the foreign language class. *Cadernos de Tradução*, 1(19), 115-131.
- Niño, A. (2008). Evaluating the use of machine translation post-editing in the foreign language class. *Computer Assisted Language Learning*, 21(1), 29-49.
- Niño, A. (2009). Machine translation in foreign language learning: Language learners' and tutors' perceptions of its advantages and disadvantages. *ReCALL*, 21(2), 241-258.

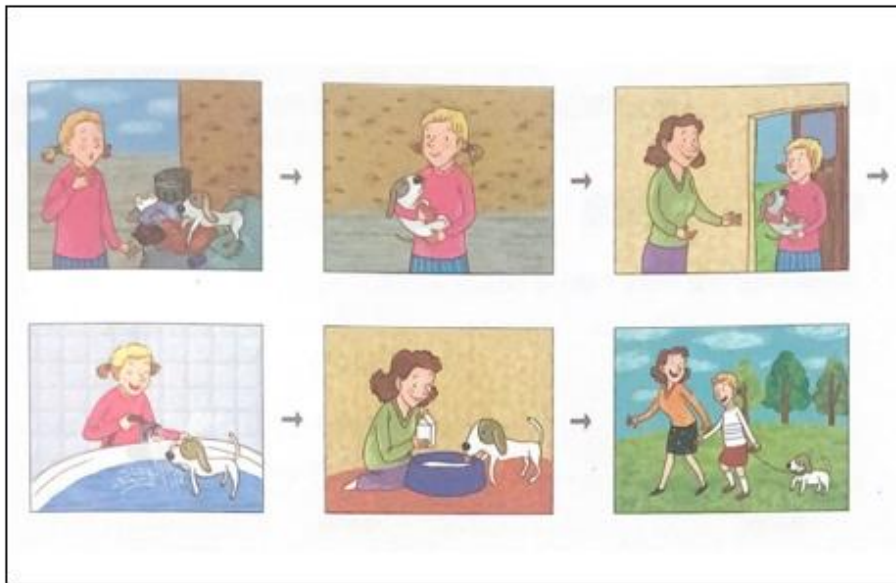
- Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30(4), 555-578.
- O'Neill, E. M. (2016). Measuring the impact of online translation on FL writing scores. *IALLT Journal of Language Learning Technologies*, 46(2), 1-39.
- Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning*, 64(4), 878-912.
- Rye, H. (2010). What's the problem in teaching English writing in secondary level? *The Jungang Journal of English Language and Literature*, 52(3), 173-197.
- Somers, H. (2003). Machine translation in the classroom. In H. Somer (Ed.), *Computers and translation: a translator's guide* (pp. 319-340). Amsterdam: John Benjamins.
- Somers, H., Gaspari, F., & Niño, A. (2006). *Detecting inappropriate use of free online machine translation by language students. A special case of plagiarism detection*. Paper presented at the 11th Annual Conference of the European Association for Machine Translation, Oslo, Norway.
- Sommers, N. (1980). Revision strategies of student writers and experienced adult writers. *College Composition and Communication*, 31(4), 378-388.
- Steding, S. (2009). Machine translation in the German classroom: Detection, reaction, prevention. *Teaching German*, 42(2), 178-189.
- 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>
- Stevenson, M., Schoonen, R., & De Glopper, K. (2006). Revising in two languages: A multi-dimensional comparison of online writing revisions in L1 and FL. *Journal of Second Language Writing*, 15(3), 201-233.

- Sukkhwan, A., & Sripetpun, W. (2014). Use of Google Translate: A survey of Songkhla Rajabhat University students. In *L-SA Workshop and Colloquium: "Speaking" for ASEAN* (pp. 88-104).
- Tavakoli, P., & Foster, P. (2011). Task design and second language performance: The effect of narrative type on learner output. *Language Learning*, 58(2), 439-473. <https://doi.org/10.1111/j.1467-9922.2008.00446.x>
- Tsai, S.-C. (2017). Effectiveness of ESL students' performance by computational assessment and role of reading strategies in courseware-implemented business translation tasks. *Computer Assisted Language Learning*, 30(6), 474-487.
- Tsai, S.-C. (2019). Using google translate in EFL drafts: a preliminary investigation. *Computer Assisted Language Learning*, 32(5-6), 510-526.
- 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. doi:10.1080/09588221.2020.1799412
- Van Praag, B., & Sanchez, H. S. (2015). Mobile technology in second language classrooms: Insights into its uses, pedagogical implications, and teacher beliefs. *ReCALL*, 27(3), 288-303.
- White, K. D., & Heidrich, E. (2013). Our policies, their text: German language students' strategies with and beliefs about web-based machine translation. *Teaching German*, 46(2), 230-250.

