

A Thorough Assessment of Google Translate's Central Kurdish English Translation: Accuracy, Linguistic Features, and Challenges

Salah Jameel Jabrael ¹, Rawand Sabah Ahmad ²

¹ English Language Department, Faculty of Arts, Soran University, Erbil, KRG – Iraq

² English Language Teaching Department, Faculty of Education, Tishk International University, Erbil, KRG – Iraq

ABSTRACT: Enhancing machine translation systems for Kurdish language, notably for Central Kurdish is significant through offering insights into specific difficulties and targeted suggestions for improvements. Hence, the study aims to thoroughly assess the accuracy of Google Translate in translation of Central Kurdish to English. The study focuses on linguistic features and the challenges confronted during the process of translation. The study also seeks to analyze the linguistic elements and identifying errors in grammatical, syntactic, semantic, stylistic nuances and the difficulties faced in translation. To achieve this, the study employs as a qualitative research design and textual analysis. Likewise, the study applies the (Costa et al., 2015) methodology to assess linguistic errors in the aforementioned linguistic features. In this sense, the research article utilizes eight randomly selected reports from both Rudaw and K24 satellite media sources from various types of texts. The results showed that Google Translate normally struggles with grammatical, semantic, stylistic and cultural complexities inherent in Kurdish language. Significant mistranslations, omissions, and syntactic wider field of machine translation, focusing on the necessities required for improving underrepresented languages.

Keywords: Assessment, Errors, Google Translate, Central Kurdish-English Translation, Linguistic Features

1. Introduction

The introduction includes the background of the study, research problem, research objectives, research questions and hypothesis of the study.

1.1. Background

Translation is the procedure of transmitting the text from the original language into the intended language while taking into account the meaning of both tongues (Fadlan & Ichsan, 2022). The translation process intends to ensure that the target language's principles are followed and that the translated content has the same relevance. In this sense, with the assistance of Google Translate (henceforth, GT), people are capable of translating various material throughout multiple languages, making it a vital instrument for global interaction. The value GT provides stems from its practicality and affordability for ongoing interaction, manifesting linguistic disparities around the world. In spite of being a widely recognized tongue mostly in the Kurdish geographical positions, a great number of Central Kurdish (Sorani Kurdish) speakers challenge with translating Central Kurdish into English because of highly complexity grammar, semantics stylistics, lexis, abundance of colloquial expressions, and contextually dependent values, etc. between the two varieties of communication. In this sense, precise rendering from Central Kurdish to English is necessary for successful interaction as well as for maintaining cultural heritage and enabling the Kurdish individuals to obtain and exchange knowledge within English environments. Therefore, it is particularly critical to evaluate technology for such tongues, including GT.

1.2. Research Problem

Determining the difficulties and constraints GT encounters when translating Central Kurdish to English is the main research issue. Perhaps, the complex linguistic features of Central Kurdish are frequently a challenge for GT, which causes a number of problems. Such challenges in the translation of Central Kurdish to English by virtue of GT involve mistranslations of popular phrases in grammar, syntax, semantics, stylistics, etc. that might not have obvious English counterparts and the deletion of crucial elements of language necessary to accurately represent the purpose of the source language.

Furthermore, the GT system could be unable to convey the minute social and historical factors present in Central Kurdish, which could result in translations that are either excessively exact or entirely wrong. The arrangement of words, word structure, the meaning of expressions, the style of language, the usage of affixes in Kurdish are some examples of linguistic variations between the two languages that make interpretation more challenging and frequently lead to renditions that can be hard to comprehend or undermine their initial significance. Resolving such problems is essential to raising GT's preciseness and dependability between the two languages.

1.3. Research Objectives

The main objectives of the current study come as follows:

1. To assess the accuracy of Google Translate's Central Kurdish-English translation.
2. To analyze the linguistic features involved in the translation process.
3. To identify common challenges and errors in the translation.
4. Research Questions and Hypotheses

This study has the following research questions:

1. How accurate is Google Translate in translating Central Kurdish to English?
2. What are the linguistic features most affected during translation?
3. What are the main challenges faced in the translation process?

Also, it is hypothesized that the translations from Central Kurdish to English provided by GT will not always be accurate, especially when it comes to maintaining syntax, morphology, grammar, semantics, lexis (word selection), and stylistics. It is also hypothesized that grammatical issues, meaning in context missing, and incorrectly understanding phrases unique to a culture will create difficulties in translating.

1.4. Significance of the Study

The field of machine translation including, especially about the central Kurdish-English language combination, could be greatly impacted by this study. The results can help to raise the level of precision and reliance of machine translation programs such as Google Translate by identifying particular issues and offering focused fixes for the Central Kurdish language. The research can shed light on the distinctive linguistic characteristics of Central Kurdish that influence translation performance, which is of great benefit for machine translation. The research can assist in elaborating on the complexity of Kurdish grammar, syntax, semantics, and unique cultural characteristics by identifying frequent linguistic errors and challenges while translating Sorani- Kurdish into by GT. This development is essential to improving the efficiency of automatic translators for transferring between both languages.

Likewise, via more precise renditions of the two varieties, perhaps, Kurdish speakers can communicate more effectively while communicating with native speakers of English, obtain knowledge faster online, and preserve their culture better.

Also, the investigation provides researchers with useful information that they can utilize to improve and enhance translator programs, thereby producing more reliable and socially conscious about instruments for rendition. The investigation can lay the groundwork for additional study on neural networks and automated translation for scholars who are engaged with Kurdish language technology, thereby addressing the separation between Kurdish and Central Kurdish. It might also serve as a basis for further studies regarding other tongues that are marginalized in the field of automatic translation. Such a strategy highlights the wider implications the study, emphasizing its value to different users and how it may contribute to further translate text using machines.

