An Analysis of Josa and Eomi in Translating Korean TV Dramas Into English With Artificial Intelligence

Koh, Sungran

Abstract

The goal of this study is to find out why the English subtitles of Korean TV dramas have frequent errors. It is anticipated that the findings would shed light on innovative ways for machine translation technology to agglutinate languages. To do this, as a first step, Korean-English subtitles were grammatically tagged according to the category part of speech (POS) to find out which POS has the most frequent errors in each language. Thirty-one groups were analyzed and categorized by tagging the part of speech. Then, for the Korean language, the Kokoma Korean morpheme analyzer was run to tag the Korean script according to the category noun, verb, adjective, etc. These were categorized into forty-five groups. This categorization included nine subsets of josa (postposition) and fourteen of eomi (ending), which are the most difficult parts of Korean to translate into English due to differences in linguistic structure. As a next step, the subtitles were scored and graded as the most corrected and the least corrected by Korean-American bilinguals. The results show that the most frequent error of josa is JX (auxiliary particle) among nine groups whereas the frequent error of eomi is EPT (tense prefinal ending).

Keywords: translation, function words, eomi and josa, agglutinating languages, POS tagging

Applicable levels: secondary, tertiary

1 Corresponding author, Lecturer, College of General Education, Kookmin University, 77 Jeongeung-ro, Seongbuk-gu, Seoul, 02707, Korea (E-mail: sungran@kookmin.ac.kr)

Received: April 14, 2022
Revised: May 19, 2022
Accepted: May 24, 2022

Copyright: © 2022 The Society for Teaching English through Media (STEM)

This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
I. INTRODUCTION

Korean TV drama has greatly spread all over the world during the past dozen years. Along with this spreading, the subtitles of the dramas play a crucial role for the viewer to understand them appropriately. Subtitling in Audiovisual Translation (AVT) has now received significant attention for its complicated theories and practices (Basari & Nugroho, 2017). In addition, the technical aspect of subtitling is considered important. In particular, machine translation technology has recently developed with a rapid speed in the initial stage of the Fourth Industrial Revolution. It is expressed by emerging technological breakthroughs in a lot of areas such as artificial intelligence and the Internet of Things, etc. Consequently, the statistical approach to the similarity between machine translation and human translation has been widely investigated. Ali (2016) mentions, “machine translation may play a pivotal role in helping language experts in their daily work in general and in aiding non-professionals to understand and create text in target languages in particular” (p. 55). Also, the errors of machine translation (MT) are still argued even though it has developed over a number of years. Koh (2020) states, “MT is not only related to mathematics, computer science, and linguistics but also various syntactic and semantic levels of languages have caused its error problems. MT still lacks in recognizing the proper synonym, collocation or word meaning” (p. 158). Especially, the area of spoken language is still challenging in the developing phase.

The most prominent reason for the MT errors is the difference between the first language and the second language. English and Korean are very different in terms of cultural and linguistic characteristics. Korean L2 learners have been affected by their mother tongue regarding the thought processes and their production of English. Korean and English have two different surface structures. Korean has SOV word order whereas English has SVO word order, and this causes errors in the interlingual basic order. These errors are due to learners incorrectly applying L2 basic sentential order to that of L1.

Besides being an SOV and postpositional language, the Korean language is one of the agglutinative languages with rich morphological features such as Japanese, Finnish, Hungarian, Mongolian, and Turkish. Turkish is an agglutinating language that builds words by gluing morphemes together. Dincer, Karaoglan, and Kisla (2008) remark, “the number of minor part of speech categories for Turkish are much richer than any analytical language like English” (p. 683). Agglutinating languages that are more likely to be inflected have the sparsity problem, which leads to the difficulties in POS tagging of natural language processing (Can, Ustün, & Kurfalı, 2016). For example, Turkish has six cases such as nominative, accusative, genitive, dative, ablative and locative. Studies on the L1 acquisition of Turkish show that all types of case markings have the high error rate except for nominative and ablative case markings (Gonulal, Spinner, & Winke, 2016). The case system is the hardest aspect of Turkish morphology for English-speaking learners.

While SOV languages such as English have the semantic content of the particles added to the nouns and verb stems, the agglutinative languages have the chains of particles attached to complicated suffixes, morphemes per word structures showing many syntactic characteristics due to word-form sparsity, and variable word order. Korean words are functional words divided into two categories: josa and eomi which are well defined and frequently used (Hong, Koo, & Yang, 1996). Kim, Chae, Snyder, and Kim (2014), mention, “Josa is used to define nominal cases and modify other phrases, while eomi is an ending of a verb or an adjective to define a tense, show an attitude, and connect or terminate a sentence” (p. 637). Consequently, English and Korean have different morphological and syntactic structures. The examples are as follows (Figure 1):

**FIGURE 1**
Morphological and Syntactic Structures of English and Korean (From Kim et al., 2014, p. 637)
Kim et al. (2014) state that agglutinative languages are computationally difficult due to word-form sparsity and variable word order. Consequently, it is also difficult to translate Korean-English exactly due to the huge differences in the linguistic structure between them.