APPENDICES

Appendix 1.

A Picture-based Narrative Task



Appendix 2.

A Background Questionnaire

A background questionnaire	
Previous English learning experience	3-2. Which MT do you use for your English writing?
1. How have you learned English so far?	<ul style="list-style-type: none">- Google Translate- Papago- Other
<div style="border: 1px solid black; height: 70px; width: 300px;"></div>	
2. Have you studied English in English speaking countries?	4. How much do you agree with the following statements?
<ul style="list-style-type: none">- Yes- No	MT is a useful tool for English writing.
2-1. In case when you answer 'yes' to the above question, where did you study? And how long have you studied there?	<ul style="list-style-type: none">- Strongly agree- Somewhat agree- Somewhat Disagree- Strongly disagree
<ul style="list-style-type: none">- Where: _____- How long: _____	5. Please describe freely why you think so.
Previous experience on MT use	<div style="border: 1px solid black; height: 70px; width: 220px;"></div>
3. I have previously used MT for English writing.	
<ul style="list-style-type: none">- Yes- No	
3-1. In case when you answer 'yes' to the above question, how many times have you used MT for your English writing?	
<ul style="list-style-type: none">- Everyday- Less than once a week- Less than once a month	

Appendix 3.

A Post Questionnaire

A post questionnaire	
How much do you agree with the following statements?	
Choose the number that represents below.	
① Strongly agree, ② Somewhat agree, ③ Agree, ④ Somewhat disagree, ⑤ Strongly disagree	
1. I've found GT was helpful in revising my writing.	5. GT is accurate in the <u>grammar</u> of English writing.
① ② ③ ④ ⑤	① ② ③ ④ ⑤
2. GT is helpful for <u>content improvement</u> in English writing.	6. I was satisfied with using GT for today's task.
① ② ③ ④ ⑤	① ② ③ ④ ⑤
3. GT is helpful for <u>vocabulary use</u> in English writing.	Please explain freely: I was / was not satisfied with using MT for today's writing task because...
① ② ③ ④ ⑤	<div></div>
4. GT is helpful for the use of <u>sentence pattern</u> in English writing.	7. I will use GT for English writing in the future.
① ② ③ ④ ⑤	① ② ③ ④ ⑤
	Please explain freely: I will use/not use MT for the revision of English writing in the future because...
	<div></div>

Appendix 4.

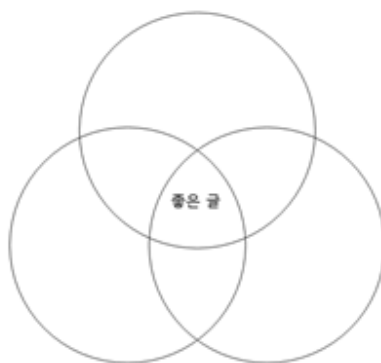
Instructional Material for Instruction Group 1

『한국 중학생 영작문 수정에 번역기 활용 연구』에 관한 수업자료

수업 목표: 영작문 수정에 구글번역기를 활용하는 방법에 대해 이해하고 배운 것을 활용
해 문장의 오류를 수정할 수 있다.

▶ 좋은 글의 요건: 더 좋은 글을 쓰기 위해 영작문을 ‘수정한다’는 의미는?

- ① 어법상 틀린 문장 고치기
- ② 내용을 풍성하게 보충하기
- ③ 글의 흐름 매끄럽게 하기



▶ 영작문 수정에 번역기를 활용하는 방법

① 어법상 틀린 문장 고치기

1. 내가 쓴 영어 문장을 한글로 번역하여 _____이 되는지 확인한다.
2. _____을 거쳐 오류를 찾는다.
3. 내가 쓴 영어 문장과 번역기 결과물을 _____어법 오류를 수정한다.

