

본 논문은 아래의 저작권 정책을 가지고 있으니, 이용에 참고하시기 바랍니다.

• 저작권 정보 (Copyright Policy)

-학술지 발행기관

• 재사용 정보 (CC License)



-저작자를 밝히면 자유로운 이용이 가능하지만 영리목적으로 이용할 수 없습니다.

• 셀프아카이빙 정보 (Author Self-Archiving)

-Gray : 검토 중 · 비공개 · 무응답 · 기타

• 원문 접근 정보 (Reader Rights)

-이용자 접근정책 : CCL 유형에 따른 재사용 가능

-무료 DB : 과학기술학회마을 / KCI / 국외 : EBSCO /

A Comparative Study of ChatGPT and Papago Korean-to-English Translations of Dialogue from the Adaptation of Pachinko

Jeong-Hwa Lee ^{1,*}

¹ Dankook University; Visiting Professor; 2019jhlee@naver.com

* Correspondence

<https://doi.org/10.5392/IJoC.2025.21.3.139>

Manuscript Received 26 February 2025; Received 4 July 2025; Accepted 12 August 2025

Abstract: The purpose of this study is to compare the performance of two Korean-to-English machine translations of Pachinko: ChatGPT-3.5 (ChatGPT) and Naver Papago (Papago). The drama dialogues were analyzed based on 590 pairs of samples translated using both ChatGPT and Papago to identify translation accuracy as well as lexical, grammatical, omission, and mechanical errors. These errors were further categorized into 11 subtype error patterns. The research found that ChatGPT produced 343 correct instances (58.1%) out of 590 attempts, while Papago achieved 223 correct translations (37.8%) for the same dataset. With regard to the four categories examined in this study, ChatGPT generated 190 lexical errors, 31 grammatical errors, 18 omissions, and 8 mechanical errors. In contrast, Papago yielded 301 lexical errors, 48 grammatical errors, 10 mechanical errors, and 8 omissions. Among the 11 error subtypes, the most frequent were incorrect word meaning errors, with 148 instances for ChatGPT and 226 for Papago, followed by pronoun errors, observed in 23 instances for ChatGPT and 40 for Papago. This comprehensive analysis of the primary error classifications reveals limitations in both ChatGPT and Papago. These findings underscore the need for future updates to enhance sensitivity to the unique aspects of the Korean language in both translation systems.

Keywords: Korean-to-English Machine Translations; Pachinko; ChatGPT; Papago; Errors

1. Introduction

The Fourth Industrial Revolution has enabled people of different nations to access Korean cultural content through over-the-top (OTT) media services like Netflix, Disney Plus, Apple TV, Watcha, and Wavve. With the development of machine translation, OTT services provide auto-generated subtitles [1]. However, since various human languages often transcend simple semantic communication, and language is deeply rooted in human-to-human relationships as well as socio-cultural phenomena, it is difficult for translation software to mechanically convert one language to another.

Korean spoken language is eclectic in its usage, often imagined as a series of connections mixed with rags [2], to reflect that it does not follow the logical and grammatical requirements inherent in Korean's written form. With regard to the sentence structure, the spoken language is complicated by the previously delineated human-to-human relationships and sociocultural phenomena that often go unspoken. It is this feature of spoken word that affects the accuracy of translation and triggers translation errors that occur in Korean-to-English translations when utilizing AI translators [3].

Research on the quality of Korean-to-English translation using machine translators gained attention with the introduction of artificial neural network translation systems (ANNTS) [4]. With the release of Chat GPT-3.5 (ChatGPT), in November 2022, interest in the use of cutting-edge artificial intelligence driven by user prompts has revolutionized language translation tools [5, 6]. However, there is currently a dearth of research on the features and performance of ChatGPT's Korean-to-English translation capability compared to its ANNTS forbearers such as Papago. Specifically, such research can offer valuable insights into the strengths and weaknesses of ChatGPT and Papago in Korean conversational translations to guide future AI translation research.

Pachinko depicts the social situation in Korea, Japan, and the United States between 1910 and 1989. As such, dialogue is eclectic in its use of three languages including a variety of Korean dialects, honorifics, colloquial speech, and proverbs depicting the dynamic and evolving nature of language use reflective of the time. Pachinko has a significant proportion of Gyeongsang Province dialect. For instance, the Gyeongsang dialect uses distinctive interrogative endings such as “나?” (na?), “노?” (no?), “가?” (ga?), and “고?” (go?) in response to pitch changes in phrases or words [7]. Moreover, it is unique in that it addresses people with respect by utilizing the honorific particle “예” (ye). In this way, the unique linguistic features used in “Pachinko” serve as important elements that reflect the historical, regional, and cultural context. This linguistic complexity poses a significant challenge, especially in the translation process, and requires special attention and analysis to ensure an accurate and natural translation.

This investigation seeks to differentiate itself from earlier studies on Korean-to-English machine translation in Korea that relied on the growth of increasingly complex decision trees for translating Korean into English. In this study, the researcher used ChatGPT-3.5 (ChatGPT) and Naver Papago (Papago) English translations of Korean dialogue spoken by the characters in two episodes of the Apple TV serial Pachinko to evaluate the characteristics of these English translations. This research focus is described by the following three questions.

1. How accurate are ChatGPT and Papago’s Korean-to-English translations of the TV drama “Pachinko”?
2. How often do vocabulary, grammatical, and content issues occur in ChatGPT’s and Papago’s translations of these dialogues?
3. What is the frequency of 11 subtypes of error patterns, with illustrative examples, when utilizing ChatGPT and Papago to translate the Korean dialogues in Pachinko?

2. Literature Review

2.1 Global Trends in AI-Based Translation Research

Diverse AI-based chatbots and machine translators are being utilized in current studies to evaluate translation quality across multiple languages and subjects. Rao, McGee, and Seideman [8] compared the accuracy of ChatGPT and Google Translate (GT) in translating medical texts into various languages. ChatGPT performed better than GT in translating English documents into Spanish, with a minimized error rate in Spanish. However, GT did better in translating Vietnamese. The two algorithms generated low-quality translations (35.6% and 41.6%) for Russian.