In this paper, Korean-English subtitles are grammatically tagged according to the part of speech (POS) to find out which POS has the most frequent errors in each language. POS tagging is the marking of the morphemes based on the context meaning leading to a list of tagged morphemes. To do this, first of all, the English subtitles are parsed by the part-of-speech tagger (POS tagger) tool, which is software that reads English text and tags parts of speech to each word, such as noun, verb, adjective, etc. It was developed by Stanford University and is called the Stanford Log-Linear Part-Of-Speech Tagger. It is utilized to analyze and categorize into thirty-three groups by tagging the part of speech. Then, the Kokoma Korean morpheme analyzer is run to tag the Korean script according to noun, verb, adjective, etc. It is categorized into forty-five groups by tagging the parts of speech. As a second step, the subtitles are scored by the bilinguals who are Korean-Americans. The scores are graded as the most corrected (5 points) and the least corrected (1 point). The point 5 to point 3 are categorized into correct translation and the point 2 to point 1 are categorized into wrong translation. Even though the subtitles are grammatically wrong or awkward, if they are understandable and usable in an informal situation, the subtitles are grouped into over 3 points. As a final step, the least corrected of the two groups which have point 1 and 2 are analyzed as to which POS has the most common errors.

II. LITERATURE REVIEW

1. Korean Morphological Analysis

Korean is an agglutinative language which is like “beads on a string” with high productive morphology and has a rich set of derivational and inflectional suffixes. Huh and Laporte (2005) note, “derivational suffixes are markers of verbalization, adjectivization, and adverbialization. They are appended by applying transducers. Inflectional suffixes comprise all other types of suffixes. A single stem can be combined with up to 5,500 different sequences of inflectional suffixes” (p. 4). All part of speeches and words are structured by suffix from a noun or a verb root and words are formed by cumulating morphemes in a regulated order. As an example, 가시+것+습니+가 (go + honorific + future + highest + interrogative) consist of a verb stem and four morphemes. In addition, Korean has case marking affixes that mark three cases such as nominative, accusative, and so on. These affixes are attached to nouns, marking the case of the noun. Verbs inflect tense, mood, honorific (highest, high, middle, and low). According to Davis (2007), “this rich morphology and its agglutinative nature poses a formation problem, as in order to derive the full meaning of a highly inflected form, it is necessary to segment each morpheme and derive the full meaning from these morphemes” (p. 4).

Morphological analysis is the description of the internal structure of words and parts of words, such as root words, stems, prefixes, affixes, and suffixes. Park (2011) shows that features of Korean morphological analysis is as follows:

| Ambiguity of part-of-speech | Ambiguity of segmentation of morpheme |
|-------------------------------|---------------------------------------|

He mentions that ‘가시는’ in Korean has several meanings.

- 가시noun + 는/josa’ (thorn, prickle) e.g. 그 선인장의 가시는 참 따가웠다.
- 가시/verb + 는/eomi (leave, disappear, relieve) e.g. 물을 마셨더니 갑자기 가시는 가시는 가시는 가.
- 가/verb + 시/eomi + 는/eomi (go) e.g. 할머니께서는 집에 가시는 길이었다.
- 갈/verb + 시/eomi + 는/eomi (grind, sharpen) e.g. 아저씨의 갈을 가시는 모습은 인상적이다.

© 2022 The Society for Teaching English through Media (STEM)
As mentioned above, Koreans have difficulty in distinguishing them to foreigners because the same ‘가시는’ word varies depending on the context. Even though it has the exact same spelling, the word could mean different depending on the sentence and situation. For example, ‘가시/noun + 는/josa’ is used as a noun form where in other sentences it’s a verb or adjective. In addition, words like ‘가시/verb + 는/eomi,’ ‘가/verb + 시/eomi + 는/eomi’ and ‘갈/verb + 시/eomi + 는/eomi’ could mean different verb as well. Moreover, since Korean language does not have a clear word boundary, it leads to difficulty in translating into/ from Korean. As with many other languages, the ambiguity in Korean causes major problems in POS tagging and translation into English.

2. POS Tagging

Words in a written text are composed of basic information about part of speech (POS). POS tagging is the marking of the morphemes based on the context meaning leading to a list of tagged morphemes. Dincer, Karaoglan, and Kisla (2008) state, “…the terms ‘tag’ and ‘tagging’ which are in fact interchangeable with ‘code’ and ‘codes’ …, and it is the first rule-based tagging program based on a large set of hand-constructed rules and a small lexicon to handle the exceptions” (p. 680). Taggers have been developed in many different languages.

In Korean, the unit separated by white space is called Korean word eojeeol (so-called word-phrase), which is usually made up of the content word and one or several function word such as josa (postposition) or eomi (ending). For example, the word 켰다:kôssta ‘was big’ has a stem, ㅋ:k- ‘big’, and two functional morphemes, -었-ôss- (past) and -다-ta (declarative). In contrast to functional morphemes, stems are all lexical morphemes. The surface form of a morpheme shown in a sentence may depend on adjacent morphemes and is different from its base form or lexical form. For instance, the surface form of the stem ‘big’ is ㅋ:k- before the suffix -었-ôss- (past), but its base form is 크:keu-. Changing the endings could change the verb tenses, and this often requires to change the final consonant.

That is, the variation of the first letter to express verb tenses such as 켰 or 클 is the surface form. The variations are called phonotactic, and the surface variants are called allomorphs. The morphological analysis ascertains morphemes and assigns pertinent information to them. Nguyen, Vo, Shin, Tran and Ock (2019) remark, “Korean morphological analysis is to decompose a word-phrase into morphemes with three processes as follows”:

- Separating a word-phrase into morphemes;
- Recovering the original form for changed phonemes;
- Tagging the POS to each morpheme (Nguyen et al., 2019, p. 414)

This research uses Kokoma Korean morphological analyzer to do analytic tagging. This is the example using Kokoma:

```python
>>> from konlp.tag import Kkma
>>> from konlp.utils import pprint
>>> kkma = Kkma()
>>> pprint(kkma.pos ('한국어 분석은 재밌습니다!'))
[(한국어, NNG), (분석, NNG), (은, JX), (재밌, VA), (습니라, EFN), (!, SF)]
```

3. Differences Between Languages

Shaffer (2015) remarks, “Many SVO languages are prepositional, while many SOV languages are postpositional … Korean is both SOV and postpositional, while English is both SVO and prepositional” (p. 221). Accordingly, the errors are shown as the dropping of the English preposition since there is no equivalent structural constituent preceding the noun in Korean:

- a. Korean (SOV): Gene-i hakgyo-e gass-da.
- b. English (SVO): Gene went to school.
- c. English Error: Gene went school. (Adapted from Shaffer, 2002, p. 221)
Errors of the Korean postpositional particle -eui (a) involve the Korean L2 learner associating it with English (c), which keeps the noun order of the Korean (a), and produces errors in English. *Jimseung* refers to *the beasts* and *wang* to *kings*.

