IJACT 22-9-10

An Analysis of Semantic Errors in Machine-Translated English Compositions by Korean EFL College Students

[1]Baek, Ji-Yeon

[1]Prof., Dept. of English Education, Seowon Univ., Korea
bjy0429@seowon.ac.kr

Abstract

The purpose of this research is to investigate the types of semantic errors made by MT in translating EFL college students' original drafts written in Korean into English. Specifically, this study attempts to find out 1) what types of semantic errors are most frequently committed by MT? and 2) how students feel about the quality of the MT-produced output? The findings from this study indicated that MT produced the errors related to accuracy (47%) the most, followed by the errors related to fluency and ambiguity (14.6% respectively). Students were well aware of the errors with accuracy and fluency but had limited ability to check the errors with ambiguity. Based on the findings, this study suggests pedagogical implications which can be implemented in L2 writing classrooms.

Keywords: EFL learners, Machine translation, English composition, Error analysis, Semantic errors

1. INTRODUCTION

The emergence of artificial intelligence (AI) and big data has led to the rapid development of machine translation (hereafter MT) capabilities. In recent years, MT as a state-of-the-art tool has begun to find its way into a wide range of fields including international business, entertainment, traveling, as well as education. Most importantly, great attention has been given to the integration of MT tools into language teaching and learning. As a matter of fact, a number of researchers and language instructors have started to highlight the benefits of using MT as a pedagogical tool in L2 classrooms [1-7]. Comparing with dictionaries, which provide definitions and usages of a word, MT can provide more natural and sensitive translation outputs as it translates from big data using AI [8]. Lee [4], who compared students' initial versions of L2 writing and the revised versions with MT assistance, pointed out that MT can play a crucial role to enhance the quality of students' writing and their writing skills. Nevertheless, it does not mean that MT produces entirely accurate translation results. Therefore, there has been a great concern about the quality of translations produced by MT, and the focus of many recent studies has been on the classification and analysis of the errors in machine-generated outputs to promote the quality of MT [9-13]. The issue of errors made by MT has been continuously drawn, but there are still limited studies on error classification and analysis focused on semantic errors in the MT outputs. Therefore, the present study aims to identify and classify the types of

semantic errors made by MT in translating students' initial drafts written in Korean into English. It also attempts to make pedagogical implications which can be implemented in L2 writing classrooms. Specifically, this study focuses on 1) what types of semantic errors are most frequently made by MT, and 2) how students feel about the quality of the MT-produced outputs.

2. LITERATURE REVIEW

Translating a language into another language is a complex process, which involves at least two different languages (the source language and the target language) and two different cultures. The process of translation generally includes three steps: analyzing the source text (pre-translation), transferring it into the target language (translation), and reviewing or reconstructing the translated output (post-translation) [14]. The human or manual translation process usually involves the above-mentioned process as a sequence of three basic steps. However, these basic steps for translation cannot be carried out by machine translators. Post-translating the MT-generated outputs is still needed by the human translator. Considering the fact that the MT output is still in need of human touch, a number of researchers in this field have concentrated on identifying, analyzing, and classifying the errors made by MT to promote the quality of the translations [9-13].

As for the quality of the MT results, it was claimed that one of the big differences between human translation and machine translation is the style of writing, and it may be difficult to narrow this gap with the current translation technology in a short period of time [10]. Ahn [11], who evaluated English to Korean translations, reported that quite a few semantic and syntactic errors were included in the machine-translated results. She pointed out that the errors made by MT were mainly due to the differences in the linguistic structure between the source language (Korean) and the target language (English). Simply put, the greater the structural or grammatical difference between the two languages, the lower the quality of the MT results. In some of the recent studies focused on the quality of translations and error recognition, it was revealed that lexical errors were most frequently made by MT [15-16]. Additionally, it was pointed out that those lexical errors, which can be considered as the improper use of words or phrases in a certain context, seem to be caused by contextual confusion between the two languages [17].

Quite a number of previous research studies have been conducted to identify the types of syntactic and semantic errors made by MT along with their causes. However, there are only a few studies that focus on the specific types of semantic errors and guidelines for their application in a real-world classroom setting. In addition, little research has been exploring how EFL learners perceive the MT-produced results in terms of the raw translation quality. Thus, this study attempts to investigate the types of semantic errors in Korean-to-English translations and EFL learners' views on the quality of the MT-generated outputs. Furthermore, the findings of this study would be expected as a reference for the future use of machine translators and as guidelines for the post-translation activity in teaching and learning languages.

