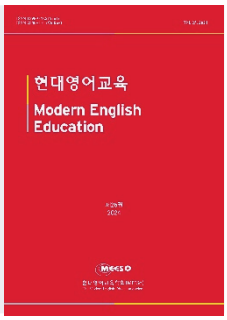
Modern English Education, Vol. 25, (2024) pp. 74-84
doi: https://doi.org/10.18095/meeso.2024.25.1.74

Modern English Education

ISSN: 1598-0782 / e-ISSN: 2586-6141
http://journal.meeso.or.kr



Analysis of Relative Clauses in English-to-Korean Machine Translation Systems' Output

Chae-Eun Kim (Pusan National University)

Received: 13 February 2024
Revised: 14 March 2024
Accepted: 27 March 2024

Kim, Chae-Eun. (2024). Analysis of relative clauses in English-to-Korean machine translation systems' output. *Modern English Education*, *25*, 74-84.

Keywords
AI, machine translation, relative clauses
인공지능, 기계번역, 관계절

Chae-Eun Kim
Associate Professor
Department of English Education
Pusan National University
alohachaeeun@pusan.ac.kr

* This work was supported by a Humanities Social-Science Research Promotion of Pusan National University (2023-24).

Abstract

The use of artificial intelligence (AI)-based tools for teaching and learning English as a foreign language has become increasingly functional in Second Language Acquisition (SLA). This study aimed to explore the potential of English-to-Korean machine translators (MTs) as interactive tools for language learning, specifically for Korean EFL learners. The study assessed the accuracy of two commonly used MTs in converting English Relative Clauses (RCs) into Korean equivalents by analyzing translation data from research articles. A total of 60 articles provided 513 examples of RCs. The study analyzed the output of MTs in three ways: (i) whether the translated meaning accurately represented the targeted form, (ii) what types of errors the Korean translations made, and (iii) how the results compared to previous works on Korean-to-English translation of RCs. The results showed that currently available English-to-Korean MTs could produce semantically correct translations at high rates (around 95%). However, they make errors such as omitting lexical elements, resulting in structural ambiguity. Compared to previous works, the accuracy rate has gradually improved. These findings suggest that MTs have great potential as aids in English reading and writing instruction.

INTRODUCTION

This study examines how two machine translators (hereafter MTs), namely Google Translator and Naver Papago, translate English relative clauses (RCs) into Korean. English RCs are complex because they require movement under transformation grammar, as noted by Aarts (2018) and Biber et al. (1999), exemplified in (1a) and (2a). For instance, in (2a), *The book* moved to an antecedent because it is relativized and it left a gap which is indicated with _. Similar processes occur in forming Korean RCs, as demonstrated in (1b) and (2b). However, a distinction arises between English and Korean RCs due to the structural differences between the two languages: English being a head-initial language and Korean being a head-final language. This variance influences the position of relativized items, as word-order sequence plays a crucial role in RC formation. Additionally, Korean exhibits the use of different case markers; for example, *-ul* (the accusative case maker) is

employed in subject RCs, such as in (1b), while *-i* (the nominative case marker) is used in direct object RCs, illustrated in (2b).

(1a) Subject RCs: the boy [that _ read the book]
(1b) Subject RCs: [_ chayk-ul ilk-un] sonyen
(2a) Direct object RCs: the book [that the boy read _]
(2b) Direct object RCs: [sonyen-i _ ilk-un] chayk

This crosslinguistic influence (Sharwood & Kellerman, 1986) results in 'avoidance' or differential rates in acquiring certain L2 syntactic structures. The mixture of different language systems has been considered as an obstacle to learning a target language. Previous research on Korean and English relative clauses (RCs) has shown that both first and second language (L1 & L2) learners encounter more difficulties comprehending and producing direct object RCs compared to subject RCs (C.E. Kim, 2015, 2018). This regularity happens both in English and Korean as L2 acquisition as well as L1 acquisition. In particular, when the structure shows straightforward contrasting traits at the syntactic level, L2 learners manifested much more difficulties in comprehending and producing the target structure (VanPatten, 2014). In this context, many L2 learners have shown systematic errors in using English RCs (C.E. Kim, 2018).