1. Korean (SOV): jimseung-eui wang
2. English (SVO): king of beasts
3. English Error: beast of king

Therefore, this study addresses the following research questions:

1) Among the content or function words in English, which has more errors?
2) Which has the most frequent errors among seven division (N, V, M, I, J, E, X) in Korean?
3) What are the common errors among *josa* or *eomi* in Korean?

Accordingly, it is hypothesized that the translation of Korean to/from English will cause errors in the postpositional particle. It is supported the hypothesis that the translation of Korean to/from English causes errors in the post-positional particle.

**III. METHOD**

Korean to English subtitles are grammatically tagged according to the POS in order to determine which POS has the most frequent errors in each language. This study is guided by this fact. It involved —

1) parsing English subtitles by the POS tagger, which reads English text and tags parts-of-speech to each word, such as nouns, verbs, and adjectives. The POS tagger is a natural language parser program that analyzes the grammatical structure of sentences and is called the Stanford Log-Linear Part-Of-Speech Tagger. The data were analyzed and categorized into 31 groups by tagging the parts-of-speech: DT, RP, CD, NN, NNS, NNPS, EX, PRP, PRPS, POS, RBS, RBR, RB, JJS, JJR, JJ, MD, VB, VBP, VBZ, VBD, VBG, WDT, WP, WPS, WRB, TO, IN, and CC. This categorization is included in the Appendix. The Korean subtitles were first processed in Openkoreatext, an open-source Korean text processor that handles Korean normalization and tokenization. It is categorized into 14 groups by tagging the parts-of-speech: nouns, verbs, and adjectives. They were categorized into 45 groups by tagging the parts-of-speech: NNG, NNP, NNB, NNM, NR, NP, VV, VA, VXX, VXA, VCP, VCN, MDN, MDT, MAG, MAC, IC, JKS, JC, JKG, JKO, JKM, JKI, JKQ, JC, JX, EPH, EPT, EPP, EFN, EFQ, EFO, EFA, EFI, EFR, ECE, ECS, ECD, ETN, ETD, XPV, XSN, XSV, XSA, and XR. This categorization includes nine subsets in *josa* and 14 in *eomi*, which are counted in each sentence. The data reveal which ones are the most frequent and the least frequent in English and Korean. In this study, all punctuations were removed from the data.

2) A Kokoma Korean morphological analyzer, one of the most necessary parts in natural language processing systems, is used. This is an open tool available for online use and as a downloadable application. The Kokoma Korean morphological analyzer tags the Korean script according to nouns, verbs, and adjectives. They were categorized into 45 groups by tagging the parts-of-speech: NNG, NNP, NNB, NNM, NR, NP, VV, VA, VXX, VXA, VCP, VCN, MDN, MDT, MAG, MAC, IC, JKS, JC, JKG, JKO, JKM, JKI, JKQ, JC, JX, EPH, EPT, EPP, EFN, EFQ, EFO, EFA, EFI, EFR, ECE, ECS, ECD, ETN, ETD, XPV, XSN, XSV, XSA, and XR. This categorization includes nine subsets in *josa* and 14 in *eomi*, which are counted in each sentence. The data reveal which ones are the most frequent and the least frequent in English and Korean. In this study, all punctuations were removed from the data.

3) The subtitles are scored by Korean American bilinguals. The scores were graded from “most corrected” (5 points) to “least corrected” (1 point), with points 5-3 indicating correct translation and points 2-1 as wrong translations. Although the subtitles are grammatically wrong or awkward, if they are understandable and usable in an informal situation, the subtitles are grouped into over three points. The least corrected of the two groups that have point 1 and 2 scores are analyzed to determine which POS has the most common errors.

The four participants who score the subtitles are the bilinguals who teaches English in a Korean University. They have experience staying in America over five years and they could speak both Korean and English fluently.

---

1 Please see the Appendix for the complete Kokoma Korean POS tag comparison chart.
1. Data Collection

The Korean-English subtitle and Korean of the Korean TV drama first were recorded into an Excel document and analyzed by this paper’s researcher. There are over ten websites in the United States running Korean TV dramas on the Internet: Hulu.com, Viki.com, Dramafever.com, mysoju.com, mykofan.com, hancinema.net, gooddrama.net, dramacrazy.com, mvibo.com, etc. Among them, Korean-English subtitles of Hulu.com and Viki.com are examined. Hulu is an ad-supported on-demand streaming video website of TV shows and movies, and is currently offered only to users in the U.S. and is blocked by IP address locations from outside the U.S. Viki is a video streaming website based in Singapore that provides on-demand streaming video of TV shows and movies from around the world. It is also the first and fastest platform for subtitling of video, counting on a community of thousands of volunteer translators (“Viki”, n.d.).

The Korean romantic drama The Winter, The Wind Blows ran from February to March 2013 by SBS, which is TV series starring Jo In-sung and Song Hye-kyo. This story is about what a gambler pretends the lost brother of a blind heiress, but these two people fall in true love after getting to know each other. The drama has a total of 16 episodes, and the object of this study is episode one and eight. Episode one has each 467 subtitles in Hulu.com, Viki.com and episode eight has 385 subtitles in both. The total number of the subtitles is 1,704.