Aeni, Baharuddin, Putera, and Melani [9] investigated ChatGPT’s translation accuracy in translating the 2023 Lana Journal’s Indonesian content into English, based on Snover’s HTER theory, and Schiaffino and Zearo’s translation quality index. The results showed 93.6% success rate with only 0.73% insertion errors, 1.32% deletion errors, 3.54% substitution errors, and 0.83% movement errors.

Gu [10] examined how the semantic role of transformative nouns—the linguistic issue that underlies the difficulty of translating attribution clauses effectively in Japanese-to-Chinese translation linguistics—affects the choice of translation patterns. A pre-edit approach is proposed to improve translation accuracy to overcome these issues. Additionally, they suggested a brand-new two-step prompt technique that combines ChatGPT—currently the most popular big language model—with their pre-edit scheme. They reported this prompt strategy has been shown to benefit translation input efficiency in zero-shot situations.

For translating ancient Chinese poetry, Gao, Lin, Zhao, and Cai [11] assessed ChatGPT’s translation using two distinct prompts against Google Translate and DeepL Translator in terms of faithfulness, fluency, and language style. Based on the findings, ChatGPT performed better than other systems in all categories and, when asked to keep the poem’s rhythm and meter, showed a remarkable capacity to preserve this structure of poetic language. Furthermore, it showed that ChatGPT was able to understand the poem’s semantic, imagistic, and figurative components by retaining them in a cohesive and smooth translation. In the findings of this study, ChatGPT appears to be superior to the current machine translation tools evaluated when it comes to comprehending classic works of literature.

Khoshafah [12] revealed that ChatGPT-3.5 translated basic Arabic text into English but broke down for some sophisticated translations that call for human assistance by contrasting its output with that of experienced translators. While ChatGPT-3.5 usually yields good translations, texts such as legal documents, medical reports,

scientific studies, and literary works employing figurative context, may not be efficiently and effectively translated. In these instances, human intervention and caution will be required to ensure accuracy.

Various studies have documented that ChatGPT's translation performance differs depending on the language and text type. For instance, ChatGPT translated well between Spanish and Indonesian, exhibiting high accuracy and low error rates. It also showed promise for usage in Spanish medical translation. In regard to Chinese and Japanese, ChatGPT translated attributive sentences from classical Chinese and Japanese poetry with remarkable grammatical fluency and poetic structure. However, its performance was underwhelming with high error rates when translating Russian and Vietnamese to English. Regarding Arabic to English translations, ChatGPT comprehended fundamental contexts in Arabic translations, but it was hampered in its ability to translate literary, legal, and medical writings. These findings suggest that ChatGPT requires more advanced algorithms to be versatile across all languages and text varieties.

2.2 Research Trends on Quality of ChatGPT Based on Korean and English

Research on the quality of Korean-to-English translations using ChatGPT has been conducted on a variety of texts. Park and Choi [13] explored whether ChatGPT might translate irony, a literary instrument that is challenging to comprehend, and compared it with existing machine translation and human translation to assess and evaluate the impact of ChatGPT's interactive features on translation quality. Machine translation has faced difficulties in identifying and translating irony, using appropriate strategies. However, ChatGPT, while not translating as accurately as a human, has shown a more accurate interpretation and translation of the original meaning and intent compared to other machine translation methods.

Yoon [14] converted three distinct types of texts using ChatGPT: a literary, Yoon Dong-ju's poetry "Self-Portrait," a linguistics textbook [15], and 21 significant Korean sayings from a Korean textbook for foreigners. Four native English speakers assessed the English translations, and none of them identified grammar mistakes. Yoon revealed that all three types of text translations produced by ChatGPT showed clear grammatical errors. Moreover, the investigation results from the linguistics textbook had the highest translation accuracy, averaging 4.375 out of 5. Proverb accuracy was lowest at 2.75 and literary work accuracy was second at 3.25. Likewise, when the literal translation failed to capture the sense of a Korean proverb, ChatGPT, contrary to previous machine translation engines, offered further explanations of the proverbs' metaphorical connotation.

Yim [16] examined ChatGPT, Google Translate, and human translation when translating Korean sentences without subjects into English. Following the translation of 343 phrases taken from business reports and an analysis of subject substitution patterns, the most prevalent tendency across all corpora was the substitution of personal pronouns for subjects. While "we" was preferred and used more often by ChatGPT and machine translators, proper and common nouns were used more often by human translation. Short passive structures were also frequently used excessively by machine translators, while this was less common in ChatGPT translation.

Research on utilizing ChatGPT to translate Korean literature into English has evaluated whether irony in works of literature might be translated and evaluated the accuracy of translations from various texts, including proverbs, poetry, and English textbooks. Studies include the evaluation of ChatGPT's capability for Korean to English translations of irony in Korean literature as well as the comparison of its translation quality of Korean proverbs, poetry, and a textbook. Likewise, by translating Korean business documentation with omitted subjects into English, an investigation was carried out to examine the distinctions between Google Translate, ChatGPT, and human translation. The research focused on Korean conversational language, which has not been examined in previous studies, to compare ChatGPT with Naver Papago. Papago has been documented for its exceptional performance in Korean text-based translation. The research used the Korean language, including an abundance of the Gyeongsang dialect from the TV series *Pachinko* to carry out a comprehensive examination of the translation characteristics of the two AI translators.

3. Methods

3.1 Data Corpus

This study utilized a set of 590 Korean dialogues from Season 1, Episodes 1 and 2 of the American drama series *Pachinko*, which streamed on Apple TV+. The drama portrays the life of a Korean woman, Kim Seon-ja, and her family that included four generations living in Yeongdo, Busan in Korea, as well as Osaka and

Yokohama, Japan, and the United States [17]. The characters in this drama primarily used Korean, with Japanese and English as supplementary languages. The dominant setting is Yeongdo, Busan, where the main character, Seon-ja, grew up, using the Gyeongsang dialect which is characteristic of the region. The characteristics of the Gyeongsang dialect include unique suffixes. Out of the total 590 sentences, 384 sentences (65.1%) use the Gyeongsang dialect, and the remaining 206 sentences (34.9%) are composed of elements such as Standard Korean Language. These figures suggest that the Gyeongsang dialect accounts for a high proportion of the drama's linguistic composition.