In recent years, numerous researchers have been studying how L2 learners enhance their English proficiency through the use of AI-based MTs (C.E. Kim, 2022, 2023; H. K. Kim & S. Han, 2021). Specifically, MTs have become a popular and accessible tool for L2 learners to assess the correctness of their productive items, such as writing outcomes. Additionally, L2 learners may use MTs to verify the meaning of sentences with complex structures when they read them. For this reason, the present work analyzed how well two famous MTs produce outcomes when they translate English RCs into Korean RCs. These findings aim to provide English as a foreign language educators and learners with insights into the reliability of MTs for teaching and learning purposes.

AI technologies have gained significant popularity in educational settings due to their widespread availability as free software systems and convenient tools (C.E. Kim, 2023). Google Translate (GT) and Naver Papago (NP) are both AI-based machine translators known for providing fairly reliable translation outcomes. The translation capabilities of machine translators contribute to the potential for self-directed learning (Knowles, 2021), where learners take the initiative in identifying their learning needs, goals, and assuming responsibility for their learning. As online-based learning tools, MTs provide meaningful translation outputs, making them beneficial for learners both with and without teachers, both in and out of the classroom. The utilization of AI technology in language pedagogy is inevitable; the question then arises: how effectively do MTs produce translation outputs? Depending on the answers to this question, our position on the adoption of MTs in language learning and teaching may be decisive. To answer this question accurately, educators and learners need to understand how reliable MTs are. Recent studies have regarded MTs as reliable educational tools due to higher accuracy in translation outcomes (C.E. Kim, 2023). Hence, the use of AI-based MTs appears inevitable in EFL settings as an indispensable tool. For the effective utilization of MTs both inside and outside the classroom, researchers must address concerns regarding the reliability of translation outcomes raised by policymakers and educators. Thus, well-designed research that examines grammatical patterns in translation allows us to determine the extent to which we rely on these tools. Significantly, AI-based translation has demonstrated continuous improvement compared to previous iterations, owing to the influx of vast amounts of data into the dataset. The databases of MTs in both languages are sufficiently robust to produce clear translations (C.E. Kim, 2022, 2023). This study examines the effectiveness of the MTs most commonly used by EFL learners in translating English RCs into Korean RCs.

LITERATURE REVIEW

English Relative Clauses

English relative clauses are considered as one of the complex syntactic structures because they have a movement evoked by transformation (Aarts, 2018). According to Comrie and Keenan (1979), relative clauses exhibit a universally accessible hierarchy, as depicted in (5). Subject relative clauses are easier to produce and comprehend compared to other types of relative clauses situated lower in the hierarchy. This intriguing discovery has been corroborated by previous studies in L1/L2 acquisition, sentence processing, and foreign language learning/teaching.

(3) the boy [that kicked the ball]
(4) the ball [that the boy kicked_]
(5) subject > direct object > indirect object > oblique > genitive > comparative

A vast amount of literature has been collected to support Comrie and Keenan's work, and various theoretical claims have also been proposed. In (6a-f), the subjects (the woman), direct objects (the man), indirect objects (the boy), obliques (the desk), genitives (the brother), and comparatives (Mary) that are present in the unfolded clauses are relativized.

(6a) Subject RC: the woman that _ loves the man
(6b) Direct object RC: the man that the woman loves _
(6c) Indirect object RC: a boy that the girl is giving a bag to _
(6d) Oblique RC: the desk that the boy put the book on_
(6e) Genitive RC: the brother whose _ the car is big
(6f) Comparative RC: Mary who John is taller than _

The following are three popular hypotheses in the field: (i) canonical word order, (ii) structural distance hypothesis, and (iii) linear distance hypothesis. Three explanations account for the difficulty in producing and comprehending a certain type of RC in many languages. First, canonical word order can be an effective factor for the processing of RCs. According to Slobin and Bever (1982), subject RCs are much easier to process than direct object RCs because subject RCs follow canonical word order, while direct object RCs do not. Secondly, the difficulty of RCs was explained by the structural distance hypothesis (hereafter SDH) (O'Grady, 1997). His approach explains that direct object RCs are more difficult than subject RCs because the former case has two maximal projections (V, S), while the latter one has only one node (S) between the filler and the gap. Thirdly, the linear distance hypothesis (hereafter LDH) (Gibson, 1998) explains that the distance between the filler and the gap is longer in direct object RCs than in subject RCs.