2. Literature review

The review of literature consists of three main sections including a synopsis of GT, the Central Kurdish language presenting difficulties when attempting to translate its linguistic elements into English, and previous studies.

2.1. An Overview of Machine Translation and Google Translate

Machine Translation

In an attempt to support human translators and interpreters', academics who study basic languages developed machine translation (henceforward, MT). Thus, MT is the process of translating written materials from a single spoken language to another language via utilizing computer (Hutchins, 1986). In this sense, the objective of (MT) is to generate meaningful translation in a specific setting, with the final product demanding only a small number of modifications to be deemed satisfactory by humans (Hutchins & Somers, 1992).

Of particular interest here is that a wide range of parameters influence both the overall standard of MT and the translation itself. Certain languages seem more appropriate for MT than others. In this context, (MT) could function poorly and is therefore not very helpful when translating heavily conjugated languages such as Greek and Polish (Lagoudaki, 2008). Conversely, the Nordic language family paired with English contributes to MT production quality dependence on language combination. However, German, grammatically differs significantly from English. On the contrary, French-Russian production, is of lower caliber. English-Spanish is an additional language combination that is regarded as more developed than other languages (Lotz & Van Rensburg, 2014). The evaluation of MT platforms is essential because the results demonstrate the system's consistency level and are applied to enhance the structure of the system (Dorr et al., 2011). On the other hand, the number of words and degree of idiomatic expressions in the input sentences hinder the MT system's ability to translate English into Kurdish (Taher, 2017). Additionally, the MT for Sorani Kurdish in terms of vocabulary and transferring policy constraints result in inaccurate translations, making it difficult to apply general principles across contexts (Kaka-Khan, 2018).

Google Translation

Google Translate (henceforth, GT) is an accessible “online service” provided by “Google Inc” which enables anyone swiftly translate words, sentences, texts including webpages (Lotz & Van Rensburg, 2014). Because GT is a freely accessible online tool and translates greater number of languages than the rest of its rivals, it constitutes one of the foremost popular services for translation. Google subsequently launched Google Translate, a cost-free mobile application (Groves & Mundt, 2015).

While eighty languages are supported at present by GT, not all of its information for each of them is readily accessible; for groups of languages where there are not plenty of data accessible through the devices where the translation system typically gathers translations. In this sense, for specific language groups, GT's accuracy in translating is accordingly less than for combinations of languages for which a substantial collection of patterns is currently being generated. “Afrikaans and English” is a far less common linguistic pair than “French and English”, both well-known international languages and popular language pairings. As a result, there would be a significantly greater number of “French-English” text combinations online than “Afrikaans-English” text groups, giving GT more material from which to gain knowledge. Hence, it is far more probable that the “French-English” result from GT will be superior to the “Afrikaans-English” result (Groves & Mundt, 2015).

2.2. Central Kurdish Language: Challenges in translating Linguistic features of Central Kurdish into English

Sorani and Central Kurmanji are other names for Central Kurdish, which is mostly spoken in the "Kurdistan Region of Iraq and parts of Iran." It includes a variety of accents, such as Mokri and Ardalani in Iran and Slaimani, Sorani, and Garmiani in the Kurdistan Region of Iraq (Jabrael, 2019). In addition, (Hasanpour, 1992) highlights that Central Kurdish is predominantly communicated “in the North-Eastern parts of Iraq and Western parts of Iran”. In spite of being expressed in the “Latin alphabet” for the Kurmanji dialect, Central Kurdish (Sorani Kurdish) is represented using the Arabic alphabet (Muhealddin et al., 2024). In terms of both population and geographic distribution, Central Kurdish is the second most widely spoken dialect among Kurds, after Kurmanji (Northern Kurdish) and before Southern Kurdish dialects. Besides, the utilization of diverse informal expressions by Central Kurdish speakers in disparate geographical locations is the major contribution of the literature study on the aforementioned accents. This is due to the fact that it is challenging for GT to render all informal expressions within such accents and subdialects in Central Kurdish into English.

Likewise, due to dissimilarities of linguistic features such as Grammar (including morphology, syntax), semantics, omission of linguistic expressions, and stylistics, etc., between Sorani and English, translating Central Kurdish into English with GT provide additional difficulties. The linguistic differences between Sorani Kurdish and English make it difficult to translate Sorani Kurdish into English using GT morphologically. In this regard, agglutinative construction of Central Kurdish possesses a difficult word construction (Taylan, 2020). It virtually utilizes agglutinative morphology, in which numerous morphemes are connected to a root word to manifest several grammatical functions including tense, aspect, mood, person, and number (Ahmadi, 2021). This means that inaccurate translations may result from GT's inability to properly evaluate and understand these intricate morphological forms. Morphology of Central Kurdish includes both derivational and inflectional morphemes which differs from that of English (Ahmadi, 2021).

In the former case, singular nouns usually have the suffixes '-an' or '-kan' attached to them, which is the number for pluralization as a branch of inflectional bound morpheme (Rasul, 2022).

On the other hand, according to (Haig & Matras, 2002), Kurdish language in general including Central Kurdish syntactically possesses a subject-object-verb phrase structure along with a tense-aspect-modality and person marking framework. Furthermore, transitive verbs in the past tenses in Central Kurdish are identified with an agentive case that does not correspond to the nominative case since the dialect is split-ergative (Manzini et al., 2015). According to (Esmaili & Salavati, 2013), Central Kurdish utilizes simply distinct pronominal enclitics for ergative-absolutive symmetry. Because the syntax of Central Kurdish and English differ, there can be difficulties while translating between the two languages utilizing GT. More precisely, the SOV structure in Central Kurdish could be inaccurate for SVO in English, which would result in an inaccurate word order in the English rendering by GT. Also, the split-ergative structure of Central Kurdish, particularly in past tenses, can be challenging when translating (Yates, 2018). In this sense, sentence construction may be inaccurate as a result of GT's inability to accurately distinguish the subject from the object. In this respect, an English translation of a Central Kurdish statement with an agentive case may give rise to an incorrect subject-object agreement.