② 내용을 풍성하게 보충하기 & ③ 글의 흐름 매끄럽게 하기

4. 내가 표현하고자 하는 의미를 ‘_____ 한국어’ 문장으로 입력한다.
5. 번역기 결과물 영어 문장을 내가 이해할 수 있는지 확인한다.
6. 이해할 수 있다면, _____을 거쳐 오류를 찾고 수정한다.
7. 이해할 수 없다면, 입력하는 한글 문장을 더 ‘간단하게’ 수정하여 내가 이해할 수 있는 영어 문장이 나올 수 있도록 한다.

Appendix 5.

Instructional Material for Instruction Group 2

<Pair Work>

- ▶ 짝과 함께 구글번역기를 이용해 다음 문장들을 수정해봅시다.

The school have good facilities.
Eating sounds is relaxing.
We discussed about the issue.
It can be used to search information.
My dream is to become doctor.
I can imagine the another world.
When I was young, my dream is to become MC.
My favorite hobby is reading novel.
People could use it easily and fastly
The sound is very funny and relaxing.
It is more interesting to imagine mystic world.
I do these activites usually at night.

<Group Work>

- ▶ 모둠원들과 함께 구글번역기를 이용해 다음 글을 수정해 봅시다.

Let me introduce my hobby. When I was young my parents took me to the soccer stadium. From then on, my hobby is playing and watching soccer. I think soccer is the most interesting sports. First it can't play alone, only with a team. Second, soccer game is very thrilled, some people think that soccer game is boring, but in fact it is not. When my team get a first score, My feeling is gonna be happier and happier. Third, The rule is very simple, soccer game's rule is so simple that first person that watching soccer game can understand what is going on. That's why I like the soccer. Thank you for reading and I suggest you strongly. How about watching soccer?

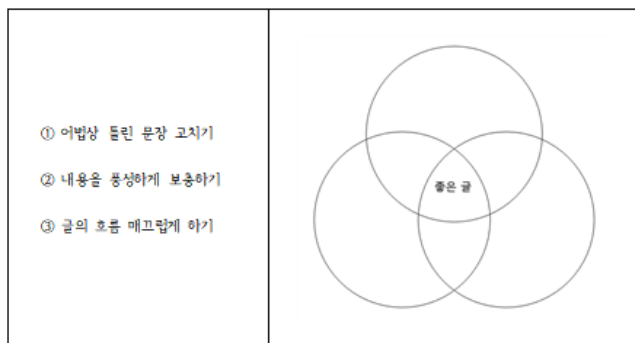
Appendix 5.

Instructional Material for No-instruction Group

*한국 중학생 영작문 수정에의 반역기 활용 연구에 관한 수업자료

수업 목표: 1. 영작문 수정의 의미를 이해하고 문장과 글을 수정할 수 있다.
2. 구글번역기를 사용할 수 있다.

▶ 좋은 글의 요건: 더 좋은 글을 쓰기 위해 영작문을 '수정한다'는 의미는?



<Pair Work>

▶ 짝과 함께 다음 문장들을 수정해 봅시다.

The school have good facilities.
Eating sounds is relaxing.
We discussed about the issue.
It can be used to search information.
My dream is to become doctor.
I can imagine the another world.
When I was young, my dream is to become MC.
My favorite hobby is reading novel.
People could use it easily and fastly
The sound is very funny and relaxing.
It is more interesting to imagine mystic world.
I do these activites usually at night.

<Group Work>

▶ 모둠원들과 함께 다음 글을 수정해 봅시다.

Let me introduce my hobby. When I was young my parents took me to the soccer stadium. From then on, my hobby is playing and watching soccer. I think soccer is the most interesting sports. First it can't play alone, only with a team. Second, soccer game is very thrilled, some people think that soccer game is boring, but in fact it is not. When my team get a first score, My feeling is gonna be happier and happier. Third, The rule is very simple. soccer game's rule is so simple that first person that watching soccer game can understand what is going on. That's why I like the soccer. Thank you for reading and I suggest you strongly. How about watching soccer?