A considerable amount of SLA research has analyzed various aspects of teaching and learning English relative clauses (RCs) in L2 contexts. However, the majority of previous studies have primarily focused on the grammatical errors made by L2 learners when producing English RCs. Furthermore, numerous studies have investigated how L2 learners process specific RC structures from a psycholinguistic perspective. In order to use RCs appropriately, ESL/EFL learners must grasp the fundamental meanings and syntactic differences of RCs in both languages.

Korean Relative Clauses

The following examples illustrate Korean subject RCs, as shown in (7a), and Korean direct object RCs, as demonstrated in (7b). Previous studies have shown that Korean subject RCs are easier for both L1 speakers (C.E. Kim, 2015) and L2 speakers (C.E. Kim & O'Grady, 2016) to comprehend and produce compared to direct object RCs. Both learners of Korean and English languages have shown difficulty in processing a specific type of RC. While syntactic movement occurs in English RCs, Korean RCs are constructed with a gap. In accordance with the operator-movement analysis, prevalent in studies on Korean RCs, it is posited that the gap undergoes movement, thereby resulting in the presence of a null operator trace (C.H. Han, 2013). In Korean, RCs are constructed using a relativized link and are indicated by case markers such as *-ka, -ul/-lul, and -eykey*, among others. Due to this cross-linguistic difference, both Korean L2 learners of English and English-speaking L2 learners of Korean face challenges in comprehending and producing RCs (C.E. Kim, 2015; C.E. Kim & O'Grady, 2016). Therefore, the process of producing RCs in Korean poses a challenge for non-native speakers of the language.

(7a) [_ kong-ul chan] sonyen
ball-Acc kick-RC boy
'the boy that kicked the ball'
(7b) [sonyen-i _ cha-n] kong
boy-Nom kick-RC ball
'the ball that the boy kicked'

According to C. E. Kim (2021) and H. Sohn (2001), the Korean RCs are composed of four types of RCs as shown in (8). As in (8), subject, direct object, indirect object, and oblique RCs indicated syntactic differences compared to English counter-examples discussed in (6).

(8a) Subject RC: [namca-lul salangha-nun] yeca
man-ACC love-RC woman
'the woman that loves the man'
(8b) Direct object RC: [namca-ka salangha-nun] yeca
man-NOM love-RC woman
'the woman that the man loves'
(8c) Indirect object RC: [yeca-ka kabang-ul cwuko-iss-nun] namca
woman-NOM bag-ACC give-Prog-RC man
'boy that the girl is giving a bag to'
(8d) Oblique RC: [namca-ka chayk-ul ollyenoh-un] chayksang
man-NOM book-ACC put-RC desk
'the desk that the boy put the book on_'

Similar to English RCs depicted in (6), the subjects (the woman), direct objects (the man), indirect objects (the boy), and obliques (the desk) present in the unfolded clauses are relativized. However, they are positioned on the right-hand side due to Korean's head-final language structure. Moreover, Korean utilizes a case-marking system to express grammatical functions in RCs, whereas in English, the case is determined by the word order system. The confluence of these various factors causes processing difficulties for L2 learners across two linguistic systems. Therefore, the discernment and adept application of these distinctions are imperative prerequisites for the accurate execution of MTs properly.

As a trial to evaluate this particular issue, C.E. Kim investigated the translation of Korean RCs into English RCs by the two MTs, achieving an overall accuracy rate of 64.50% for GT and 62.96% for NP. Her study revealed the accuracy of translated outcomes across subject, oblique, direct object, and indirect object RCs, albeit with a slightly different order of accuracy compared to the present study. Notably, she identified crucial errors in the translation of Korean RCs into English, primarily stemming from difficulties in producing proper RC structures. According to her study, various case markers in Korean RCs are the huge influences on forming English RC structures. To address the challenges posed by MTs, obtaining a thorough understanding of Korean case markers is crucial.

Problems of MTs

Numerous researchers have focused on analyzing the outcomes of MTs. C. E. Kim (2021, 2022, 2023) discovered that the accuracy of two MTs has improved every year. This was supported by research analyzing various Korean-to-English and English-to-Korean translation processes of RCs and passive constructions. These findings demonstrate how the translation process transfers content from the original language into the target language.