In addition, verb morphology and pronominal enclitics are utilized variably in Central Kurdish than in English (Ahmadi, 2021). These differences might not be appropriately addressed by GT, resulting in translations that have inaccuracies in subject-verb concordance, aspect, or tense. In this situation, an English sentence written in the past tense in Central Kurdish could potentially be translated into the present tense. Furthermore, perhaps unique pronominal enclitics that Central Kurdish employs for ergative-absolutive parallelism do not translate correctly into English, which could result in inaccurate or missing links to pronouns by GT (Ahmadi & Masoud, 2020).

Semantics presents another difficulty when translating Central Kurdish into English. Given that semantics takes a word, phrase, or sentence at its literal meaning. Since each regional dialect in Sorani Kurdish has its own speech pattern, GT needs help accurately translating linguistic expressions in Central Kurdish, particularly idiomatic expressions.

The linguistic aspect of stylistics presents another potential difficulty in the GT of translating Central Kurdish into English. (Sawant & Hingoli, 2013) maintain that stylistic evaluation involves examining grammar, vocabulary, tropes, context, and cohesion, etc. Therefore, examining language use, such as tone including formality, and informal phrases, is crucial for understanding the text's significance and efficacy in the translation of both target language and source language by GT. (Munday et al., 2022) emphasize the importance of tone, register, and form in a writer's intention. They argue that stylistic problems can arise from inaccurate word selections and variable inaccuracies when the target language is not accurately represented in the translation results.

Another challenge in the translation of Central Kurdish into English via GT is the lexis in particular the omission of some linguistic expressions. Due to the significant differences in culture and language between Sorani Kurdish and English, deletions are a common issue in converting between the two tongues using GT. Numerous challenging structures of grammar, slang terms, and culturally specific phrases exist in Sorani Kurdish which possess no direct equivalent in English. Because of its reliance on feature detection and device intelligence, GT frequently fails to appropriately recognize single statements, despite the fact that those words could be crucial to communicating a passage's complete meaning. Consequently, the machine might completely ignore those expressions, which would result in a deletion of significant psychological or situational complexity. Another aspect of difficulty is added by the fact that Sorani Kurdish depends on situation to express sense; this means that exclusions can happen because Google Translate may not completely comprehend the circumstances. As a result, there may be inconsistencies or even incorrect renditions, which complicates the general interaction between people of the two different tongues (Way, 2013).

2.3. Previous Studies

This section reviews the studies surrounding this research study, offering a critical investigation of the theories, approaches and outcomes that have formed the study. The literature review seeks to shed light on the important additions while highlighting gaps and fields for further studies. Through synthesizing similar and relevant literature, this part of the research acts as a foundation for the present study, manifesting its significance and application within the educational and academic fields. Recently, many studies have been conducted for this purpose. In a study by (Pham et al., 2022) aimed to show the perceptions of university students employing GT concerning the problems and the solutions, the

researchers asked 250 students at the university utilizing a Likert-scale survey questionnaire and interviews to explore their perceptions about employing GT. Their findings revealed that learners had positive views on using GT and found it supportive for learning language because of the characteristics such as multi-language translation and ease of use. A related study by (Mastang et al., 2024) at State Islamic University used descriptive qualitative approach by distributing questionnaires and interviews showed that students had positive perceptions about easy access of GT and negative attitudes about the accuracy of MT results.

In a different study from the above articles, (Tongpoon-Patanasorn & Griffith, 2020) examined the quality of academic abstracts translated from Thai to English utilizing GT. They identified the most common errors and evaluated their comprehensibility and usability. The results manifested that the quality is not suitable for abstract writing. In line with these findings, (Sun et al., 2022) explored types of errors found in GT from Chinese-to-English translations by comparing GT translations to the original English abstracts and analyzing the problems in the translations. In similar attempts to the present study, (Rasul, 2022) evaluated the accuracy of English-Sorani Kurdish translations generated by GT through adopting textual analysis. Findings revealed that GT is accurate at morphological and syntactic levels, however, at the level of grammatical and idiomatic levels, the translation needs improvement. In a similar vein (Muhealddin et al., 2024), aimed to explore and evaluate the quality of GT in English-to-Sorani Kurdish translations by analyzing machine translations from a qualitative and quantitative perspective, and investigating morphological, semantic, and syntactic characteristics in comparison to human translation by employing BLEU, NIST, and BERT Score. Their results indicated that GT acts at a lower level in comparison to the performance of Kurdish original language. Contrary to previous study, (Ameen & Ahmed, 2023) assessed the performance of MT from Kurmanji Kurdish-to-English by employing GT and Bing Translator. They used error analysis method to evaluate the quality of political and economic texts. The results showed that both systems performed well in terms of orthography, lexis, and grammar, yet GT slightly outperforms Bing Translator.

Apart from the last three studies, most of the previous studies are unrelated to the present research. What differentiates this research study from them is that the current research is conducted more comprehensively by seeking to uncover deeply the errors found in Kurdish-English GTs through employing both textual analysis method and the guide sheet for error identifications and classifications.

3. Methodology

This section consists of research design, data collection, error identification and analysis.