2. Data Analysis

This research shows that the content words such as noun, verb, adjective, and adverb have even more frequent errors than function words such as article, conjunction, preposition, pronoun in English. The rank of the most frequent errors is PRP, NN, VB, RB, DT, and IN. PRP and NN are ranked first and second, respectively (see Table 1).

| POS  | Percentage (%) |
|------|----------------|
| RP   | 60             |
| NN   | 50             |
| VB   | 43             |
| RB   | 37             |
| DT   | 35             |
| IN   | 35             |

According to the analysis of the Openkoreatext, the most frequent is content words such as Noun, Verb, and Adjective in Korean which is the same line as in English. The noticeable thing in Korean is the errors of josa and eomi which is postposition particle having in agglutinative language. Josa is a Korean noun-ending while eomi is a Korean verb-ending, which would mean high error frequency of josa and eomi indispensable because they are morphologically dependent on stems such as nouns and verbs (see Figure 3).

![Figure 3](image-url)
At this point of the research, it is needed to investigate *josa* and *eomi* more detailed and specify. The version of Kokoma (KKMA) provided by KoNLPy has a large set of 56 tags versus 14 tags for the Openkoreatext. In this research, KKMA Korean morphological analyzer is used for the *josa* and *eomi*. *Josa* is subcategorized into nine groups such as JKS, JKC, JKG, JKO, JKM, JKI, JKQ, JC, and JX and the data is as follows (see Table 2):

**TABLE 2**

| Type | Correct translation | Wrong translation | Difference (%) |
|------|---------------------|-------------------|----------------|
| JKS  | 28.49               | 26.32             | -              |
| JKC  | 0.84                | 1.75              | -0.91          |
| JKG  | 3.02                | 5.26              | -2.24          |
| JKO  | 14.31               | 18.42             | -4.11          |
| JKM  | 21.02               | 25.44             | -4.42          |
| JKI  | 0.89                | 0.88              | -              |
| JKQ  | 0.67                | 0.00              | -              |
| JC   | 2.80                | 1.75              | -              |
| JX   | 27.96               | 20.18             | -              |

*Mean Value* -2.92

The data shows that the most frequent errors are JX, JKM, JKO, and JKG. *Eomi*, the verb-ending, is largely categorized into four groups in the following: EP, EF, EC, and ET.

**FIGURE 4**

*The Errors of the Four Types of Eomi*

*Eomi* is subcategorized into fourteen such as EPH, EPT, EPP, EFN, EFQ, EFO, EFA, EFI, EFR, ECE, ECS, ECD, ETN, ETD and the data is as follows:

**TABLE 3**

| Type | Correct translation | Wrong translation | Difference (%) |
|------|---------------------|-------------------|----------------|
| EPH  | 2.10                | 0.36              | 1.33           |
| EPT  | 9.10                | 11.24             | 2.13           |
| EPP  | 0.00                | 0.00              | 0.00           |
| EFN  | 16.11               | 12.36             | 3.75           |
| EFQ  | 3.14                | 2.25              | 0.90           |
| EFO  | 0.27                | 1.12              | 0.85           |
| EFA  | 0.30                | 0.00              | 0.30           |
| EFI  | 0.30                | 0.56              | 0.26           |
| EFR  | 0.07                | 0.00              | 0.07           |
| ECE  | 12.75               | 14.04             | 1.30           |
| ECS  | 12.12               | 12.36             | 0.24           |
| ECD  | 25.27               | 25.84             | 0.57           |
| ETN  | 1.20                | 1.12              | 0.07           |
| ETD  | 17.26               | 18.54             | 1.28           |
The above result reveals that the most frequent errors are EFN as 3.75%, and the next orders are EPT, EPH, ECE, and ETD.

IV. DISCUSSION

1. Josa (Korean Noun-Ending)

Josa. Korean noun-ending is particle which comes after a word or clause to represent a certain relationship with another word on the sentence, or add a certain meaning to a word on the sentence. It has case particle, connective particle and auxiliary particle. Particles are not used alone and they should necessarily be attached to the end of a noun.

The form of particles does not change at the place where they are. The most frequent errors are JX which has meanings as in the case of adverb or English prepositions and complete its meaning to the preceding word. It can be added to objects, complements, adverbs as well as subjects, and readers or speakers should be paid attention to the JX.

![Correct Translation and Wrong Translation of Josa](image)

TABLE 4

| Tagging                                      | Korean-to-English translation                        |
|----------------------------------------------|------------------------------------------------------|
| [(왜, MAG), (나, NP), (데, JX), (하, VV), (의, ECD), (?, SF)] | Why are you only stomping on me?                     |
|                                               | Why only bully me?                                   |
| [(왜, NNG), (왜, NNG), (이, NNG), (입, JKM), (예, JX), (كه, VV), (는, ETD), (수, NNB), (은, MAG), (안, MAG), (을, VV), (는, ECD)] | Normally, people are not named chikll Soo but don’t you this it’s cool? Normally, we don’t name a person after Chikll Soo. |
| [(원래, NNG), (이, JKM), (풀, VV), (이, ECD)] | For one year, rot in jail                          |
| [(원래, NNG), (이, JKM), (풀, VV), (이, ECD)] | For one year, rot in hell                           |
| [(이렇게, NP), (원, NNM), (데, JX)] [(원, NNG), (예, JKM), (풀, VV), (이, ECD)] | “Parkland. Celebrating 24 Years of quality clothing at great prices” |
| [(이렇게, NP), (원, NNM), (데, JX)] [(원, NNG), (예, JKM), (풀, VV), (이, ECD)] | “Parkland. Celebrating 24 Years of quality clothing at great prices” |

JK (Josa for case marker), case particle affixed to the end of a noun plays a certain role in the sentence. As mentioned before, it has seven particles and account for the most frequent errors such as JKM, JKO, and JKG. JKM,
Adverbial Particle is used in a wide range of forms such as -와/과, -에(서), -(으)로, -에서, etc. and change the word with those kinds of particles into adverbial particles.