Issues with MTs have been discussed since their inception. Early studies on MTs aimed to highlight the discrepancies in translation outcomes. Consequently, most results indicated that MTs lacked sensitivity to the contextual nuances present during translation, regardless of whether the translator was human or machine. Linguistic differences, such as homophones and syntactic distinctions, are major contributing factors to the challenges encountered in machine translation. The main challenges arose because early MTs in the 1950s generated translation data through a statistical machine translation (SMT), which selected possibly connected matched items. SMT approach involves analyzing a text corpus to establish connections between source data and expected results. Parallel corpora have been stored to learn patterns and relationships between words and phrases in two languages. In recent years, Neural machine translation (NMT) has gained a reliable reputation due to its ability to produce more accurate translations (Stahlberg, 2020). Because the accuracy that MTs produce increases, research on utilizing MTs for foreign language teaching has been a significant concern. However, there has been limited research to date focused on analyzing the outcomes of MTs from syntactic and semantic perspectives This study contributes to foreign language teaching and learning by utilizing AI-based MTs to translate English RCs into Korean. The primary aim of the study is to assess the accuracy of translating English RCs into Korean sentences using major MTs and to evaluate their suitability for Korean EFL contexts. The main reason for focusing on RCs in this study is that using them appropriately necessitates an understanding of their morpho-syntactic functions. This examination helps determine the suitability of using MTs as a tool for enhancing language competence among L2 learners.

This study examines the accuracy of English-to-Korean translations of English RCs generated by two MTs. The study addresses the following research questions:

(i) Do the two MTs (GT and NP) produce accurate translations of English RCs into Korean?
(ii) What types of errors do the two MTs make when translating English RCs into Korean?
(iii) How do the results compare to previous research (C.E. Kim, 2021) on the accuracy of Korean-to-English MTs' translations of RCs?

THE STUDY

The following section outlines the data collection, translation process, and subsequent analysis.

Data Collection

A total of 513 English RCs were extracted from academic research papers on English RCs available in the Korean Citation Index, which contains scholarly papers in Korea. To obtain meaningful examples of RCs, academic papers focusing on Korean and English RCs were collected for data analysis. Selecting papers from the KCI for data collection helps maintain the reliability of chosen sentence examples, as they have been certified by experts in the field of linguistics and language education. A total of 60 journal articles found in the KCI provided 513 examples of English RCs. All sentences were entered into an Excel spreadsheet. Two researchers, including the author, used GT and NP to translate the English RCs into Korean. To ensure the accuracy of the results, the researchers categorized the errors accordingly.

Data Analysis

All the English RCs translated by the two MTs were analyzed to determine: (i) whether the outcomes' grammatical and semantic meaning is correct as intended; (ii) whether the outcomes' grammatical and semantic meaning is correct but not as intended; and (iii) what types of errors are produced. The subsequent illustration (9a) demonstrates instances where MTs yielded results that are grammatically and semantically precise. Moreover; (9b) adheres to grammatical correctness though generating a passive voice construction instead of the intended active voice. Conversely, (9c) illustrates an erroneous outcome characterized by grammatical inaccuracies.

(9a) GT:
Subject RCs: Every boy that loves Mary

meli-lul	salangha-nun	motun	sonyen-tul
Mary-Acc	love-RCs	every	boy-PL

(9b) Non-targeted accurate form: Passive voice
The monkey that the tiger hit

holangi-eykey	mac-un	wenswungi
tiger-DAT	hit-PASSIVE	money

(9c) GT: ungrammatical outcome
Direct object RC: the bear that the dog pushed

*kay-ka	mil-lin	kom
dog-NOM	push-PASSIVE	bear

The analysis is conducted because two MTs sometimes produce outputs that are semantically accurate but not in the intended form, similar to a human language processor. When reviewing the outcomes from the two MTs, it is necessary to check whether they are semantically and grammatically accurate despite not being in the intended form. For each MT, the English RCs along with their translations were entered into an Excel sheet for coding and analysis. To ensure reliability, the translations were initially analyzed by a native Korean-speaking researcher and subsequently reviewed by the author and another native Korean speaker. This data-checking process contributes to ensuring the reliability of the data (Krippendorff, 2011). The translations were analyzed twice: initially for overall meaning and grammaticality, and then to categorize all errors according to four types of Korean RCs.

RESULTS

GT and NP were utilized to translate the 513 English RCs into Korean, and the results are summarized in Table 1. Out of the 513 English RCs collected, subject RCs accounted for the highest proportion at 42.30%, followed by direct object RCs at 37.42%. Oblique RCs comprised 13.84% of the total, while indirect object RCs accounted for 6.43%.