3.1. Research Design

In order to assess how accurate, the translations of GT from Central Kurdish into English are, this research utilizes a qualitative approach. The data comprises eight reports that were randomly chosen from the official websites of Rudaw and K24, two famous Kurdistan television networks. To give a thorough foundation for assessment, the textual analysis is used in the current research where the texts address an extensive variety of topics, such as political thought, culture, society, media, movies, athletics, medical care, and perspective. Also, finding faults in the translations was the main method of study utilized to assess GT's accuracy. The goal of the investigation is to discover translation errors that GT makes, focusing closely on grammatical, syntactical, morphological, semantic, lexical (omission) and stylistic errors. Through the identification of the aforementioned errors, textual analysis affects translation quality. This analysis demonstrates the way the flaws impact the accuracy and intelligibility of GT's Kurdish to English translations, bringing to light issues including grammatical faults and misunderstandings that reduce the general efficacy of translation. The research also evaluates linguistic faults in the aforementioned errors by using the (Costa et al., 2015) approach. Following error detection and human correction, the investigation analyzes the lingual elements required for precise rendering.

3.2. Data Collection

The materials selected for this investigation were determined by a number of factors; initially the selected texts were issued in Central Kurdish by reputable and widely recognized media organizations, Rudaw and K24, which are regarded as reliable providers of local reporting. To guarantee diversity in materials and setting, the study addressed a broad spectrum of subjects, including social and political issues, societal, creative (film), athletics, physical well-being, and perspective. To assess how effectively GT manages various topic content and language variations, this diversity is crucial.

Another problem concerns the procedure for translating. In this sense, GT was employed to convert every single chosen text from Central Kurdish to English. The procedure proved simple where with any assistance from humans or modification, the source materials were entered into GT, and the matching English translations were compiled. This phase made guarantee that the assessment was limited to the programmed results generated by GT, giving an objective foundation for error assessment.

3.3. Error Identification and Analysis Categorization of Errors

The subsequent categories applied to the errors found in the analysis of GT:

1. Grammatical errors which comprise problems with sentence construction, tense, subject-verb concord, and incorrect application of articles or prepositions. 2. Lexicon inaccuracies which include improper vocabulary selection, particularly when a translated term is inappropriate for the situation. 3. Syntactic Errors: Problems related to word order and sentence construction that affected the coherence or meaning of the translation. 3. Syntactical inaccuracies which covers issues with sentence structure and placement of words which impacted the consistency or content of the translation. 4. Situational inaccuracies which contain translation flaws that result from a failure to properly interpret various cultural or societal distinctions. Also, it prevents the primary text's purposeful sense from being accurately conveyed. A brief mention of a comparative analysis of GT with human translations is particularly pertinent to this section. In this way, the identical texts were translated by humans and GT to evaluate the quality of the output. Consequently, there were instances in which GT did not accurately translate the texts from Sorani Kurdish into English in terms of linguistic features of grammar, syntax, morphology, lexis (lexical deletion, and stylistics).

4. Results and Discussions

Based on the types and kinds of error analysis, the errors are classified and analyzed under grammatical, syntactic, semantic, omission and stylistic errors. This method of analysis and categorization is sometimes referred to as "Linguistic Error Analysis". While analyzing errors, the main issue is that sometimes certain items have multiple errors, as in the case of translating the title of the first type of text "The state raided the city like a winter forty". This translation contains multiple syntactic, collocational, and mistranslation errors resulted in a loss of meaning and understanding. It appears that the problems are originated from mistranslating Kurdish idiomatic and cultural items that lead to a nonsensical English sentence.

Grammatical Errors

There are various grammatical errors in the eight chosen texts. In the first example of the text, in examples such as "What Murad tells in a modern way", the subject "What Murad tells" is better followed by a more accurate predicate like "portrays" or "depicts" so as to show the natural methods of storytelling. When reading expression such as "In a high poetic language", it is grammatically better and artistically more natural to write "in highly poetic language" or "with elevated poetic language". In the second text (Rudaw, 2024b), the use of preposition "on" in the sentence "Hiding Saddam Hussein" will be screened in Kurdistan Regional Government (KRG) cinemas on November" is incorrect when referring to a month. In this case, it must be "in November". Moreover, there is redundancy in the context of the same translated text, for instance, phrases such as "left it as a suggestion for viewers" and "those over the age of 15 can see the film" are monotonous and could be shortened. The translated text also shows inconsistent use of tenses. While the rest of the text employs past tense, the tense in "The director is awaiting a response from the Iraqi Ministry of Culture, Tourism and Archaeology" is shifted to present tense, therefore, depending on the context, it is better to be matched to the rest of the text. Similarly, in the third text (Rudaw, 2024c), there is an orthographic error such as punctual which encompasses an error involving improper use of punctuation with a conjunctive error in the sentence "...can now walk and his movements have returned to normal" in the third passage 'where a comma (,) and 'and' should be placed after 'walk', so that the sentence it can be accurate in grammar.

Additionally, in the fifth text (Rudaw, 2024e), there is a grammatical error in the sentence "Dr Hami Salih Harki, an management and economic expert who also owns a transportation company...". Given that the indefinite article "a" must occur before words that begin with constant, "a management" is the right phrase in this context rather than "an management. In addition, in the fifth text (Rudaw, 2024e), there is inconsistency in the use of tenses in both "will be supplied" and "it is exported" as switching to present tense in the same situation would confuse the tense timeline.

According to the first translated text from (Kurdistan24, 2024a), the sentence, "We have benefited greatly

from the ideas of Anne Hidalgo, the mayor of Paris through a project that saved Paris from drowning traffic and Erbil wants from today Let him take that step," should be correctly punctuated by using comma after "Paris" in "We have benefited greatly from the ideas of Anne Hidalgo, the mayor of Paris, through a project that saved Paris from drowning in traffic. Erbil wants, from today, to take that step." In the same text, the phrase "It starts July 26 and runs until August It will continue," is improperly capitalized. Moreover, the sentence, "Erbil wants from today Let him take that step" should be corrected in terms of pronoun as "Erbil wants to take that step starting today." On the other hand, in the seventh text by (Kurdistan24, 2024b), in terms of punctuation and use of commas, the sentence, "The Iraqi government has no plans for the housing sector, and the quality of investment projects in Iraq is very poor," the structure is poor, the second part is a different thought and it should be connected with a semicolon or broken into two separate sentences. In the same text, the quotation mark around the expert's statement at the beginning of the structure, is missing. Therefore, it should be corrected as "The Iraqi government has no plans for the housing sector and the quality of investment projects in Iraq is very poor," an expert in commercial and investment law stated."