3.2 Data Procedure

In this research, a total of 590 dialogues from the TV drama Pachinko were classified into 36 segments based on scene division, temporal flow, and thematic relevance. The dialogues were translated from Korean to English using ChatGPT and Papago versions that were publicly available between January and February 2024. For the purposes of comparison, the same dialogues were translated using both ChatGPT (N=590) and Papago (N=590). Korean-to-English translation was performed using both ChatGPT (N=590) and Papago (N=590). This delimiter was employed as current translation software, despite advances in capability, struggles with the figurative context inherent in languages. In utilizing ChatGPT, the instruction "Translate the following Korean text into English" followed by an element enclosed in quotation marks was the prompting template. This instruction was the only criterion that separated ChatGPT from Papago regarding human intervention. As a cross reference for validity and reliability of ChatGPT and Papago, 50 percent of the ChatGPT (N=295) and Papago (N=295) samples (ChatGPT: 590 and Papago: 590) were compared with translations which were made by a language specialist with native proficiency in both Korean and English. Examination of inter-rater reliability using Kappa coefficient demonstrated significance at the $k = .738$ ($p=.000$) level, reflecting a high level of agreement in inter-rater reliability concerning the precision of segment assessments, consistent with Landis and Koch's criteria [18].

3.3 Data Analysis

In this investigation, the erroneous frequency rates (EFR) of translations generated by ChatGPT and Papago and the reliability and validity of the EFR judgment were assessed employing IBM SPSS Statistics 23. The translation error criteria in this study were developed based on the standards utilized in previous research [19-23], with a specific focus on the Korean conversational language used in Pachinko dialogues, to assess the kinds of errors in the translations made by ChatGPT and Papago. Four types of errors and 12 subcategories were identified as shown in Table 1. The first error category included errors of incorrect meaning, homonym errors, title recognition errors, and cultural misinterpretations. The second category grouped errors in grammar such as word order, subject-verb agreement, and pronoun reference. A third error category defined as omission errors included translations that excluded content relevant to the original Korean meaning. These errors were further subcategorized into omitted words and omitted phrases or clauses. The fourth category of errors was mechanical translation errors, including proper noun and spelling errors.

Table 1. Classification of four categories with eleven subtypes

Category	Subtype
Lexical error	Incorrect word meaning
	Homonym errors
	Honorific title errors
	Cultural differences between Korean and English
	Determiner
Grammatical error	Tense
	Pronoun
	Word omission
Omission error	Phrase or clause omission
Mechanical error	Proper
	Spelling
Total	11 Subtype errors

4. Results

4.1 Results on Accuracy of ChatGPT and Papago Translations

This part illustrates the results of the accuracy of Korean-to-English translations of dialogue from the drama *Pachinko* by ChatGPT and Papago in Table 2. A total of 590 translated sentences were examined for ChatGPT translation accuracy, 343 (58.1%) were correctly translated, while 247 (41.9%) were inaccurate. Out of the 590 sentences translated using Papago, 223 (37.8%) were found to be accurate translations, and 367 (62.2%) were found to have errors. Given these results, ChatGPT tended to perform with greater accuracy, however, both tools included a sizable number of errors when translating phrases.

Table 2. Evaluation of translation accuracy

Translation Output	ChatGPT	Papago
Correct Translation	343 (58.1%)	223 (37.8%)
Incorrect Translation	247 (41.9%)	367 (62.2%)
Total	590 (100%)	590 (100%)

4.2 Frequency Counts for Four Error Types

This portion summarizes the lexical errors, grammatical errors, omission errors, and mechanical errors that occurred in the translations by ChatGPT and Papago in Table 3. Regarding ChatGPT's performance, vocabulary errors accounted for the largest percentage at 190 (76.9%), followed by grammatical errors at 31 (12.6%), 18 (7.3%) omission errors and 8 (3.2%) mechanical errors. In terms of Papago translation errors, there were 301 (82.0%) lexical errors, 48 (13.1%) grammatical errors, 10 (2.7%) mechanical errors and 8(2.2%) omission errors. Both translation tools include lexical errors as the primary issue; however, the frequency of omission errors is comparatively lower in Papago than in ChatGPT. This indicates that further difficulties exist with lexical interpretation in the Papago translation logic.

Table 3. Frequency of four error categories

Category	ChatGPT	Papago
Lexical Error	190 (76.9%)	301 (82.0%)
Grammatical Error	31 (12.6%)	48 (13.1%)
Omission Error	18 (7.3%)	8 (2.2%)
Mechanical Error	8 (3.2%)	10 (2.7%)
Total	247 (100.0%)	367 (100.0%)

4.3 Frequency Counts for Subtype Errors with Examples

4.3.1 Frequency of 11 Subtype Errors

This section reports on the frequencies of 11 subtype errors featured over the four error categories in Table 4. When looking at the results of the analysis of the 11 subtype errors for ChatGPT and Papago, incorrect word meaning errors were the most frequent at 148 (59.9%) instances for ChatGPT, followed by pronoun errors accounting for 23 (9.3%) instances. Additionally, 21 (8.5%) honorific title errors, 15 (6.1%) mistranslations of homonyms, 12 (4.9%) phrase or clause omission errors, 7 (2.8%) errors due to cultural differences between Korean and English, 7 (2.8%) proper noun errors, and 6 (2.4%) determiner errors were made. Finally, 4 (1.6%) tense errors, 3 (1.2%) word omissions, and 1 (0.4%) spelling error were identified.