TABLE 1
Percentages of RCs

Types of RCs	N	%
Subject	217	42.30%
Direct object	192	37.42%
Indirect object	33	6.43%
Oblique	71	13.84%
Total	513	100%

The translations of all 513 RCs by the MTs were analyzed to address the first research question. The initial analysis evaluated all translations for accuracy in meaning, revealing that NP slightly outperformed GT in terms of accuracy. The outcomes can be categorized into three groups: (i) targeted forms as semantically and grammatically accurate RCs, (ii) non-targeted forms as semantically and grammatically accurate RCs, and (iii) semantically and grammatically inaccurate. Table 2 illustrates the proportions of semantically and grammatically accurate RCs as targeted forms for each MT: 94.54% for GT and 95.32% for NP. Additionally, both MTs generated a small number of semantically and grammatically accurate outcomes as non-targeted forms: 0.39% for GT and 0.58% for NP. Furthermore, both MTs produced semantically and grammatically inaccurate forms: 9.55% for GT and 8.77% for NP.

TABLE 2
Korean Translations of English RCs: Rate of Semantically and Grammatically Accurate Translations Produced by the Two MTs

Category	GT	NP
Targeted form: Semantically and grammatically accurate	94.54% (485/513)	95.32% (489/513)
Non-targeted form: Semantically and grammatically accurate	0.39% (2/513)	0.58% (3/513)
Semantically and grammatically inaccurate	9.55% (49/513)	8.77% (45/513)
Total	100% (513/513)	100% (513/513)

For the second analysis, the data were examined based on the types of RCs. They were categorized into four types of English RCs: subject, direct object, indirect object, and oblique RCs.

TABLE 3
Korean Translations of English Subject RCs: Rate of Semantically and Grammatically Accurate Translations Produced by the Two MTs

Response Types	GT	NP
Targeted form: Semantically and grammatically accurate	94.91% (206/217)	96.31% (209/217)
Non-targeted form: Semantically and grammatically accurate	0.92% (2/217)	0.46% (1/217)
Semantically and grammatically inaccurate	4.14% (9/217)	3.22% (7/217)
Total	100% (217/217)	100% (217/217)

As shown in Table 3 above, more than 90% of all RCs translated by the two MTs were accurate (94.91% for GT and 96.31% for NP). The following examples in (10) demonstrate subject RCs produced as targeted forms by the MTs. Both properly construct the subject RCs with appropriate lexical elements and the relativizer '-nun', corresponding to the English 'that'.

(10a) GT:
Subject RCs: Every boy that loves Mary

meli-lul	salangha-nun	motun	sonyen-tul
Mary-Acc	love-RCs	every	boy-PL

(10b) NP:
Subject RCs: Every boy that loves Mary

meli-lul	salangha-nun	motun	sonyen
Mary-Acc	love-RCs	every	boy

Examples in (11) demonstrate semantically and grammatically inaccurate forms that appear in the two MTs at similar rates. In (11b), all three lexical elements, 'people', 'them', and 'money', are shown in NP; however, an indirect object applied as 'them' as in (11a) was missing, leading to ambiguity.

(11a) GT:
Subject RCs: the people that give them the money on the high street

?kilkeli-eyse	ton-cwu-nun	salam-tul
high street-LOC	money-give-RC	person-PL

(11b) NP:
Subject RCs: the people that give them the money on the high street

kutul-eykey	kilkeli-eyse	ton-cwu-nun	salam-tul
they-DAT	High street-LOC	money-give-RC	person-PL

TABLE 4
Korean Translations of English Direct Object RCs: Rate of Semantically and Grammatically Accurate Translations Produced by the Two MTs

Response Types	GT	NP
Targeted form: Semantically and grammatically accurate	90.10% (173/192)	93.75% (180/192)
Non-targeted form: Semantically and grammatically accurate	0% (0/192)	0% (0/192)
Semantically and grammatically inaccurate	9.89% (19/192)	6.25% (12/192)
Total	100% (192/192)	100% (192/192)

As demonstrated in Table 4 above, over 90% of the direct object RCs generated by the two MTs were semantically and grammatically correct (90.10% for GT and 93.75% for NP). However, when the direct object RC is targeted, inaccurate forms were produced at rates of 9.89% for GT and 6.25% for NP. In (12b), all three lexical elements, 'yesterday', 'you', 'lose', and 'key', are present in NP; however, the subject applied as 'ney' as in (12a) was missing, resulting in ambiguity.