3. METHODOLOGY

3.1 Participants

A total of seven Korean EFL college students (two females, five males) participated in this study. All participants were junior students majoring in English at a four-year university located in the central region of Korea. Sixteen students were taking a class 'Field English and Experience', and seven students out of 16 were selected purposefully by obtaining students' consents in advance. According to the semi-structured interview which was conducted at the beginning of the class, it was revealed that most of the participants

seemed to have an intermediate level of English competence (which is approximately equivalent to TOEIC 600-700). However, all the participants did not have enough confidence in writing a structured essay in both Korean and English.

3.2   Data Collection

Students were trained on how to structure an essay in Korean, translate the Korean version into English by using MT, then post-edit the MT results. They have completed three writing tasks throughout the semester following the three steps of translation. After each task was completed, they were required to submit all the drafts produced during the writing tasks. Thus, each student submitted three drafts (1) initial draft written in Korean, 2) MT-produced English translation, and 3) post-edited version of the MT output), and finally ended up submitting nine drafts throughout the semester. Thus, a total of 63 writing results were collected; however, only the MT-produced English drafts, 21 out of 63 writing samples, were analyzed in this study.

The other data for this study were self-reflective journals written by seven students during the semester. As most of them were not familiar with the reflective journal, how and what to write, and the aims of it were thoroughly explained to the students in the first week of the semester. Then, students were guided to write and submit their self-reflective journals during the last class of each week except orientation week, and mid-term and final examination weeks. Thus, a total number of 84 reflective journals (12 journals from each student) from seven students were collected. They were questioned on these three units as follows: 1) learning achievement from the writing activity with MT assistance, 2) views on the quality of the MT-generated output, and 3) plans for post-editing the MT output. The weekly submitted journals were checked by the researcher and proofread by another colleague, and overall feedback on specific themes was given to the entire class each week. The journals were written in the students' first language (Korean) in order to make students feel more comfortable writing their reflections on the activities, and they were translated into English by the researcher.

3.3   Data Analysis

Table 1. Classification of Semantic Errors (adopted from Han (2021))

Criteria	Content	Indicators
Accuracy	Is the meaning of the original text clearly translated?	mistranslations, additions, omissions, distortions, differences in nuances
Fluency	Is the translated text natural?	unnatural collocation or idioms, and parts of speech or style
Cohesion	Is the translated text naturally and logically connected?	use of inappropriate directives, pronouns, connecting words, and conjunctions
Coherence	Are the terms and expressions consistently used?	inconsistent use of terms and expressions
Ambiguity	Are there any terms or expressions which can be understood in more than one way?	use of ambiguous vocabulary and sentence structure
Conciseness	Is there any unnecessary repetition?	unnecessary repetition and translation
Functionality	Are the appropriate expressions used for the relevant field?	use of inappropriate vocabulary, expression, and writing style

In this study, semantic errors were defined as a concept that encompasses internal knowledge about the message to be transferred. In other words, seven categories of analyzing semantic errors were identified: accuracy, fluency, cohesion, coherence, ambiguity, conciseness, and functionality. In order to gain the answers to the research questions, the MT-produced drafts and students' self-reflective journals were analyzed, both quantitatively and qualitatively. As for the quantitative data analysis, Han's error classification described in Table 1 below was employed to sort out the semantic errors in the raw drafts produced by MT [18].

The frequency and percentage of the classified errors were calculated through SPSS Statistics Version 22. Additionally, the data drawn from self-reflective journals written by seven students were qualitatively analyzed by careful perusal to examine if there was any interrelationship between the findings from the journals and descriptive analysis. Specifically, the detailed description of certain aspects or issues in terms of the quality of the MT results was drawn from the raw data of the journals and reviewed by another colleague who is in the same research field.

4. RESULTS AND DISCUSSION

The MT-generated outputs from three different writing tasks throughout the semester had been collected. In order to investigate the most frequent types of semantic errors committed by MT, a part of the error analysis framework used in Han's study was adopted [18]. Table 2 presents the descriptive analytical results of semantic errors in the MT-produced results.