Appendix 6.

Samples from an IG Participant

- From student 1

1. Original Writing

One day, the girl found a dog. The dirty dog looks like very poor. The girl picked up a dog, and brang at the girl's house. The girl said the mom 'I found a poor dog. Please allowed wash a dog'. The girl's mom allowed a dog come in the house. The girl washed a dirty dog. The mom was never thought about live with a dog, but a dog looks live very poor. The mom noticed a dog seems like hungry, so mom give a milk to dog. The mom thought a dog can't live at the outside, so she decided live with a dog. The dog became a family, and live happily ever after.

2. Revised Writing

One day, the girl found a dog. The dirty dog looked very pitiful. The girl took the dog and brought it home. The girl said to her mother, 'I found a poor dog. Please allow me to wash the dog'. The girl's mother allowed the dog to enter the house. The girl washed the dirty dog. The mother had never thought of living with a dog, but the dog looked very pitiful. The mother noticed that the dog looked hungry, so she gave the dog milk. The mother thought that the dog couldn't live outside, so she decided to live with the dog. The dog became part of their family, and they lived happily ever after.

- From student 2

1. Original Writing

One day, a girl named Suji found a dog which was in the trash. Suji thought that the dog was thrown away, So she took the dog and went to her home. After she went to her home, her mom is happy because she likes dog. She said to Suji that wash the dog. So Suji washed the dog and gave some milk. The dog drank it quickly. When the dog drank it all, Her mom and Suji took a walk with the dog.

2. Revised Writing

One day, a girl named Suji found a dog abandoned in a trash can. Suji thought the dog had been abandoned. So she took the dog home. After she went to her home, her mom was happy because she likes dogs. She told Suji to wash the dog. So Suji washed the dog and gave it some milk. The dog drank it quickly. When the dog finished drinking, her mother and Suji went for a walk with the dog.

Appendix 7.

Samples from NG Participants

- From student 1

1. Original Writing

Grace was going to the convieant store to buy some snaks. She was going to ---- so she stoped. Suddenly she hard strangy sound behind her. Behind her there was trash. But the sound didn't stoped. So she looked carfully what made the sound and there was a small dog. The dog was very small and darty. She deside to keep the dog and wash it and feed it. She went bake to her home. Grace's mom was suprized, but hre mom agreed to keep the dog. She washed the dog and feed the dog with her mom. They leaved happy for long thime.

2. Revised Writing

Grace was going to the convenience store to buy some snaks. She stopped at the crosswalk waiting for the traffic light to change. Suddenly she heard strangy sound behind her. But behind her was only a pile of garbage. However the sound didn't stoped. So she looked carefully what made that strange sound, surprisingly, there was a small dog. The dog was very small and darty. So she deside to keep the dog. She went back to her home. Grace's mom was a little suprized, but hre mom agreed to keep the dog. She washed the dog and feed the dog with her mom. Grace and her mom loved the dog very much, so the dog loved them and followed them well. They lived happy for a long time with love

- From student 1

1. Original Writing

The dog is thrown away outside. He sits with trash in a street. Then the girl walks that street. She doesn't let him. She brings the dog in her house. In a house, her mom welcomes them. She loves him. The girl washes the dog in the bathroom. And the girl's mom feed him with milk. Lastly, the girl and her mom turns their clothes, they walk in the street with together. They are live a long day with the dog.

2. Revised Writing

The dog is thrown away outside. He sits with trash in a street. Meanwhile, the girl walked down that street since she had an appointment. She doesn't let him. She brings the dog in her house. At house, her mom welcomed them when she saw them coming. She loves him. The girl washes the dog in the bathroom. The girl's mom wanted to feed him, thus she gave him some milk which she had bought from a nearby supermarket. Lastly, the girl and her mom turns their clothes, they walk in the street with together. After that, they live together all their lives.