In the eighth text by (Kurdistan24, 2024c), the phrase "An Arab citizen of Erbil named Ziad Abdulqadir, 49 years old," is not a full sentence. It could be revised and linked to the previous sentence as "Lieutenant Hawkar Kamal, an officer in the Soran Civil Defense, told Kurdistan 24 that the drowned Arab citizen was named Ziad Abdulqadir, 49 years old." In the first sentence of the eighth text (Kurdistan24, 2024c), "An officer in the Soran Civil Defense said: The citizen came to the place for a trip and drowned while swimming and teams continue to search for his body," could be restated and use comma instead of a colon that requires quotation marks to show reported speech as "An officer in the Soran Civil Defense said, 'The citizen came to the place for a trip and drowned while swimming. Teams continue to search for his body.'"

Syntactic Errors

In the first type of text, it is better to use conjunctive adverbs or relative clauses to link sentences like "Like Yasar Kemal, he is a Kurds" and "What Murad tells in a modern way is the story of streets of Amed". Moreover, in the same text, syntactically it is smoother and clearer to write "the poverty and helplessness experienced by the Kurdish nation" rather than "of the poverty and helplessness". In the second type of text, the use of "so" in "so those over the age of 15 can see the film" implies a causal relationship, yet the structure of the sentence is awkward. It might be more appropriate to include "therefore" and write the sentence as "therefore, viewers over the age of 15 are also permitted to see the film". Furthermore, in the same text, in terms of punctuation, it would be clearer to replace the use of comma before "so" in "but did not specify the age requirement and left it as a suggestion for viewers, so those over the age of 15 can see the film" with a period to construct to independent sentences. Correspondingly, in the third text by (Rudaw, 2024c), there is the use of redundant phrases. For instance, repeating introducing phrases such as "Hiwa Mustafa, a specialist in cardiovascular surgery and supervisor of Halabja Governor Azad Tofiq's surgical team" and reusing phrases like "Dr. Hiwa Mustafa, a specialist in cardiovascular surgery" is a syntactic redundant phrase as one introduction phrase is sufficient for clarity. Moreover, in the third passage, for the sake of clarity and better flow, the sentence "The expert said people undergoing this type of surgery should reduce fatty foods and pay attention to daily physical activity, even avoid smoking cigarettes and hookahs", would be rephrased as "The expert advised that people undergoing this type of surgery should reduce their intake of fatty foods, pay attention to daily physical activity, and avoid smoking cigarettes and hookahs."

In the fourth passage, the name of "Luka Modric" is used almost in every sentence. Although, it is not incorrect, this sort of repetition conveys awkward syntactic tone. The use of pronouns "He has been with Real Madrid since 2012..." would make the sentence more accurate. This sentence structure "Luka Modric has been with Real Madrid since 2012 and has been one of the legends of the club and world football for the past 12 years" is better to be rephrased as "Luka Modric has been with Real Madrid since 2012 and has become a legend in both the club and world football over the past 12 years." Moreover, the use of parallelism in "including six Champions League titles, five Club World Cups, four European Super Cups, four La Liga titles, two Copa del Rey titles, and five Spanish Super Cups" could be better concisely rewritten as "six Champions League titles, five Club World Cups, four European Super Cups, four La Liga titles, two Copa del Reys, and five Spanish Super Cups."

Moreover, there is a syntactic error in the text five in the phrase "It is" in the clause "In addition to exporting to Afghanistan, the Kurdistan Region's gasoline products are also exported from Bandar Abbas to other countries. It is first transported to Dubai and from Dubai to China and many other countries". This is because a plural verb was needed to accompany the plural noun "it is referred to" ('products,' in this

case). Therefore, rather of saying "It is first," the proper form is "They are first." This is due to the fact that the inflectional morphology of "they", i.e., "products" in terms of number must agree with a plural verb (are); here, "they" is an anaphora for the plural noun "products". In other words, "they" must coordinate with the verb "are" to morphologically convey plurality.

Furthermore, in the fifth text (Rudaw, 2024e), there is a run-on sentence. for instance, "Gasoline and oil products from the Kurdistan Region are transported from refineries to Parwizkhan. They will be transferred to Afghanistan" could be better rephrased "Gasoline and oil products from the Kurdistan Region are transported from refineries to Parwizkhan. They will be transferred to Afghanistan." The structure "He believes the problem is political and is due to the fact that Iran itself and its companies export gasoline and oil to Afghanistan Do not buy Kurdistan Regional Government oil" is syntactically incorrect that could be revised as "He believes the problem is political, arising from the fact that Iran and its companies export gasoline and oil to Afghanistan, leading to reluctance to buy oil from the Kurdistan Regional Government."

Concerning syntactic errors in the sixth text by (Kurdistan24, 2024a), the phrase "It starts July 26 and runs until August It will continue," has to be revised to make a complete sentence "It starts on July 26 and will continue through August." It could be better to split the sentence "The governor of Erbil referred to the project to build a cycle path in Erbil and said: We have benefited greatly from the ideas of Anne Hidalgo, the mayor of Paris through a project that saved Paris from drowning traffic and Erbil wants from today Let him take that step," into separate, coherent structures.