### TABLE 5
The Errors of the JKM (Adverbial Particle)

| Tagging | Korean-to-English translation |
|---------|--------------------------------|
| [(희흥, NNG), (님, XSN), (사, TV), (이, ECD), (생건, NNG), (예, JKM), (있어, NNG), (의, JKG), (결혼, NNG), (을, JKO)] | I think while the CEO lives, Young’s and my wedding. I think when CEO lived (CEO was once alive), he married Young. |
| [(원해, NNG), (사람, NNG), (이름, NNG), (의, JKM), (는, JX), (저기, VV), (는, ETD), (의, NNB), (을, MAG), (안의, MAG), (는, ECD)] | Normally, people are not named chik-II Soo but don’t you this it’s cool? Normally, we don’t name a person after Chik-II Soo. |
| [(يوم, NNG), (월, JKM), (오른쪽, NNG), (으로, JKM), (10, NR), (센터, NNG), (의, MAG), (가, VV), (이, ECS), (는, VVX), (세계, EFN)] | Next time, I’ll aim 100 cm the right. Next time, I’ll aim 100 cm to the right. |
| [(우리, NP), (PL, OL), (그룹, NNG), (은, JX), (의, MDT), (사태, NNG), (가, JKS), (원히, VV), (는, ETD), (갑각적, NNG), (이, VCP), (는, ETD), (디자인, NNG), (을, JKO), (최고, NNG), (의, JKG), (배한, NNG), (으로, JKM)] | “Parkland. Celebrating 24 Years of quality clothing at great prices” |
| [(24, NR), (년, NNM), (건, NNG), (의료, NNG), (사업, NNG), (무론, NNG), (예서, JKM), (결혼, NNG), (생대, NNG), (하, XSV), (는, ETD), (정장, NNG), (을, JKO), (하, VV), (의, ECS), (은, MDT)] | For the past 24 years, he has made tremendous progress in the fashion industry. He has made tremendous progress in the fashion industry for the past 24 years. He has made tremendous progress in the fashion industry within 24 years. |

The JKO can be divided into direct object particle and indirect object particle. They play roles making the preceding word the object of the sentence. Direct object particles such as -(을/를) are the phonological allomorph. Although they have the different forms, they have the same sentence function, which could lead to confusion among speakers/writers.

### TABLE 6
The Errors of the JKO (Object Postposition)

| Tagging | Korean-to-English translation |
|---------|--------------------------------|
| [(희흥, NNG), (님, XSN), (사, TV), (이, ECD), (생건, NNG), (예, JKM), (있어, NNG), (의, JKG), (결혼, NNG), (을, JKO)] | I think while the CEO lives, Young’s and my wedding. I think when CEO lived (CEO was once alive), he married Young. |
| [(우리, NP), (PL, OL), (그룹, NNG), (은, JX), (의, MDT), (사태, NNG), (가, JKS), (원히, VV), (는, ETD), (갑각적, NNG), (이, VCP), (는, ETD), (디자인, NNG), (을, JKO), (최고, NNG), (의, JKG), (배한, NNG), (으로, JKM)] | “Parkland. Celebrating 24 Years of quality clothing at great prices” |
| [(24, NR), (년, NNM), (건, NNG), (의료, NNG), (사업, NNG), (무론, NNG), (예서, JKM), (결혼, NNG), (생대, NNG), (하, XSV), (는, ETD), (정장, NNG), (을, JKO), (하, VV), (의, ECS), (은, MDT)] | For the past 24 years, he has made tremendous progress in the fashion industry. He has made tremendous progress in the fashion industry for the past 24 years. He has made tremendous progress in the fashion industry within 24 years. |

JKG, genitive postposition, is like an apostrophe + s (’s), and makes the preceding word an adnominal phrase of the sentence and which is often called adnominal particle. However, -(의), Korean genitive postposition has diverse meaning functions depending on the surrounding words such as ‘have by someone’, ‘written by someone’, and ‘about something’ etc. Consequently, it could lead to wrong translation.

### TABLE 7
The Errors of the JKG (Genitive Postposition)

| Tagging | Korean-to-English translation |
|---------|--------------------------------|
| [(희흥, NNG), (님, XSN), (사, TV), (이, ECD), (생건, NNG), (예, JKM), (있어, NNG), (의, JKG), (결혼, NNG), (을, JKO)] | I think while the CEO lives, Young’s and my wedding. I think when CEO lived (CEO was once alive), he married Young. |
| [(우리, NP), (PL, OL), (그룹, NNG), (은, JX), (의, MDT), (사태, NNG), (가, JKS), (원히, VV), (는, ETD), (갑각적, NNG), (이, VCP), (는, ETD), (디자인, NNG), (을, JKO), (최고, NNG), (의, JKG), (배한, NNG), (으로, JKM)] | “Parkland. Celebrating 24 Years of quality clothing at great prices” |

© 2022 The Society for Teaching English through Media (STEM)
An Analysis of Josa and Eomi in Translating Korean TV Dramas Into English With Artificial Intelligence

2. Eomi (Korean Verb Ending)

Korean verb must be affixed with an ending to form an eojel. Eomi, Korean verb ending is largely divided into two categories depending on the position and function of the eomi: EP on position and EF, EC, ET on function. Also, when the verb stem is affixed with EF and EC, it has the function of predicate. On the other hand, when combined with an ET, it has the function of modifier.