Table 4. Frequency of 11 subtypes of errors

Category	Subtype	ChatGPT	Papago
Lexical error	Incorrect word meaning	148 (59.9%)	266 (72.5%)
	Homonym	15 (6.1%)	14 (3.8%)
	Honorific title	21 (8.5%)	15 (4.1%)
	Cultural differences between Korean and English	7(2.8%)	5 (1.4%)
	Determiner	6(2.4%)	4 (1.1%)

Grammatical error	Tense	4(1.6%)	5 (1.4%)
	Pronoun	23(9.3%)	40 (10.9%)
Omission error	Word omission	3(1.2%)	2(0.5%)
	Phrase or clause omission	12(4.9%)	5 (1.4%)
Mechanical error	Proper noun	7 (2.8%)	10 (2.7%)
	Spelling	1 (0.4%)	1 (0.3%)
Total	11 Subtype errors	247 (100.0%)	367 (100.0%)

Regarding Papago, incorrect word errors were the most frequent, accounting for 266 (72.5%) instances (72.5%) followed by 40 (10.9%) pronoun errors. There were 15 (4.1%) errors in honorifics, 14(3.8%) mistranslations of homonyms, and 10 (2.7%) errors in recognizing proper nouns. In addition, there were, 5 (1.4%) errors due to cultural differences between Korean and English, 5 (1.4%) tense and 5 (1.4%) phrase or clause omissions. 4 (1.1%) determiner errors, and 2 (0.5%) word omission errors and 1 (0.3%) spelling error occurred.

4.3.2 Frequency of 11 Subtype Errors

This section regroups the 11 subtypes into four broad categories including lexical errors, grammatical errors, omission errors, and mechanical errors with examples of these errors. For six of the 11 subtypes, information on subtype errors is presented under four headings: original Korean text, ChatGPT translation, Papago translation, and manual translation. For the five remaining subtypes, information was provided in a different format because errors shared for the same Korean text were not available for both AI translators.

4.3.2.1 Examples of Four Subtypes Falling in the Lexical Error Category

This section provides examples of the four subcategories of lexical errors in Table 5. The first subtype of lexical errors is related to incorrect word meaning errors generated by ChatGPT and Papago. ChatGPT translated “나의 어머니는 박복한 분이셨습니다.” into “My mother was virtuous person.” while Papago translated this Korean text into “My mother was rich man.” In this Korean text, the adjective “박복한” is created using characters whose meaning is rooted in Chinese. The “박” should have been interpreted as “little” or “not” in this sentence. while the word “복한” means “fortune,” or “lucky.” Therefore, in this segment, it is appropriate to translate “박복한” as “unfortunate” or “unlucky” rather than “virtuous” by ChatGPT and “rich” by Papago.” Moreover, Papago mistranslated the “분” as “a rich man.” that is not appropriate because My mom is a woman. This shows that Papago does not accurately recognize the gender of the subject. This finding suggests the requirement for translations that encompass the etymology, gender, contextual meaning, and idiomatic expressions of Chinese-character words to reflect the rich meaning of various Korean expressions when using AI translators.

Table 5. Examples of lexical errors

Subtype	Criteria	Examples
Incorrect word meaning errors	Korean text	나의 어머니는 박복한 분이셨습니다.
	ChatGPT	My mother was virtuous person.
	Papago	My mother was rich man.
	Manual translation	My mother was a rich man.
Homonym errors	Korean text	선자야, 와?
	ChatGPT	Sunja, come here.
	Papago	Seonja, are you coming?
	Manual translation	Sunja, what is it?
Honorific title errors	Korean text	성님, 그래두 이걸 아입니더.
	ChatGPT	Seongnim, still, this isn't right.
	Papago	But, Holy Father, this is not it.
	Manual translation	Sister, still, this isn't right.
Cultural differences between Korean and English	Korean text	에노야 야노야
	ChatGPT	Hey, you, let's go.
	Papago	Hey, hey, hey, hey, hey, hey, hey, hey, hey hey!
	Manual translation	Enoya yanoya

With respect to the homonym errors, ChatGPT translated the Korean text “선자야, 와?” to “Sunja, come here” and Papago translated it into “Seonja, are you coming?” ChatGPT translated the Korean text “선자야, 와?” to “Sunja, come here” and Papago translated into “Seonja, are you coming?” In this segment, “와?” is a shorthand for “왜 그러니?” in Gyeongsang-do dialect. “Neither ChatGPT nor Papago detected the regional meaning of this shorthand. Both machine translators translated “와” into meaning “come.”

Regarding title errors, ChatGPT translated “성님, 그래도 이걸 아녘니더” as “Seongnim.” Papago incorrectly translated the sentence to refer to “Holy Father,” however, “the correct translation is “Sister, but this is not right.” For the segment, Chatgpt recognized the title “성님” as a proper noun and translated it into Korean as it sounds. Papago recognized “성님” as “Holy Father.” In Korean, “성님” is an expression used in the Gyeongsang-do region, and is a term used to refer to a family member or close friend of the same gender who is older than the speaker. This shows that ChatGPT and Papago did not accurately reflect this Korean honorific title when translating it into English.

The following portion of the paper presents examples of translation errors that are rooted in cultural differences between Korean and English. ChatGPT translated the expression “에노야 야노야” as “Hey, you, let’ go,” while Papago translated it as “Hey, hey, hey, hey, hey, hey, hey, hey, hey!” However, both AI based translators did not provide accurate translation. “에노야 야노야” is a part of a Korean boat song, that must be translated phonetically to be accurate. Therefore, the correct translation is “Enoya yanoya.” Both ChatGPT and Papago did not accurately recognize the refrain of the Korean folk song and instead used a common refrain from some English-speaking cultures.

4.3.2.2 Examples of Three Subtypes Falling in the Grammatical Error Category

This section highlights examples of grammatical errors in Table 6. Regarding the determiner errors, ChatGPT translated “니 그 피아노가 얼마짜린 줄아나?” as “Do you know how expensive your piano is?” For this segment, “그 피아노는” is referencing “that piano” not “your piano.” Papago translated “지 배타러 가야 되는데” as “I need to go catch a boat.” In this segment “배” is not “a boat” but the boat. Here, the speaker is going to a specific boat, so “the” is correct in the context. Regarding tense errors, ChatGPT translated “할매한테 맨날 들었는데요” as “I’ve heard it from Grandma all the time.”

Regarding tense errors, ChatGPT translated “할매 한테 맨날 들었는데요” as “I’ve heard it from Grandma all the time.” However, in the context for the segment, “들었는데요” is not “I’ve heard it,” but rather “I heard it” This reflects the tendency of ChatGPT not to recognize the appropriate tense. Papago translated “그런데도 신고를 안 해?” to “But still don’t report it” For the segment, “안해” should be “didn’t” not “don’t” “But you still didn’t it?” This tense error occurred because the Papago translation did not consider the overall flow of the dialogue in this drama.