(12a) GT:
Direct object RCs: the key you lost yesterday

ecey	ilhepelin	yelsoy
yesterday	lose-PST-RC	key

(12b) NP:
Direct object RCs: the key you lost yesterday

ney-ka	ecey	ilhepelin	yelsoy
you-NOM	yesterday	lose-PST-RC	key

Example (13b) demonstrates grammatical and semantical correctness in NP; however, the passive form realized as '-lin' in (13a) appeared, resulting in an ungrammatical outcome. Interestingly, even non-targeted forms can be semantically and grammatically correct as in (13c). In order to produce direct object RCs, a dative case marker and 'mac-un'(get hit) were formed, rather than using a nominative case marker. When direct object RCs are targeted and translated, passivized subject RCs are produced in the MTs. This phenomenon is commonly observed in human language processing (for further

discussion, see C.E. Kim, 2023).

(13a) GT:
Direct object RC: the bear that the dog pushed
*kay-ka mil-lin kom
dog-NOM push-PASSIVE bear

(13b) NP:
Direct object RC: the bear that the dog pushed
kay-ka mil-ko-kan kom
dog-NOM push-AND-go bear

(13c) NP:
Direct object RC: The monkey that the tiger hit
holangi-eykey mac-un wenswungi
tiger-DAT hit-PASSIVE money

TABLE 5
Korean Translations of English Indirect Object RCs: Rate of Semantically and Grammatically Accurate Translations Produced by the Two MTs

Response Types	GT	NP
Targeted form: Semantically and grammatically accurate	84.84% (28/33)	81.81% (27/33)
Non-targeted form: Semantically and grammatically accurate	0% (0/33)	0% (0/33)
Semantically and grammatically inaccurate	15.15% (5/33)	18.18% (6/33)
Total	100% (33/33)	100% (33/33)

As shown in Table 5, indirect object RCs appear to be quite accurate, with similar rates of 84.84% for GT and 81.81% for NP. When the indirect object RCs are targeted, (14b) demonstrates grammatical and semantic correctness in NP. However, sometimes a reversal error occurs, resulting in subject RCs being produced instead of indirect object RCs. In (14a), the marker '*-eykey*' needs to be changed to the nominative case marker '-i' to form an indirect object RC.

(14a) GT:
Indirect object RC: the boy to whom Askale send a letter
*Askale-eykey phyenci-lul ponayn sonyen
Askale-DAT letter-ACC send-RC boy

(14b) NP:
Indirect object RC: the boy to whom Askale send a letter
Askale-i pheynci-lul ponayn sonyen
Askale-NOM letter-ACC send-RC boy

TABLE 6
Korean Translations of English Oblique RCs: Rate of Semantically and Grammatically Accurate Translations Produced by the Two MTs

Response Types	GT	NP
Targeted form: Semantically and grammatically accurate	84.50% (60/71)	83.09% (59/71)
Semantically and grammatically inaccurate	15.49% (11/71)	16.90% (12/71)
Total	100% (71/71)	100% (71/71)

As shown in Table 6, oblique RCs demonstrate fairly high accuracy rates, with similar rates of 84.50% for GT and 83.09% for NP. However, when oblique RCs are targeted, semantically and grammatically inaccurate forms were produced in both

MTs at similar rates of 15.49% for GT and 16.90% for NP. In (15a), subject deletion errors are observed, with one missing lexical element, 'subject you'. Additionally, (15b) indicates a semantic error, resulting in a nonsensical outcome due to the absence of a corresponding word for 'above.'

(15a) GT:
Oblique RC: A good book which you can sit on
*ancul-swu-iss-nun cohun chayk
sit on can-RC good book

(15b) NP:
Oblique RC: the bird that the bee is flying above
*pel-I nalatani-nun say
bee-NOM fly-RC bird

Overall, both GT and NP performed well in translating English RCs into Korean RCs. One possible assumption is that NP and GT, being sensitive to Korean RCs, now produce more advanced outputs than before (see C.E. Kim, 2021). This suggests that it may be possible to resolve issues regarding the accuracy of the data and explore the potential application of MTs in language teaching settings.