EP is positioned in the middle of a word as prefinal ending that attached to verbs and placed between the stem and the final ending. It is subcategorized into three group such as EPH, EPT, and EPP. EPH is a honorific prefinal ending for older people such as ‘-시’ and ‘-사오- and the speaker honors the subject of the sentence. EPT is a tense ending such as ‘-은’ and ‘-였’ and EPP is a polite particle such as ‘-옵’. Eomi is categorized into three groups depending on the function of the eomi such as EC, EF, and ET and it is subcategorized into eleven group. The finding shows EPT is the most frequent ending among EPH, EPT, and EPP.

As shown in Table 8, among the three kinds of tense, the data shows most errors are shown in the past tense, ‘-였’ which is EPT.

| Tagging | Korean-to-English translation |
|---------|-------------------------------|
| [( carry, NNG), (ወ, NP), (행, NNG), (이, JJ), (이, NNG), (을, VA), (을, ETD), (이, NNB), (을, VA), (이, ECS), (을, VV), (을, EPT), (이, EFN)] | I told you to change your name, didn’t I? I told you I didn’t like the fact that I have same name as my brother. |
| [(가, VV), (의, ECE), (의, NNG), (가, JX), (와, ECP), (을, VXV), (의, ECN), (의, EFS), (의, EFN)] | That.. That was not when mom was alive. |
| [(가, VV), (와, ECP), (의, EFN), (? SF), (의, NNG), (을, XSN), (의, MAG), (을, VV), (의, EFQ), (의, NNG), (? SF)] | Can you not hear? Bring all the guys! Soo is here, Soo! Are you deaf? Told you to call on/gather all the guys now? |
| [(가, VV), (의, ECE), (의, NNG), (가, JX), (와, ECP), (의, VXV), (의, ECN), (의, EFS), (의, EFN)] | That.. That was not when mom was alive. |
EF (Final Ending) is attached at the end of the sentence and has the function of sentence completion and describes the sentence types. EF shows six types of the sentence such as EFN, EFQ, EFO, EFA, EFI, and EFR. EFN has sixty declarative ending. Consequently, among them, the most frequent error is the EFN such as -다: da, -에: ne, -지: -ji, -이: -eo, with a large gap.

|TABLE 9|
The Errors of the EFN (Final Declarative Ending)

| Tagging | Korean-English Translation |
|----------|----------------------------|
| ([직급, MAG), (③, VV), (을, ETD), (때, NNG), (아니, VCN), (다, EFN)] | This isn’t the time to smile. It’s no time for a joke. |
| ([역수, NNG), (가, JKS), (의, VV), (네, EFN)] | There seems to be a lot of them. |
| ([이사, NNG), (주리, NP), (할, NNG), (의셔, JC), (의 gọn, NNG), (를, VA), (는, ETD), (가, NNB), (של, VA), (더, ECS), (의, VV), (이, EPT), (의, EFN)] | I told you to change your name, didn’t I? I told you I didn’t like the fact that I have same name as my brother. |
| ([다음, NNG), (예, JKM), (오른쪽, NNG), (으로, JKM), (10, NR), (형, NNG), (의, MAG), (가, VV), (의, ECS), (는, VVV), (가, EFN)] | Next time, I’ll aim 10cm the right. Next time, I’ll aim 10cm to the right. |
| ([이먹, VV), (것, EPT), (의, EFN), (? SF), (의, NNG), (는, XSN), (다, MAG), (를, VV), (것, EFO), (이고, NNG), (!, SF)] | Can you not here? Bring all the guys! Soo is here, Soo! Are you deaf? Told you to call on/gather all the guys now? |
| ([폭수, NNG), (가, JKS), (의, VV), (네, EFN)] | There’s a lot of them. |

EC connects the preceding words (word, phrase, or clause) to the following words and has the most frequent errors. It is subcategorized into ECD, ECE and ECS. The most frequent errors are shown in the ECD which has a linking role such as -지: -ji, -어요: -s, -니: -ni, -는데: -nde, -게: -ge, etc. ECD is so diverse that it is more likely to make errors in translation.

|TABLE 10|
The Errors of the ECD (Connective Dependence Ending)

| Tagging | Korean-English translation |
|----------|----------------------------|
| ([특장, NNG), (닐, XSN), (살, VV), (있는, ECD), (생존, NNG), (의, JKM), (의, NNG), (와, JKM), (의, JKG), (한, XSN), (을, ECS), (된, JKO)] | I think while the CEO lives, Young’s and my wedding. I think when CEO lived (CEO was once alive), he married Young. |
| ([에, MAG), (비, NP), (르, JX), (의, VV), (의, ECD)] | Why are you only stomping on me? |
| ([그려, VV), (는, ETD), (가, NNB), (킬, VA), (는, ECD), (? SF)] | Is that right? Isn’t it right? Isn’t it so? |
| ([일, NNG), (의식, NNG), (의, JKM), (는, JX), (저기, VV), (는, ETD), (가, NNB), (같, MAG), (한, MAG), (는, VV), (하는, ECD)] | Normally, people are not named chik-ll Soo but don’t you this it’s cool? Normally, we don’t name a person after Chik-ll Soo. |
| ([그, VV), (건, ECE), (의마, NNG), (가, JKS), (있는, VV), (의, ECD), (개시, VXA), (는, ETD), (한, NNG), (가, JKS), (의마, VVV), (는, EPT), (나소, EFO)] | That.. That was not when mom was alive. |
| ([태동, MAG), (의마, VA), (개, ECD), (그렇, MAG), (할, VV), (전, ECE), (한, MAG), (할, VV), (할, ECS)] | I can’t just live like this. Can’t I just take it easy? |
| ([그려, VV), (는, ETD), (가, NNB), (킬, VA), (는, ECD), (? SF)] | I guess not. Isn’t it? |
| ([이먹, VA), (의, ECD), (있, NNG), (의, JKG), (있, NNG), (과출부, NNG), (한, JX), (할, VV), (하고, ECE), (할, VV), (의, ECD), (을, VVV), (하고, ECE)] | Why are you working as a housemaid? You should not live as housemaid. |
| ([말, VV), (나는, ECD), (은, SF), (의, MAG), (로켜부, NNG), (가, JKS)] | I saw the mailman… No, Before the mailman... |
| ([직, NNG), (고, VV), (의, ECD), (의, NP), (할, VV), (여리, EFO)] | Just stay there. Stay quiet. |