The following are examples of pronoun errors. ChatGPT translated “선자야, 내 니한테 약속한 거 아나?” to “Sunja, do you remember what you promised me?” In the Korean text, “내” refers to me and “니” means to you. This shows that ChatGPT does not recognize Korean personal pronouns correctly. Papago translated “그분도 고생께나 하실텐네” into “He must have had a hard time, too.” However, this is an incorrect translation in the context. Here, “He” refers to a pronoun for the character Senja in the drama. Therefore, the correct translation would be “She must have had a hard time, too.” rather than Papago’s “He must have had a hard time, too.”

Table 6. Examples of grammatical errors

Subtype	Criteria	Examples
Determiner errors	Korean text	니 그 피아노가 얼마짜린 줄 아나
	ChatGPT	Do you know how expensive your piano is?
	Manual translation	Do you know how much this piano is?
	Korean text	지 배타러 가야 되는데
Tense errors	Papago	I need to go catch a boat.
	Manual translation	I need to go catch the boat.
	Korean text	할매한테 맨날 들었는데요
	ChatGPT	I’ve heard it from Grandma all the time.
	Manual translation	I heard it from Grandma all the time.
	Korean text	그런데도 신고를 안해
	Papago	But you still don’t report it.

Pronoun errors	Manual translation	But, you still didn't report it?
	Korean text	선자야, 내 니한테 약속한 거 아냐?
	ChatGPT	Sunja, do you remember what you promised me?
	Manual translation	Sunja, do you remember what I promised you?
	Korean text	그분도 고생께나 하셨겠네.
	Papago	He must have had a hard time, too.
	Manual translation	She must have had a hard time, too.

4.3.2.3 Examples of Four Subtypes Falling in the Omission Error Category

This portion demonstrates examples of word and phrase omission that occur in ChatGPT and Papago translation in Table 7. In terms of word omission, ChatGPT translated, “미국서 들어온 달디단 오렌지” as “Oranges imported from the United States,” and in this interpretation, the adjective “달디단” was omitted. The accurate translation would be “Sweet oranges imported from the United States.” In another instance, a Papago translation, “거가 훨씬 안전하데이” was translated as “It’s much safer,” and in this transaction, “거가” was omitted. A more accurate translation would be “It’s much safer there.” These analyses provide representative examples of a word omission phenomenon that occurs in machine translation and point out the possibility and identify a trend showing that semantically important elements may be omitted.

Table 7. Examples of omission errors

Subtype	Criteria	ChatGPT
Word omission errors	Korean text	미국서 들어온 달디단 오렌지
	ChatGPT	Oranges imported from the United States
	Manual translation	Sweet oranges imported from the United States
	Korean text	거가 훨씬 안전하데이
	Papago	It's much safer.
	Manual translation	It's much safer there.
Phrase omission errors	Korean text	동포애도 없으요?
	ChatGPT	Don't you have any fellow countrymen?
	Papago	Don't you have brotherhood?
	Manual translation	Don't you have any love for your fellow countrymen?

The following are examples of phrase omission made by ChatGPT and Papago. ChatGPT translated “동포애도 없으요?” as “Don’t you have any fellow countrymen?” while Papago translated it as “Don’t you have brotherhood?” Both of these translations omit the word “애.” In the case of ChatGPT, translating “any fellow countrymen” as “any love for your fellow countrymen” can be said to faithfully reflect the original text. Similarly, in the translation of Papago, translating “Don’t you have brotherhood?” as “any sense of brotherhood for your fellow countrymen” reflects the meaning of the original Korean text. This suggests that AI translators need to analyze and reflect the context more precisely so as not to omit semantically important elements.

4.3.2.4 Examples of Four Subtypes Falling in the Lexical Error Category

This portion provides samples of mechanical errors for the subcategories in Table 8. Regarding proper noun errors, ChatGPT translated “불란서에서 온 최고급 사탕” into “finest candies from Holland” and Papago translated “The finest candy from the Bulanse.” These two translations are incorrect. In the segment, “불란서” refers to France. ChatGPT and Papago’s translations were incorrectly translated as “Holland” and “Bulanse” respectively, which is a result of both translators failing to recognize the historical context of the term “불란서.” Such errors can occur when a translator fails to properly apply historical context.

In terms of spelling errors, ChatGPT incorrectly transcribed the Korean boat song refrain “에노야 야노야 어기어차” as “Heyano, Yanono, Egiocha.” The correct transcription should be “Eyano, Yano, uhgi-ee uhcha!” This case shows that ChatGPT has difficulty accurately transcribing Korean rhythmic fillers and chants into English. Such errors can occur because certain Korean syllables and pronunciations do not directly correspond to the English phonetic system. Papago translated “마, 뭐 말이 글노?” to “E, what do you mean?” For this segment, “E” would have likely been spelled in English as “Huh” depending on how it is spoken. “What do you mean?” reflecting its inability to spell the common interjection “마.” This case may be indicative of a weakness in Papago for interpreting interjections such as the one shown here.

Table 8. Examples of mechanical errors

Subtype	Criteria	Examples
Proper noun errors	Korean text	불란서에서 온 최고급 사탕
	ChatGPT	The finest candies from Holland
	Papago	The finest candy from the Bulanse
	Manual translation	The finest candy from France
Spelling errors	Korean text	에노야 야노야 어기어차
	ChatGPT	Heyano, Yanono, Egiocha
	Manual translation	Eyano, Yano, uhgi-ee uhcha
	Korean text	마, 뭘 말이 글노?
	Papago	E, what do you mean?
	Manual translation	Huh, what do you mean?

5. Discussion

5.1 Comparison of ChatGPT and Papago Accuracy with Recommendations for Improvement

Based on results for Korean-English translation of the drama *Pachinko* dialogue using ChatGPT and Papago, out of a total of 590 translated sentences, ChatGPT accurately translated 343 instances, while Papago accurately translated 223 instances for the same dialogue. This was by far the most common type of error produced by both translators. Papago recorded 120 more of this type of translation error than ChatGPT. These findings align with Lee's study, which examined the translation of Korean short stories and reported that ChatGPT had slightly fewer errors than Papago [24].