DISCUSSION AND CONCLUSION

This study aims to investigate the translation of English RCs into Korean RCs using two MTs, Google Translator and Naver Papago. A total of 513 English RCs were translated using these MTs. The results were analyzed to determine the rates of semantic accuracy and the frequency of successful translations into Korean RCs. Each of the 513 translations produced by GT and NP falls into one of three categories: (i) targeted form; (ii) non-targeted form (semantically and grammatically correct); and (iii) semantically and grammatically incorrect form. Based on the analysis of these outcomes, three main findings emerge, which are discussed below.

The first major finding of this study is that both MTs, GT and NP, translated English RCs into semantically correct Korean RCs at high rates: GT at 94.54% and NP at 95.32%. Interestingly, even when they did not produce targeted RCs, the MTs mostly provided sentences with appropriate meanings, resulting in semantically accurate Korean RC structures at rates of 0.39% for GT and 0.58% for NP. However, both GT and NP also generated semantically and grammatically inaccurate forms, with rates of 9.55% and 8.77%, respectively. Overall, this analysis suggests that both MTs perform well in translating English RCs, with sentences expressing the targeted form at rates above 90%. Although these occurrences were minimal, GT (0.39%) and NP (0.58%) also produced non-targeted but semantically and grammatically accurate forms. As highlighted by various researchers, the ability to select one structure among possible language patterns to convey the same meaning is a fundamental characteristic of human language (MacDonald, 2013). Similarly, the two MTs exhibit diverse outcomes, occasionally generating both active and passive voice sentences to convey the same meaning. Given this diversity, the analysis in this study primarily focuses on the semantic appropriateness of the MTs' output.

Secondly, both MTs exhibited instances of inaccurate responses, including missing lexical items and generating ungrammatical outcomes. Dropping lexical elements, particularly pronouns, was a common error observed in the translations, resulting in ambiguity in the sentence's meaning. An interesting observation is that both MTs tended to drop lexical items primarily when they were pronouns. Additionally, in cases where direct object RCs with active voice were targeted, the MTs often produced passive voice constructions to avoid sentences with transitivity. Despite this, the translations provided semantically correct responses for the targeted RCs.

Thirdly, in comparison to the translation from Korean RCs to English RCs, the findings of the current study align closely with those of previous research (C.E. Kim, 2021). Over the span of nearly two years, the accuracy of RC translation has significantly improved, as evidenced by the findings of the current study. GT and NP exhibited low error rates in this study, indicating meaningful and grammatically correct outcomes and instilling confidence in the use of MTs for language teaching and learning. Based on the principles of NMT, the reliable data generated from two studies translating RCs in different languages can serve as valuable guidance for the teaching and learning of English through MTs.

The analysis of all RCs proceeded as follows to yield meaningful results. Firstly, each RC was input into the two MTs to assess their accuracy, providing insight into the performance of the MTs. Additionally, errors were categorized to identify the underlying causes of linguistic difficulties encountered by the MTs, aiming to propose potential improvements. In

summary, English RCs were predominantly translated into semantically appropriate equivalents by the most popular MTs. Furthermore, the MTs demonstrated successful translations of passive voice constructions when targeting direct object RCs. These findings align closely with human translation outcomes. Human language processing often involves selecting a specific structure from various possible alternatives, occasionally employing simplification strategies to opt for simpler structures over more complex ones.

The MTs exhibit similar behavior in their performance. Overall, the results indicate that MTs hold significant potential as a language learning tool, usable by both instructors for teaching reading and writing, and by learners as an independent learning aid. The findings of this research suggest that educators and learners can confidently utilize MTs, especially in written discourse contexts. However, it is crucial to assess the applicability of MT outputs in foreign language learning and teaching. While this study demonstrated high accuracy in translating English RCs, it is essential to note that the MTs occasionally generate inappropriate patterns. These findings carry pedagogical implications, emphasizing the importance of MT reliability awareness among teachers and learners, as well as the need to explore alternative structures provided by MTs. For EFL learners in middle and high school, it is imperative for EFL teachers to comprehend English RC structures and their translation into Korean RCs. More importantly, mastering the structure of RCs in English is the crucial initial step in composing a variety of RCs, and this process can be facilitated through the use of MTs. English textbooks used in Korean middle and high schools typically introduce only two types of RCs: subject and direct object RCs, with a focus on the appropriate relative pronoun for each type. However, L2 learners need to acquire a comprehensive understanding of the diverse forms and functions of RCs in both English and Korean (Towle & Cottrell, 1992). Exposure to the structural differences between their L1 and L2 is essential for EFL learners' grammatical development (Celce-Murcia & Larsen-Freeman, 1999; Cowan, 2008; Huddleston, 1984; H.K. Kim & S. Han, 2021; Quirk et al., 1972).