ECD is coordinative connective ending after the stem connects the preceding words to the following words coordinately such as -بدل/ -myeon-, -고/go, -어/-e yeo, -이/ -eu myeo, -자/ -ja/ -ja, -만/ -jiman, -건/ -geon and has the third frequent errors. They list two or more facts and connect several clauses with similar or contrast meaning.
An Analysis of Josa and Eomi in Translating Korean TV Dramas Into English With Artificial Intelligence

TABLE 11
The Errors of the ECE (Connective Coordinative Ending)

| Tagging                      | Korean-to-English translation |
|------------------------------|-------------------------------|
| [(일, VV), (나, ECE), (이, NNG), (존여, MAG)] | Ah! This hyung, serious! Hey, you dude! |
| [(일, VV), (간, ECE), (이름, NNG), (가, JKS), (살, VV), (아, ECD), (게시, VX), (2, ETD), (재, NNG), (가, JKS), (나, NNG), (의, EFN), (2, ETD), (어, EFN)] | That... That was not when mom was alive. |
| [(메, MAG), (이름, VA), (게, ECD), (그간, MAG), (살, VV), (번째, ECE), (안, MAG), (의, VV), (어, ECS), (?, SF)] | I can’t just live like this. Can’t I just take it easy? |
| [(나, MAG), (의, VV), (2, ETD), (배부, NNG), (경보, NNG), (님라, VV), (고, ECE), (님리, NNG), (가, JKS), (나, VV), (고, ECE)] | Why are you working as a housemaid? You should not live as a housemaid. |
| [(메, MNN), (2, SF), (내, NP), (7, JKS), (존여, MAG), (바])+, VA, (내, ECE)] | No. I’m busy. Yes, I’m a little busy at the moment. |
| [(바), MAG), (의, VV), (2, ETD), (배부, NNG), (경보, NNG), (님라, VV), (고, ECE), (님리, NNG), (가, JKS), (나, VV), (고, ECE)] | They all wanted detailed information. They were all in a panic for detailed information. |
| [(일, VV), (간, ECE), (이름, NNG), (가, JKS), (살, VV), (아, ECD), (게시, VX), (나, ETD), (내, NNG), (가, JKS), (나, NNG), (의, EFN), (2, ETD), (어, EFN)] | That... That was not when mom was alive. |

ECS is the subsidiary linking eomi such as -이,-이요/-이, eo, yeo, -개/기, -지고/jigo and has the fifth frequent errors. Linking Emoi -으 in Korean could be declarative or question depending on the type of the sentences.

TABLE 12
The Errors of the ECS (Subsidiary Linking Ending)

| Tagging                      | Korean-to-English translation |
|------------------------------|-------------------------------|
| [(내가, NNG), (꼬리, NP), (의, NNG), (이명, JC), (이름, NNG), (갑, VA), (은, ETD), (거, NNB), (살, VA), (탄, ECS), (한, VV), (일, EPT), (식, EFN)] | I told you to change your name, didn’t I? I told you I didn’t like the fact that I have same name as my brother. |
| [(내, NP), (꼬리, NNG), (안, NNG), (公网, VV), (와, ECS), (?, SF)] | You’re not painting? Aren’t you painting? |
| [(이름, NNG), (에, JKM), (오른쪽, NNG), (으로, JKM), (10, NR), ( siti, NNG), (다, MAG), (가, VV), (와, ECS), (과, VV), (과, EFN)] | Next time, I’ll aim 10 cm to the right. |
| [(베, MAG), (이름, VA), (게, ECD), (그간, MAG), (살, VV), (번째, ECE), (안, MAG), (의, VV), (와, ECS), (?, SF)] | I can’t just I’ve like this. Can’t I just take it easy? |
| [(24, NR), (년, NNN), (간, NNG), (의료, NNG), (사업, NNG), (상태, NNG), (하, XSV), (2, ETD), (상장, NNG), (을, JKO), (하, VV), (과, ECS), (은, MDT)] | For the past 24 years, he has made tremendous progress in the fashion industry. |

ET is called Conversion Ending and is attached to the verb stem of the verb or adjective and make it perform the function of another part of speech. ET is divided into ETN and ETD. ETN is attached to the verb stem and change it into abdominals. ETD can be grouped by tense and aspect such as present-continuous, past-perfect, past-retrospect, past perfect-retrospect, future-guessing/will/ability. It leads to confusion to use. The data shows that the most frequent errors are ETD.