This outcome implies that ChatGPT generated higher translation accuracy than Papago, but it also illustrates that neither translator provides perfect translations. The higher translation accuracy of ChatGPT is likely due to its large-scale data learning [25] and more sophisticated context awareness. Alternatively, Papago, despite operating a translation model specialized for Korean, may have relatively low translation accuracy because it takes a more literal approach or has limitations in handling complex contexts [21]. To strengthen these two AI translators, more sophisticated natural language processing algorithms need to be developed, and machine learning models need to better interpret the context with deeper insight.

5.2 Assessment and Significance of the Four Error Categories

ChatGPT also produced 190 lexical errors compared to 301 for Papago, and 31 grammatical errors compared to 48 for Papago. ChatGPT had 18 instances of omission errors compared to Papago with 8 instances. Both AI translators had relatively few mechanical errors, with ChatGPT having slightly fewer, with 8 instances compared to 10 instances by Papago. This finding suggests that ChatGPT and Papago had the most problems with lexical interpretations over the four categories. This trend aligns with Lee and Cha's Papago Korean-English translation study. They reported that for colloquial sentences translated using Papago, errors occurred about 44% of the time with semantic errors being the most frequent [21]. In terms of the grammatical errors, Papago had a higher error rate than ChatGPT, and thus needed to produce more grammatically accurate translations. However, ChatGPT had more omission errors, which makes it more likely that the information in the original text will not be fully conveyed, so improvements are needed to enhance content preservation. Mechanical errors occurred twice as often with Papago, but there were no significant differences between ChatGPT and Papago. These findings suggest that both translators should improve their AI learning models to analyze the context more precisely. In addition, a facility that allows users to provide spontaneous feedback on errors would improve the quality of translations.

5.3 Exploration of 11 Error Subtypes with Examples

The analysis of the translation errors by ChatGPT and Papago was categorized into four types of errors including 11 subtypes. The first subtype was labeled as incorrect word meaning, including single and multiple word errors, and errors in translating common nouns phonetically, which frequently occurred. For ChatGPT, 190 out of 247 errors were of this type, while Papago produced 301 of this type of error out of its 367 total translation errors. This was by far the most common type of error produced by both translators. Papago recorded 79 more of this type of translation error than ChatGPT. This finding reflects that Papago may have greater

limitations in its ability to accurately reflect the linguistic context of the vocabulary. More generally, these results show that both ChatGPT and Papago need a richer depth of logic into the variety of forms that Korean vocabulary takes in order to improve the accuracy of translations.

Homonym errors were a second sub-type occurring when words with multiple meanings were not accurately translated. ChatGPT produced 15 homonym errors, and Papago showed 14 homonym errors. The Korean language has many homonyms that ChatGPT and Papago could not translate. Examples of this subtype of error patterns included “와,” and “말,” “서.” The two translators also mistranslated the dialect “고마” in “고마가자” to mean “Let’s go.” ChatGPT recorded it as “thank you” while Papago translated it as “Let’s go thank you” showing that the algorithms wrestled with interpreting various Korean homonyms.

With the honorific titles, ChatGPT generated 19 errors and Papago made 15 errors. In this subtype of errors, ChatGPT made somewhat more mistranslations than Papago. Koreans tend to use family titles not only for family members but also, for non-family members and colleagues. The actors use “형님,” “성님,” and “할머니” to express intimacy when they meet someone who is not familiar with them. This system of titles reflects the use of titles that do not have a direct equivalent in English. In Western cultures, friendly gestures would take using first rather than a title with the last name. This difference in the use of honorifics caused translation errors. Therefore, to interpret Korea’s use of honorific titles more accurately, translation strategies need to incorporate greater sensitivity to the cultural and social meaning of honorifics is necessary. This addition of this knowledge to the AI would significantly enhance the quality of machine Korean-English translations.

In terms of the cultural differences between Korean and English, ChatGPT produced 6 errors while Papago generated 4 translation errors in this subtype. Regarding the errors caused by cultural differences within the lexical error subtype, ChatGPT produced four errors and Papago produced five errors. The errors in this subtype occurred in the refrain of a traditional folk song entitled “배노래 (Boat Song)” for both ChatGPT and Papago. Since it is a refrain of a Korean song, it is appropriate to translate it as it sounds while maintaining the phonetic characteristics. For example, the phrase “에노야 야노야” was translated as “Hey, you, let’s go” by ChatGPT and as “Hey, hey, hey, hey, hey, hey, hey, hey, hey, hey!” by Papago. Neither of the AI translators produced a precise translation. It is desirable to translate it as “Enoya Yanoya!” (A part of a Korean boat song, equivalent to “Yo ho, yo ho” in English) to convey the cultural meaning together.

When it comes to determiners, ChatGPT produced 6 errors while Papago generated 4 errors related to the use of articles, possessive adjectives, and demonstrative adjectives. These errors indicate that both translators failed to sufficiently reflect the meaning of determiners referring to specific Korean nouns and suggest that the accuracy of determiner use based on contextual interpretation needs to be improved. These results suggest that there are differences in the way AI translators handle determiners, and in particular, they show that they sometimes do not accurately reflect the rules for determiner usage between Korean and English. Unlike English, Korean has a linguistic characteristic that does not explicitly use determiners such as articles. In English, articles such as “a,” “an,” and “the” are important factors in determining the specificity and consistency of nouns, but in Korean, these meanings are often conveyed through context. Considering these differences, it is suggested that the translation of determiners by AI translators needs to be more precisely adjusted according to the context and that learning that reflects the structural differences between Korean and English should be strengthened.

In terms of tense errors, ChatGPT made 4 errors, while Papago made 5 errors. This suggests that both translators make errors in the use of tenses. English has more tenses than Korean, which may make it more likely that difficulties will arise in selecting the appropriate tense when translating from Korean to English. In particular, Korean often infers tenses based on context, so errors may occur if the translation model cannot accurately reflect this. The fact that the error frequency between the two tools does not differ significantly shows that problems related to tense handling are common for these two AI translators. This supports the suggestion that more sophisticated context analysis and tense mapping techniques are needed to reduce tense errors.