Also, in the seventh text by (Kurdistan24, 2024b), there is a syntactic fragment in "Four million housing units should be built in Iraq by 2027 to solve the problem of random housing, the Strategic Center for Human Rights said," it would be more appropriate to rephrase it as "The Strategic Center for Human Rights stated that four million housing units should be built in Iraq by 2027 to solve the problem of random housing." In the seventh text by (Kurdistan24, 2024b), the run on sentence "while an economic researcher said: 120,000 housing units are under construction but have been distributed to officials before completion," causes confusion. The revision could be more appropriate if it is rewritten "An economic researcher noted that while 120,000 housing units are under construction, they have already been allocated to officials before completion."

Besides, there is a misplaced modifier in the text eight by (Kurdistan24, 2024c) in "An officer in the Soran Civil Defense said: The citizen came to the place for a trip and drowned while swimming and teams continue to search for his body," that could more clarified as "According to an officer in the Soran Civil Defense, the citizen was on a trip when he drowned while swimming. Teams are continuing to search for his body." There is a lack of clarity in the sentence "On Wednesday, July 17, 2024, Lieutenant Hawkar Kamal, an officer in the Soran Civil Defense told Kurdistan24: An Arab citizen of Erbil named Ziad Abdulqadir, 49 years old," that could be clearer if it is revised as "On Wednesday, July 17, 2024, Lieutenant Hawker Kamal, an officer in the Soran Civil Defense, told Kurdistan24 that the citizen was an Arab resident of Erbil named Ziad Abdulqadir, who was 49 years old."

Semantic Errors

Each of the eight texts that were chosen has a different semantic error. In the first type of text, the expression "this huge state" in the phrase "shows the dark and frightening face of this huge state" is somehow vague and unclear for the reader as they might not be familiar with the context. Therefore, it is more appropriate to write as "the oppressive face of the Turkish state" or "the overwhelming power of the state" relying on the intentional meaning. Similarly, depending on the context, it is more accurate to write a word such as "reimagines" or "retells" instead of using "rewrites" in the sentence "He rewrite the story of poverty and helplessness of the Kurdish nation" as it might imply that he is changing the existing narrative rather than introducing it in a modern style. In the second text (Rudaw, 2024b), the sentence "but did not specify the age requirement and left it as a suggestion for viewers, so those over the age of 15 can see the film" is ambiguous and unclear because it implies a suggestion which is not clarifying specifically age of restriction. In terms of the word choice in the same type of text, the word "partner" is likely vague and unclear semantically in "Hane, the main partner of the Norwegian government". Therefore, it might be better to say "the principal collaborator" instead.

The phrase "transplanted five heart arteries" in the third text by (Rudaw, 2024c), is ambiguous semantically, usually for better semantic accuracy, the term "bypass" is employed in this situation "performed a quintuple bypass." Furthermore, the expression "had a blockage in his heart" is confusing since it is not clear whether there was a single blockage or multiple blockages is more accurate. In the fourth text by (Rudaw, 2024d), there is an ambiguous sentence including "The Croatian captain is one of the first

players to win six Champions League titles, which will remain forever in the history of Real Madrid football and his career” that could be clarified and rephrased for accuracy.

There is a semantic ambiguity in the fourth passage in the sentence “On an individual level, Luka Modric won the 2018 World Player of the Year award and the Ballon d’Or.” Since it might be vague to readers that both the Ballon d’Or and World Player of the Year are the same in terms of purpose and appreciation, the reader or audience should have background knowledge about them. Also, there is another semantic error in terms of wrong choice of word in the sentence “Modric has renewed his contact with the club”, here the word ‘renewed’ suggests that the contract was done personally by Modric. However, it was the club that did contract. In this sense, the right replacement word is “extended” which indicates that the contract was done by both club and Modric. In text five, the expression “This is Kurdistan Regional Government gasoline, but Iranian officials only mention Iraq” is ambiguous, because it is not clear if the gasoline is from Kurdistan or whether it is originated from Iraq. In the same text, there is an inconsistent use of terminology, the use of “gasoline” and “oil products” could confuse the reader. It could be more accurate and concise to distinguish between “gasoline” and “broader oil products.” In terms of semantics, text six by (Kurdistan24, 2024a), contains an ambiguous phrase such as “Erbil wants from today Let him take that step,” where using “him” semantically causes confusion. In the seventh text by (Kurdistan24, 2024b), the phrase “The Strategic Center for Human Rights said,” is semantically vague and could be better revised as “According to the Strategic Center for Human Rights, four million housing units need to be built by 2027 to address the issue of informal settlements and random housing.”

In the eight text by (Kurdistan24, 2024c), there is unclear term such as “team”, as it is vague that could be used more specifically such as “rescue teams” or “search teams,” to offer a more clarified comprehension.

Omission Errors

The first text translated by Google Translate published by (Rudaw, 2024a) contains some missing parts that are skipped. For instance, expressions like "به تورکی و به کوردی ژان و ناسۆری نه ته وه که ی دهنووسپته وه" and "سێکردنی" are not translated. Another instance of omission is "نووسه ی ئه وه به ره هه مه زۆر به ئاسانی ده توانی له تۆره" . In the second type of text, Google Translate skipped translating this "کۆمه لایه تییه کاندایه که ی خۆی، وه ک خۆی ب لای و بگاته وه" into English. Sometimes Google translate confuses the source language text, for instance, titles like “Dr.” without names and the omission of “Fazel Badrani” in the first place and mentioning it in the text without the title in “In an interview with Rudaw, Dr.” “We are open... Fazel Badrani” confuses the readers. Moreover, Google Translate, mistranslates “ئیمه ده ستمانی کراوه یه” as “we are open” and misses translating “به یز وه زیی روشنیی” in the context of the original text.

Also, there is a subject omission in the third text by (Rudaw, 2024c) “they transplanted five heart arteries”, where the referent of “they” is ambiguous”; hence, “they” as a plural pronoun should refer to “the team of surgical operation including Dr. Hiwa himself”, since he is a specialist in the cardiovascular. Likewise, in the fourth text by (Rudaw, 2024d), in the sentence “Modric has won 26 different titles with Real Madrid, including six Champions League titles, ...”, some part of the sentence is omitted such as during his professional path, at the beginning of the sentence to refer to the context of his achievements in football in the in the past.