This study has identified several limitations. Firstly, in order to enhance the reliability of translated data from English RCs to Korean, it is imperative to meticulously address errors generated by MTs, especially for educational applications. Secondly, further research is warranted to devise effective pedagogical approaches that leverage the capabilities of MTs to teach English grammar, including RCs, to EFL learners within classroom environments.

References

Aarts, B. (2018). *English syntax and argumentation*. Bloomsbury Academic.

Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E., & Quirk, R. (1999). *Longman grammar of spoken and written English* (Vol. 2). Longman.

Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The grammar book: An ESL/EFL teacher in course* (2nd ed.). Heinle/Cengage Learning.

Cowan, R. (2008). *The teacher's grammar of English with answers: A course book and reference guide*. Cambridge University Press.

Comrie, B., & Keenan, E. L. (1979). Noun phrase accessibility revisited. *Language*, *55*(3), 649-664.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*(1), 1-76.

Han, Chung-Hye. (2013). On the syntax of relative clauses in Korean. *Canadian Journal of Linguistics*, *58*(2), 319-347.

Huddleston, R. (1984). *Introduction to the grammar of English*. Cambridge University Press.

Kim, Chae-Eun. (2015). Korean children's acquisition of relative clauses. *Language Research, 51*(2), *421-441.*

Kim, Chae-Eun, & O'Grady, W. (2016). Asymmetries in children's production of relative clauses: Data from English and Korean. *Journal of Child Language, 43*(5), 1038-1071.

Kim, Chae-Eun. (2018). Korean EFL learners' use of complex RCs in a CMC context. *Modern English Education*, *19*(1), 29-37.

Kim, Chae -Eun. (2021). A study on relative clauses in Korean-to-English machine translation systems. *English 21, 34*(4), 279-305.

Kim, Chae -Eun. (2022). Analysis of Korean-to-English machine translation systems' treatment of passives. *Journal of Pan-Pacific Association of Applied Linguistics, 26*(1), 87-103.

Kim, Chae -Eun. (2023). Analysis of passives in English-to-Korean machine translation systems' output. *Modern English Education, 24*(1), 1-9.

Kim, Hye-Kyung, & Han, Sumi (2021). College students' perceptions of AI-based writing learning tools: With a focus on Google Translate, Naver Papago, and Grammarly. *Modern English Education, 22*(4), 90-100.

Knowles, M. A. (2021). Five motivating concerns for AI ethics instruction. *Proceedings of the Association for Information Science and Technology*, *58*(1), 472-476.

Krippendorff, K. (2011). Agreement and information in the reliability of coding. *Communication Methods and Measures*, *5*(2), 93-112.

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology, 4*(1), 1-16.

O'Grady, W. (1997). *Syntactic development*. University of Chicago Press.

Quirk, R., Greenbaum, S., Leech, G. N., & Svartvik, J. (1972). *A grammar of contemporary English*. Longman.
Slobin, D. I., & Bever, T. G. (1982). Children use canonical sentence schemas: A crosslinguistic study of word order and inflections. *Cognition*, *12*(3), 229-265.
Sharwood, M. S., & Kellerman, E. (1986). Crosslinguistic influence in second language acquisition: An introduction. In E. Kellerman, & M. S. Sharwood (Eds.), *Crosslinguistic influence in second language acquisition* (pp. 1-9). Pergamon Press.
Sohn, Ho-Min. (2001). *The Korean language*. Cambridge University Press.
Stahlberg, F. (2020). Neural machine translation: A review. *Journal of Artificial Intelligence Research*, *69*(1), 343-418.
Towle, A., & Cottrell, D. (1996). Self-directed learning. *Archives of disease in childhood*, *74*(4), 357-359.
VanPatten, B. (2014). Input processing in adult SLA. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (pp. 125-146). Routledge.