Regarding pronoun errors, ChatGPT generated 23 errors while Papago made 40 translation errors. This subtype of error is the second most common error among the 11 sub-errors in both ChatGPT and Papago. This subtype of pronoun errors is caused by the frequent omission of subjects and objects in Korean conversational speech, and also because the two AI translators were based on the Gyeongsang dialect, and thus failed to accurately recognize pronouns such as “나(ni)” and “나(nae).” This suggests that AI translators need to understand the overall flow of conversational sentences to translate appropriate pronouns, and that pronoun recognition in various Korean dialects is important.

With respect to single-word omissions, ChatGPT produced 2 single-word mistranslations and Papago made 5 errors. This suggests that after translation, it is necessary to check whether the meaning is properly

conveyed by comparing it with the original text, with special attention given to whether words such as adjectives and adverbs are not omitted. To reduce this subtype of errors caused by the omission of subjects and objects in spoken Korean, context-based translation technology and large-scale Korean conversation data should be refined to recognize frequently occurring omission patterns so that algorithms can translate appropriate subjects or objects.

In terms of phrase omissions, ChatGPT made 12 errors whereas Papago produced 5 errors. This outcome indicates that ChatGPT tends to omit more information than Papago. That is, ChatGPT frequently abbreviates or simplifies certain elements of the source text during the translation process, which increases the possibility of missing some nuances or details contained in the source text. This shows that ChatGPT tends to prioritize sentence fluency and readability over literal translation, while Papago is likely to aim for a more intuitive and faithful translation of the source text. The results of this study suggest that it is necessary to expand the parallel corpus and design the data in a way that maintains the sentence structure and length of the source text to minimize information omission during the translation process.

In reaction to proper noun errors, ChatGPT produced 7 errors while Papago generated 10 errors. ChatGPT and Papago made mistakes in translating people's names and country's names. For example, ChatGPT translated the proper noun "선자," the name of a female protagonist in *Pachinko*, as "the old man," while Papago incorrectly added "It's" to the translation. For the country name, ChatGPT translated "볼란서" as "Holland," while Papago translated it as "Bulanse" which is an old Korean expression for "France." When AI translators do not recognize the meaning of people's names or place names, translations are liable to be distorted. Therefore, the AI translators need to be improved so that they can more accurately reflect proper nouns.

In terms of spelling errors, ChatGPT and Papago produced a single instance each. This error was the least common among the 11 subcategories. This reflects that AI translators produce word-for-word conversions, but struggle when dealing with more sophisticated tasks involving word meaning and cultural context. Therefore, in addition to spelling correction functions, it is necessary to have more sophisticated context-based meaning interpretation and reflection of cultural differences to improve the performance of the AI translator. This suggests that Papago has limitations in accurately interpreting expressions such as Korean exclamations. ChatGPT needs to strengthen the appropriate correspondence between Korean syllables and English pronunciations to accurately transcribe Korean rhythmic exclamations. Learning diverse Korean spoken data contained in traditional songs or dialects can improve accuracy. Papago needs to learn data that reflects exclamations and colloquial expressions used in conversational speech to more accurately interpret expressions such as Korean exclamations.

6. Conclusions

Given the rapid development of the OTT industry as well as the growth in popularity of Korean content around the world, the demand for English subtitles for various Korean content is increasing significantly. For this study, a popular drama was selected for study. Korean-to-English translations of dialogues using ChatGPT and Papago dialogues were evaluated for accuracy of the AI translations. Specifically, the frequency of four major types of translation errors was compared, and 11 subtypes of the errors were analyzed to assess the precision of the AI translators. This study is significant in that it goes beyond a simple comparison of translation performance and sheds light on the limitations of AI translation quality.

Results served as a basis for future directions to improve these tools by closely analyzing the characteristics and occurrence patterns of each type of translation error. Results of the study showed ChatGPT had a higher overall translation accuracy except for grammatical errors with both translators having the most frequent occurrence as lexical errors. As the dialogue of *Pachinko* is mainly composed of the Gyeongsang dialect, there is a possibility that different results might be found if the target of translation was Korean Standard Language. However, Korean has a variety of dialects. In addition to the Gyeongsang dialect, various regional dialects such as Jeolla-do, Chungcheong-do, Jeju-do, etc. are utilized throughout Korea.

Therefore, AI translators need to be strengthened in their learning of these diverse Korean dialects and need to improve their algorithms to reflect linguistic and cultural contexts more precisely and naturally. With more improved technological advancements, the quality of AI-based translations of Korean content supporting the provision of high-quality English subtitles will be possible in the future. In addition, AI-based translation technology has a wide range of potential applications beyond simple subtitle creation such as real-time translation in a variety of settings. This will contribute to expanding the accessibility of Korean content to global markets and increase the opportunities for more viewers to experience Korean culture and language.

In this study, the performance of AI translators analyzing dialogue from the drama was the focus. In the future, research might be expanded to evaluate the quality of AI translations for other types of Korean content such as news articles, literary works, academic papers, spoken language, and other forms. Through this process, it is expected that results will more clearly identify the types and patterns of translation errors. Then, it will also be possible to clarify the differences in translation and make them better capture the nuanced characteristics of the text. It is hoped that this research has significantly contributed to the improvement of AI translation models.

Acknowledgments: This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF- R-2023-01065)

Conflicts of Interest: The author declares no conflict of interest.