Meanwhile, there is a missing preposition of ‘on’ and period omission with lack of capitalization of (it) in the text six by (Kurdistan24, 2024a), in the sentence “It starts July 26 and runs until August it will continue”. The corrected sentence is “It will start on July 26 and runs until August. It will continue”.

Stylistic Errors

There are some stylistic errors in the selected texts. In the first text written by (Rudaw, 2024a), there is a redundancy in the expression “how did these sentences escape the scissors of state censorship”, where ‘escape’ and ‘scissors’ could suggest avoiding something limited; therefore, one of them should be removed; the whole phrase can be rephrased as following “how did these sentences circumvent state censorship”. Likewise, in the second text by (Rudaw, 2024b), there is another stylistic error, it would be better not to repeat “at” in these phrases “The documentary will be screened at Empire Cinema in Erbil at 14:00, 16:30, 19:00 and 20:00, and at City Cinema in Sulaimani at 16:00, 18:00 and 20:00 and 22:00) and Duhok Mall Cinema in Duhok will have two screenings at (15:30 and 19:00)”. Since it is also redundant, it is more accurate to rephrase them as “The documentary will be shown at the following times: in Erbil's Empire Cinema (14:00, 16:30, 19:00, 20:00); Sulaimani's City Cinema (16:00, 18:00, 20:00, 22:00); and Duhok Mall Cinema (15:30, 19:00).” When translating the second text by (Rudaw, 2024b) through Google Translate, there are instances of informality in terms of choosing lexical items, hence it is more appropriate for

example to rephrase these expressions "I don't want to defend him," Iraqi Deputy Minister of Culture, Tourism and Archeology Fazel Badrani told Rudaw "Mr. Minister of Culture, but his office is always open and crowded for artists, intellectuals, writers and academics" as "I am not here to defend him," said Iraqi Deputy Minister of Culture, Tourism, and Archaeology Fazel Badrani in an interview with Rudaw. "However, the Minister of Culture's office remains open and busy, welcoming artists, intellectuals, writers, and academics" as the tone is informal and to some extent disjointed. Another example in the second type of text is that the original text is translated as "The film is scheduled to be released worldwide this year and had its first international exhibition in November 2023 in Amsterdam, its first Arab exhibition in Jeddah, Saudi Arabia and has had private exhibitions in several European and Arab countries". For the sake of clarity and consistency, it is better to break the structure into two or more independent sentences, and change the inconsistent use of "has had" in the sentence to generate a text like this "The film is scheduled for worldwide release this year. It had its first international exhibition in November 2023 in Amsterdam, its first Arab exhibition in Jeddah, Saudi Arabia, and has been privately exhibited in several European and Arab countries." Therefore, in order to improve the quality of the text, it would be better to rephrase the structures as "In an interview with Rudaw, Iraqi Deputy Minister of Culture, Tourism, and Archaeology Fazel Badrani said, "We are open. I will welcome him in my private office in Baghdad on Sunday. I am not here to defend the Minister of Culture, but his office remains open and is always bustling with artists, intellectuals, writers, and academics."

In addition, there is a stylistic error in the third text by (Rudaw, 2024c), in the phrase "Wednesday, July, 17", which contains a redundancy. It can better be rephrased as "On Wednesday" to avoid repetition and redundancy. In the third text by (Rudaw, 2024c), the repetitive use of governor's name and titles "Azad Tofiq, the governor of Halabja" is unnecessary. Besides, there is a mixture use of formal and informal language, for instance, the phrase "even avoid smoking cigarettes and hookahs" could be more readable and consistent and is rephrased as "The expert recommended avoiding tobacco products, including cigarettes and hookahs. In the fourth text by (Rudaw, 2024d), there is an overuse of titles and honors when listing achievements. This could be rephrased as "Modric's time at Real Madrid is marked by an incredible 26 titles, including..." for a more compelling narration. In the fifth text by (Rudaw, 2024e), the repetition of "Kurdistan Region's gasoline products," "Iran," and "Afghanistan" could be reduced by adding pronouns and various sentence structures to improve readability. In order to produce a clearer and precise text, the sentence "The price of a ton of diesel in the Kurdistan Regional Government (KRG) is about \$500, while in Afghanistan it will rise to \$650 to \$700 a ton after increased transportation and customs costs," could be revised as "Diesel costs about \$500 per ton in the KRG, but rises to \$650-\$700 per ton in Afghanistan due to higher transportation and customs fees."

In the sixth text by (Kurdistan24, 2024a), the repetition of "governor" and "French" could be reduced to improve the accuracy and the style. For instance, instead of using "The governor of Erbil," "The governor thanked", and "the governor of Erbil referred to" could be rewritten in various sentence structures such as "Governor Omid Khoshnaw thanked... He also referred to..." Besides, there is no stylistic consistency in using titles. For instance, instead of using "French Consul," "French Consul General," and "French Consul in the Kurdistan Region," a single consistent title is better to be employed throughout the text. Likewise, there is an ambiguity and verbosity in the passage six in the sentence "We have benefited greatly from the ideas of Anne Hidalgo, the mayor of Paris through a project that saved Paris from drowning traffic and Erbil wants from today Let him take that step"; the whole sentence is vague in terms of meaning. It can be improved as follows "..., through a project that lessens traffic jams in Paris. The objective of our project is to implement comparable policies in Erbil". Currently, the sentence is more obvious and understandable for readers. Furthermore, in the text six by (Kurdistan24, 2024a), the verb phrase "...expressed his pleasure for the participation and presence of the governor of Erbil in the opening ceremony of the Paris Olympic Games" contains multiple style errors involving conjunctions. Thus, "...expressed his pleasure for the participation of the governor of Erbil" would be a better way to frame it. This enhances readers' comprehension of the content.