Table 2. Descriptive Analysis of Semantic Errors produced by MT

Criteria	Frequency (n)	Percentage (%)
Accuracy	119	47.0
Fluency	37	14.6
Cohesion	25	9.9
Coherence	15	5.9
Ambiguity	37	14.6
Conciseness	3	1.2
Functionality	17	6.7
Total	253	100

As shown in Table 2 above, a total of 253 semantic errors were identified in the MT results. The highest frequency of errors was seen in *accuracy* at the rate of 47%, followed by *fluency* and *ambiguity* at the rate of 14.6% respectively. Examples of the three most frequently committed errors are as follows: *accuracy* (Korean source text: *오늘날 AI 스피커는 사람이 편리하게 사용할 수 있도록 개발되었다.* English translation: *Today, AI speakers have been developed for people to use in a convenient [way].), *fluency* (Korean source text: *기계의 도덕적 선택 시스템을 개발하기는 어렵습니다.* English translation: *It is difficult to develop a machine moral selection system.), *ambiguity* (Korean source text: *다른 언어사용자들은 당신의 언어에 혼돈을 느낄 것이다.* English translation: *Other language speakers will feel confused about your language.). The results presented in Table 2 indicate that MT tools are more likely to make errors related to the aspects of clearly translating the meaning of the original text and less likely to make errors related to repeating or translating any unnecessary vocabulary or expressions.

In order to see how students thought about the quality of the MT results, their reflective journals, which had been written while they were performing the writing tasks with MT assistance, were analyzed by

carefully reading their journals multiple times. The following excerpts are the two most predominantly mentioned issues in all the students' reflective journals. The first excerpt is from one of the female students' journals. She mentioned that the MT-produced results did not accurately translate her initial intention.

> "I thought the MT results were not grammatically wrong, but there were some parts that were different from my initial intention. For example, if a subject was omitted in the original Korean text, *you*, as a subject, were mostly used in the MT-produced English translation. However, there were many parts that needed to be corrected into *I*."

Looking at the second excerpt from one of the other female students' journals, it can be indicated that she mentioned about the fluency of the MT-generated output.

> "I wrote my initial draft by using a colloquial expression in Korean, but the MT-produced output was written in a literary style. And there were quite a lot of parts in translations that were not natural. So, I thought I needed to revise the sentences by accurately placing the subject, verb, and object, and omitting unnecessary vocabulary or expressions."

On the other hand, students were not likely to find out the errors which were related to *ambiguity*. It was revealed from the student-teacher conference session. Students stated that they mostly compared the Korean source language and the MT-produced English text at a word level, so it was difficult to grasp the parts which were likely to be interpreted in several different ways.

5. CONCLUDING REMARKS

The purpose of this study was to investigate the types of semantic errors committed by MT in translating EFL college students' original drafts written in Korean into English. Specifically, the present study attempts to find out 1) what types of semantic errors are most frequently made by MT? and 2) how students feel about the quality of the MT-produced output? The findings from this study revealed that MT produced the errors related to accuracy (47%) the most, followed by the errors related to fluency and ambiguity (14.6% respectively). As mentioned in the students' self-reflective journals, students were relatively well aware of the errors with *accuracy* and *fluency* but had limited ability to check the errors with *ambiguity*.

Based on the findings, this study suggests that: 1) it is necessary for MT users including L2 learners to understand the structural differences between the source and target languages. They need to consider and reflect on the differences when writing the source text to obtain more semantically accurate translation results. 2) L2 language teachers need to provide thorough guidelines for checking semantic errors, which learners should focus on while post-editing the MT translations in order to reduce the errors in the MT-produced results. As we are living in the era of International Technology, and MT technology is rapidly developing, it is unavoidable to use MT tools in language classrooms. Therefore, this study should lead to further research on investigating the changes of the semantic errors in the MT-produced results and students' revised ones.

REFERENCES

[1] I. Garcia, "Can Machine Translation Help the Language Learner?," in Proc. International Conference ICT for Language Learning, pp. 1-4, 2016.