References

- [1] J. H. Lee and K. W. Cha, "An analysis of the errors in the auto-generated captions of university commencement speeches," *Journal of Asis TEFL*, vol. 17, no. 1, pp. 143-159, spring 2020, doi: <http://dx.doi.org/10.18823/asiatfl.2020.17.1.9.143>.
- [2] M. H. Kim, "A study on the characteristics of spoken and written Korean," *Korean Language Research*, no. 15, pp. 23-74, Dec. 2004.
- [3] J. H. Lee and K. W. Cha, "Human interpretation and Machine translations based upon interviews with director Joon-ho Bong," *Korean Journal of English Language and Linguistics* vol. 23, pp. 204-210, Jan. 2023, doi: <https://doi.org/10.15738/kjell.23..202303.204>.
- [4] NAVER Cloud, How to Use! Implementing your own translation service using Papago NMT API, Accessed: Feb. 10, 2025. [Online] Available: https://m.blog.naver.com/n_cloudplatform/222272171633
- [5] I. Kostka and R. Toncelli, "Exploring applications of ChatGPT to English language teaching: opportunities, challenges, and recommendation," *The Electronic Journal for English as a Second Language*, vol. 27, no. 3, Nov. 2023, doi: <https://doi.org/10.55593/ej.27107int>.
- [6] T. N. Fitria, "Artificial intelligence (AI) technology in OpenAI ChatGPT application: a review of ChatGPT in writing English essay," *Journal of English Language Teaching*, vol. 12, no. 1, pp. 44-58, Mar. 2023, doi: <https://doi.org/10.15294/elt.v12i1.64069>.
- [7] NAVER Knowledge Encyclopedia, Language of Busan, The Digital Local Culture Encyclopedia of Korea, Accessed: Jan. 20, 2025. [Online] Available: <https://terms.naver.com/entry.naver?docId=2822070&cid=55772&categoryId=55807>
- [8] P. Rao, L. M. McGee, and C. A. Seidman, "A comparative assessment of ChatGPT vs. Google Translate for the translation of medical documents," *Journal of Medical Artificial Intelligence*, vol. 4, no. 1, pp. 1-8, Jun. 2024, doi: <https://doi.org/10.21037/jmai-24-24>.
- [9] R. Aeni, B. Baharuddin, L. J. Putera, and B. Z. Melani, "The accuracy of ChatGPT in translating linguistics text in scientific journals," *Didaktik: Jurnal Ilmiah PGSD STKIP Subang*, vol. 10, no. 1, pp. 59-68, 2024. [Online] Available: http://eprints.unram.ac.id/44286/2/Jurnal_Rusmita%20Aeni_E1D019184.pdf
- [10] W. Gu, "Linguistically informed ChatGPT prompts to enhance Japanese-Chinese machine translation: A case study on attributive clauses," *arXiv preprint*, vol. 2303.15587, Mar. 2023, doi: <https://doi.org/10.48550/arXiv.2303.15587>.
- [11] R. Gao, Y. Lin, N. Zhao, and Z. G. Cai, "Machin translation of Chinese classical poetry: a comparison among ChatGPT, Google Translate, and DeepL translator," *Humanities and Social Sciences Communications*, vol. 11, no. 1, pp. 1-10, Jun. 2024, doi: <https://doi.org/10.1057/s41599-024-03363-0>.
- [12] F. Khoshafah, "ChatGPT for Arabic-English translation: Evaluating the accuracy," *Research Square*, Apr. 2023, doi: <https://doi.org/10.21203/rs.3.rs-2814154/v2>.
- [13] S. J. Park and E. S. Choi, "A study on the applicability of ChatGPT in translating irony," *The Journal of Translation Studies*, vol. 24, no. 2, pp. 131-160, Jun. 2023, doi: <https://doi.org/10.15749/jts.2023.24.2.005>.
- [14] Y. B. Yoon, Analysis of ChatGPT's Accuracy on Korean-English Translation, *The Journal of Elementary, Education*, vol. 34, no. 4, pp. 215-231, Dec. 2023, doi: <https://doi.org/10.20972/kjee.34.4.202312.215>.
- [15] S. B. Jeon, *Introduction to English Phonetics*, Eulyoo Publishing, Seoul, 2005.
- [16] J. Yim, "Strategies for translating zero-subject Korean sentences into English: A comparison among human translation, Google Translate, and ChatGPT," *Linguistics*, vol. 32, no. 3, pp. 1-22, Sep. 2024, doi: <https://doi.org/10.24303/lakdoi.2024.32.3.1>.
- [17] Y. H. Shin, "A study on MinJin Lee's Pachinko as a diaspora narrative," *Homo Migrants*, vol. 30, pp. 105-135, May. 2024, doi: <https://doi.org/10.32715/hm.2024.30.004>.
- [18] J. P. Landis and G.G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, pp. 159-174, Mar. 1977, doi: <https://doi.org/10.2307/2529310>.

- [19] D. Vilar, J. Xu, L. F. D'Haro, and H. Ney, Error Analysis of Statistical Machine Translation Output, Proceedings of the Fifth International Conference on Language Resources and Evaluation, LREC'06, pp. 697-702, 2006. [Online] Available: <http://www.lrec-conf.org/proceedings/lrec2006/pdf/413.pdf>
- [20] J. H. Lee and K. W. Cha, "An error analysis of English writings of Korean university students with the assistance of machine translators," Journal of Learner-Centered Curriculum and Instruction, vol. 20, no. 19, pp. 133-163, Oct. 2020, doi: <https://doi.org/10.22251/jlcci.2020.20.19.133>.
- [21] J. H. Lee and K. W. Cha, "A study of error types in Korean-English translation from Korean spoken language to Papago" Modern English Education vol. 23, no. 1, pp. 56-65, Feb. 2022, doi: <https://doi.org/10.18095/meeso.2022.23.1.56>.
- [22] B. H. Seo and S. Y. Kim, "An analysis of errors in machine translation," The Journal of Translation Studies, vol. 19, no. 1, pp. 99-117, Mar. 2018, doi: <http://doi.org/10.15749/jts.2018.19.1.004>.
- [23] J. Y. Choi and J. H. Jeon, "An analysis of error types in machine translation as a learning tool for Korean middle and high school students," Journal of Language Sciences, vol. 29, no. 1, pp. 179-200, Feb. 2022, doi: <https://doi.org/10.14384/kals.2022.29.1.179>.
- [24] C. S. Lee, "Differences in English literary translation among ChatGPT, Papago, and human translators," The Journal of Translation Studies, vol. 25, no. 2, pp. 11-37, Jun. 2024, doi: <https://doi.org/10.15749/jts.2024.25.2.001>.
- [25] Y. R. Cha, "Big data analysis of news articles related to ChatGPT, Journal of Practical Research in Advertising and Public Relations, vol. 17, no. 3, pp. 83-110, Aug. 2024, doi: <https://doi.org/10.21331/jprapr.2024.17.3.004>.



© 2025 by the authors. Copyrights of all published papers are owned by the IJOC. They also follow the Creative Commons Attribution License (<https://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.