TABLE 13
The Errors of the ETD (Adnominal Transformative Ending)

| Tagging                      | Korean-to-English translation |
|------------------------------|-------------------------------|
| [(지금, MAG), (또, VV), (또, ETD), (때, NNG), (아니, VCN), (또, EFN)] | This isn’t the time to smile. It's no time for a joke. |
| [(그리, VV), (2, ETD), (가, NNB), (간, VA), (저, ECD), (?, SF)] | Is that right? Isn’t it? |
| [(내가, NNG), (꼬리, NP), (의, NNG), (이름, JC), (이름, NNG), (간, VA), (2, ETD), (가, NNB), (살, VA), (탄, ECS), (한, VV), (일, EPT), (식, EFN)] | I told you to change your name, didn’t I? I told you I didn’t like the fact that I have same name as my brother. |
| [(지키, VV), (2, ETD), (2, NNB), (?, SF)] | Chik-II Soo? |

Journal of English Teaching through Movies and Media, Vol. 23, No. 2. (2022)
V. CONCLUSION

The goal of this study was to find out why the errors happens frequently in the subtitles of Korean TV drama. This finding showed that the content words have even more frequent errors than function words in English as hypothesized. The most frequent errors were ranked PRP, NN, VB, RB, DT, and IN. Likewise, the most frequent errors were content words such as noun, verb, adjective in Korean. The result of the error analysis through POS tagging and scoring showed that the most common errors are Korean function words are EPT, EPH, ECE, and ETD. As mentioned before, Korean has rich morphology and agglutinative nature, and it is regarded as the hardest aspect in translating Korean into English or vice versa. The morphological analysis is to ascertain morphemes and assign pertinent information to them.

This is conducted for Korean language among agglutinating languages, but if this study is conducted for other agglutinating languages such as Turkish and Japanese etc., it would be meaningful and worthwhile to enlighten the method of teaching second language. In addition, it is suggested that the future study should be conducted based on this study. For example, after L2 learners are taught the basic knowledge of Korean functional words, on josa or eomi, they are tested to translate Korean sentences into English to demonstrate their “before and after” degree of improvement in English writing. Further, as the importance of the functional words places an emphasis on Korean-to-English translation, our findings are also useful for machine translation of agglutinating languages in the Fourth Industrial Revolution. Ryoo and Cho (2017) mentions “translation is a special kind of cross-cultural communication, involving two linguistically different cultures. Cross-cultural awareness means the translator’s perception of the cultural elements of the languages involved in the process of translation” (p. 96). It could have a critical effect on the quality of subtitles translated Korean-to-English as Korean culture and media has become one of the most important exports in the world.

REFERENCES

Ali, A. (2016). Exploring the problems of machine translation from Arabic into English language faced by Saudi university student of translation at the faculty of arts, Jazan University Saudi Arabia. IOSR Journal of Humanities and Social Science, 21(4), 55-66. https://doi.org/10.9790/0837-2104025566
APPENDIX

| Tag   | Description                        | Tag   | Description                        |
|-------|------------------------------------|-------|------------------------------------|
| CC    | Coordinating conjunction           | RB    | Adverb                             |
| CD    | Cardinal number                    | RBR   | Adverb, comparative                |
| DT    | Determiner                         | RBS   | Adverb, superlative                |
| EX    | Existential there                  | RP    | Particle                           |
| IN    | Preposition or subordinating conj. | TO    | to                                 |
| JJ    | Adjective                          | VB    | Verb, base form                    |
| JJR   | Adjective, comparative             | VBD   | Verb, past tense                   |
| JJS   | Adjective, superlative             | VBG   | Verb, gerund, or present participle|
| MD    | Modal                              | VBN   | Verb, past participle              |
| NN    | Noun, singular or mass             | VBP   | Verb, non-3rd person singular present|
| NNS   | Noun, plural                       | VBZ   | Verb, 3rd person singular present  |
| NNP   | Proper noun, singular              | WDT   | Wh-determiner                      |

Penn Treebank Part-of-Speech Tags
| KKMA (ntags=30) | Tag | Description |
|-----------------|-----|-------------|
| **N** | NN (noun) | NNG (general noun) |
| | NNP (proper noun) | NNB (dependent noun) |
| | NNM (bound noun) |  |
| | NR (numeral) | NR (numeral) |
| | NP (pronoun) | NP (pronoun) |
| **V** | VV (verb) | VV (verb) |
| | VA (adjective) | VA (adjective) |
| | VX (auxiliary) | VXV (auxiliary verb) |
| | | VXA (auxiliary adjective) |
| | VC (copular) | VCP (positive designator) |
| | | VCN (negative designator) |
| **M** (modifier) | MD | MDN (numeral determiner) |
| | MDT (common determiner) | MDT (common determiner) |
| **I** (interjection) | IC | IC (interjection) |
| **J** (josa-postposition) | JK (josa for case marker) | JKS (subjective postposition) |
| | JKC (complementary postposition) | JKG (genitive postposition) |
| | | JKO (object postposition) |
| | | JKM (adverbial particle) |
| | | JKl (josa for vocative case) |
| | | JKQ (josa for quotative case) |
| **JC** (connective postposition) | JC (connective postposition) |  |
| | **E** (ending) | EP (prefinal ending) | EP (prefinal ending) |
| | | EPT (tense prefinal ending) |
| | EF (final ending) | EFN (declarative final ending) |
| | | EFO (imperative final ending) |
| | | EFA (propositive final ending) |
| | | EFI (exclamatory final ending) |
| | | EFR (final ending) |
| | **X** | XP | XP (substantive prefix) |
| | | XPV |
| | **XR** (radix) | XR (radix) |

Note. Team KKMA. (n.d.). Korean POS tags chart. Retrieved from http://kkma.snu.ac.kr/documents/?doc=postag

Note. Williams, S. (2013). An analysis of POS tag patterns in ontology identifiers and labels (Technical report No. 2013/02). Retrieved from http://computing-reports.open.ac.uk/2013/TR2013-02.pdf