[2]   J. R. Jolley and L. Maimone, "Free Online Machine Translation: Use and Perceptions by Spanish Students and Instructors," in A. Moeller (Ed.), Learn Language, Explore Cultures, Transform Lives (pp. 181-200). Minneapolis: 2015 Central States Conference on the Teaching of Foreign Languages.

[3]   S. M. Lee, "Korean College Students' Perceptions toward the Effectiveness of Machine Translation on L2 Revision," Multimedia- Assisted Language Learning, Vol. 22, No. 4, pp. 206-225, 2019.

[4]   S. M. Lee, "The Impact of Using Machine Translation on EFL Students' Writing," Computer Assisted Language Learning, Vol. 33, No. 3, pp. 157–175, 2020. DOI: 10.1080/09588221.2018.1553186.

[5]   E. O'Neill, "Measuring the Impact of Online Translation on FL Writing Scores," The IALLT Journal, Vol. 46, No. 2, pp. 1-39, 2016.

[6]   K. D. White and E. Heidrich, "Our Policies, their Text: German Language Students' Strategies with Beliefs about Web-based Machine Translation," Die Unterrichtspraxis/Teaching German, Vol. 46, No. 2, pp. 230-250, 2013.

[7]   J. Y. Baek and K. H. Rha, "Korean College Students' Self-Directed Use of Machine Translator in an English Writing Class During the Digital Era: Is This a Crisis or a Chance?," Journal of Safety and Crisis Management, Vol. 11, No. 6, pp. 27-34, 2021. DOI: 10.14251/jscm.2021.6.27.

[8]   S. Doherty and D. Kenny, "The Design and Evaluation of a Statistical Machine Translation Syllabus for Translation Students," The Interpreter and Translator Trainer, Vol. 8, No. 2, pp. 295-315, 2014. DOI: 10.1080/1750399X.2014.937571.

[9]   A. Niño, "Exploring the Use of Online Machine Translation for Independent Language Learning," Research in Learning Technology, Vol. 28, pp. 1-32, 2020. DOI: 10.25304/rlt.v28.2402.

[10] C. S. Lee, "Stylometric Comparative Analysis of Style in Human vs. Machine Literary Translations," The Journal of Translation Studies, Vol. 20, No. 2, pp. 111-130, 2019. DOI: 10.15749/jts.2019.20.2.005.

[11] M. Y. Ahn, "MT Problems and its MTPE Suggestion with Regard to Structural Differences between English and Korean," The Mirae Journal of English Language and Literature, Vol. 25, No. 1, pp. 103-130, 2020. DOI: 10.46449/MJELL.2020.02.25.1.103.

[12] Y. Hwang, Y. Kim, and K. Jung, "Context-Aware Neural Machine Translation for Korean Honorific Expressions," Electronics, Vol. 10, No. 13, 1589, 2021. DOI: 10.3390/electronics10131589.

[13] K. Alasta and L. Sujarwati, "Lexical Errors Produced by Google Translate in Translating "Putri Serindang Bulan" to English Language," *Journal of Development and Innovation in Language and Literature Education*, Vol. 2, No. 2, pp. 200-211, 2021. DOI: 10.52690/jadila.v2i2.196.

[14] M. Yamada, "A Study of the Translation Process through Translators' Interim Products," *Interpreting and Translation Studies*, No. 9, pp. 159-176, 2009.

[15] J. Vardaro, M. Schaeffer, and S. Hansen-Schirra, "Translation Quality and Error Recognition in Professional Neural Machine Translation Post-Editing," *Informatics*, Vol. 6, No. 3, pp. 1-29, 2019. DOI: 10.3390/informatics6030041.

[16] N. F. M. Sholikhah, and R. N. Indah, "Common Lexical Errors Made by Machine Translation on Cultural Text," *Jurnal Linguistiks Terapan dan Pendidikan Bahasa Inggris*, Vol. 8, No. 1, pp. 39-51, 2021.

[17] M. P. A. Llach, "A Critical Review of the Terminology and Taxonomies Used in The Literature on Lexical Errors," *A Journal of English and American Studies*, No. 31, pp. 11-24, 2005.

[18] S. H. Han, "Factor Analysis of Machine Translation Quality," *Interpreting and Translation Studies*, Vol. 25, No. 2, pp. 147-170, 2021. DOI: 10.22844/its.2021.25.2.147.