On the other hand, there are other stylistic errors in terms of repetition and redundancy in the seventh text by (Kurdistan24, 2024b), in the noun phrase "expert on trade and investment" and the verb phrase "The Iraqi government has no plans for the housing sector, and the quality of investment projects in Iraq is very poor." These stylistic errors are found in the passage "The Iraqi government has no plans for the housing sector, and the quality of investment projects in Iraq is very poor," an expert in commercial and investment law said. Mukhlis Salim, a legal expert on trade and investment, told Kurdistan24: Iraqi citizens do not benefit from investment projects. The Iraqi government has no plans for the housing sector, and the

quality of investment projects in Iraq is very poor". To enhance the text's coherence and clarity and prevent repetition and redundancy, it is suggested to rephrase it as follows. "Mukhlis Salim, a legal expert on trade and investment, told Kurdistan24 that the Iraqi government has no plans for the housing sector, and that the quality of investment projects in Iraq is very poor. He also mentioned that Iraqi citizens do not benefit from these projects". Additionally, this textual improvement sustains the attention of the reader by making the text straightforward to follow and comprehend. In the seventh text by (Kurdistan24, 2024b), the phrase "The Iraqi government has no plans for the housing sector, and the quality of investment projects in Iraq is very poor," is repeated which would be more accurate to rephrase it as "According to Salim, the Iraqi government lacks strategic planning for the housing sector, leading to the poor quality of investment projects." The sentence "we are proud of the investment law of the Kurdistan Region," should be paraphrased in order to maintain a formal tone because it discusses the policies of government. It could be rephrased as "The investment law of the Kurdistan Region is considered exemplary."

In the eighth text by (Kurdistan24, 2024c), the phrase "Teams continue to search for his body," is unnecessarily repeated twice." It could be concisely written as "Lieutenant Kamal added that the teams continue to search for his body, marking the ninth drowning case in the region this year." Besides, the phrase "This is the ninth drowning case in the region this year," would be more formally written as "This marks the ninth drowning incident in the region this year," to match with the tone of news reports.

The study concludes that there is inconsistency in the accuracy of Google Translate. Due to the grammatical, cultural, and stylistic complexities in Central Kurdish, Google Translate is typically unable to provide precise translations. Mistranslation and omission are often found in translation process results in clarity and potential misunderstandings. Based on the findings, syntactic and grammatical structures and semantic errors are the frequent linguistic features. Word order, tense, aspect and morphology of central Kurdish poses challenges that Google Translate fails to translate accurately. While Google Translate works as a practical machine translation tool for basic translation tasks. For Central Kurdish, the translation by Google Translate requires improvements. The results of the study provide future enhancements in Google Translate, eventually mastering better communication and comprehension between both languages.

5. Conclusion

The study concludes that there is inconsistency in the accuracy of Google Translate. Due to the grammatical, cultural, and stylistic complexities in Central Kurdish, Google Translate is typically unable to provide precise translations. Mistranslation and omission in translation often results in clarity and potential misunderstandings. Based on the findings, syntactic and grammatical structures and semantic errors are the linguistic errors. Besides, word order, tense, aspect and morphology of central Kurdish poses challenges that Google Translate fails to translate them accurately. While Google Translate works as a practical machine translation tool for basic translation tasks, for Central Kurdish, the translation by Google Translate requires improvements. The results of the study provide future enhancements in Google Translate so as to master better communication and comprehension between both languages.

6. References

- [1] Ahmadi, S. (2021). A Formal Description of Sorani Kurdish Morphology. <https://doi.org/10.48550/arXiv.2109.03942>
- [2] Ahmadi, S., & Masoud, M. (2020). Towards Machine Translation for the Kurdish Language. <https://doi.org/10.48550/arXiv.2010.06041>
- [3] Ameen, H. H. M., & Ahmed, H. A. (2023). Assessing the Quality of Machine Translation from Kurmanji Kurdish into English. *Academic Journal of Nawroz University*, 12(3), 503-517. <https://doi.org/10.25007/ajnu.v12n3a1690>
- [4] Costa, Â., Ling, W., Luís, T., Correia, R., & Coheur, L. (2015). A linguistically motivated taxonomy for Machine Translation error analysis. *Machine Translation*, 29, 127-161. <https://doi.org/10.1007/s10590-015-9169-0>
- [5] Dorr, B., Snover, M., & Madnani, N. (2011). Part 5: Machine translation evaluation (Vol. 936).
- [6] Esmaili, K. S., & Salavati, S. (2013). Sorani Kurdish versus Kurmanji Kurdish: an empirical comparison. *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*,
- [7] Fadlan, A., & Ichsan, I. (2022). A Linguistic Error Analysis of Google Translate Translation Result From Bahasa Indonesia To English. *Humanitatis: Journal of Language and Literature*, 8(2), 193-204. <https://doi.org/10.30812/humanitatis.v8i2.1560>
- [8] Groves, M., & Mundt, K. (2015). Friend or foe? Google Translate in language for academic purposes. *English for Specific Purposes*, 37, 112-121. <https://doi.org/10.1016/j.esp.2014.09.001>
- [9] Haig, G., & Matras, Y. (2002). Kurdish linguistics: a brief overview. *STUF-Language Typology and Universals*, 55(1), 3-14. <https://doi.org/10.1524/stuf.2002.55.1.3>
- [10] Hasanpour, A. (1992). Nationalism and Language in Kurdistan. In: San Francisco: Mellon Press.
- [11] Hutchins, W. J. (1986). Machine translation: past, present, future. Ellis Horwood Chichester.
- [12] Hutchins, W. J., & Somers, H. L. (1992). An introduction to machine translation. Academic Press Limited.