국 문 초 록

기계번역기술이 급격히 발전하며 번역기(MT)를 교육적으로 활용하는 것에 대한 논의가 이루어지고 있으나 여전히 중등 학교 교실에서의 번역기 사용에 대해서는 거의 알려져 있지 않다. 따라서 본 연구는 한국 중학교 영어 학습자들의 번역기 사용을 연구하고자 한다. 또한, 어린 학습자들일수록 번역기 활용에 관한 적절한 지도의 필요성이 크기 때문에, 본 연구는 번역기 활용에 관한 수업이 한국 중학생들의 영어 글쓰기 수정에 미치는 영향을 조사하고자 한다.

본 연구는 부산의 한 중학교에서 진행되었으며, 2개 반, 총 42명의 학생이 그림 기반 글쓰기 과제와 수정 과제를 수행하였다. 글쓰기 과제 후, 실험군 학급은 두 차시에 걸쳐 영작문 수정에의 번역기 활용에 관한 수업을 받았다. 수업은 번역기에 관한 설명, 연구자의 시연, 그리고 연습 활동으로 이루어졌다. 이에 반해 대조군 학급은 번역기를 자유롭게 사용할 수 있도록 노출되었을 뿐 관련 수업은 제공되지 않았다.

번역기 수업의 영향을 연구하기 위해 참가자들의 영작문 결과물과 수정 양상, 그리고 번역기에 대한 참가자들의 인식을 윌콕슨 부호순위 검정과 맨 휘트니 검정을 사용해 분석하였다.

결과 분석에 의하면, 실험군의 경우 철자 오류 및 문법 오류가 크게 줄어든 반면 대조군의 경우 문법 오류만 감소하였다. 또 실험군의 경우 T단위 평균 길이 (mean length of T-unit), T단위 당 절의 수 (clauses per T-unit), 절 평균 길이 (mean length of clauses) 모두 변화가 없었던데 반해, 대조군의 경우 T단위 평균 길이와 절 평균 길이가 유의미하게 증가했다. 다만 대조군의 이러한 향상은 개인차가 컸다. 어휘 사용과 관련해서 실험군의 경우 학술어휘

(AWL)사용이 유의미하게 증가한 반면, 대조군의 경우에는 단어 사용에 변화가 없었던 것으로 나타났다. 이러한 차이점들은 중학교 학습자들이 번역기를 사용하여 철자 및 문법 오류를 줄이고 보다 학문적인 단어를 사용하여 글을 수정하는데 번역기 활용 수업이 효과적이었음을 시사한다.

다음으로 학생들의 영어 글쓰기 수정 양상에 대한 결과를 살펴보면, 실험군 학생들이 대조군 학생들에 비해 모든 범주에서 유의미한 수준의 높은 빈도를 보여주었다. 단순히 수정 빈도가 높은 것 외에도, 수정 시 특정 수정 단위, 목적, 활동에 제한되었던 대조군 학생들에 비해 실험군 학생들은 다양한 단위의 수정 단위(문자, 단어, 구 등)와 다양한 목적(철자, 어휘, 문법, 표현, 및 내용 추가), 그리고 다양한 활동(추가, 삭제, 대체, 융합)을 보여주었다. 이러한 결과는 번역기 활용 수업이 한국 중학생 학습자들이 보다 적극적이고 다양한 수정 과정에 참여하도록 이끄는 데 성공적이었음을 보여준다.

학습자들의 번역기에 대한 인식과 관련하여, 번역기의 유용성, 번역기에 대한 만족도, 향후 번역기 재사용 의향을 살펴본 질문들에 실험군 학습자들은 대조군 학습자들에 비해 훨씬 긍정적인 반응을 보여주었다. 이는 번역기 활용 수업이 학습자들로 하여금 번역기에 대해 긍정적인 인식을 갖도록 하는데 효과적이었음을 나타낸다.

본 연구는 표본 크기와 방법론적 한계를 내재함에도 불구하고, 번역기가 중학교 교실 현장에서 영어 학습을 위한 교육적 도구로 어떻게 활용될 수 있는지에 대한 통찰을 제공한다. 연구 결과를 토대로 교육적 함의와 향후 연구에 대한 제안도 논의하고자 한다.

주요어: 기계번역, 구글번역기, 수업, 영어 작문, 영작문 수정, 컴퓨터 지원 언어 학습

학번: 2020-28070