A Text Typological Assessment of Automatic Translation: A Case Study of English into Kurdish Translated Texts by OpenAI GPT and Google Translate

Fereydoon Rasouli[1], Keivan Seyyedi[1], Alan H. Azaldin[2]

[1]Department of Translation, Cihan University-Erbil,
Kurdistan Region, Iraq
[2]Department of Technical Translation, Polytechnic University- Erbil,
Kurdistan Region, Iraq

***Abstract*—The extensive use of artificial intelligence in translation projects has motivated this study to evaluate the quality of two popular systems, GPT OpenAI and Google Translate. To assess the systems' performance, 15 sentences were chosen from a variety of text types based on Katharina Riess's text typology informative, expressive, and vocative, and were used as input data for translation into the Kurdish language. The output translations were then assessed using two evaluation metrics: Bilingual Evaluation Understudy (BLEU) and translation edit rate (TER). The findings revealed that overall, GPT outperformed Google Translate, as it achieved a higher BLEU score reflecting better choices in equivalence and sentence structure and a lower TER score, indicating fewer necessary corrections in the translated text compared to the human (reference sentences) translation. In particular, GPT presents a better performance in the translation of expressive and vocative texts, where understanding emotions and persuasive language is more difficult.**

***Keywords*—Artificial intelligence, Google translate, Kurdish language, Machine translation, OpenAI GPT, Translation quality assessment.**

I. Introduction

The rapid advancements in artificial intelligence (AI) have influenced numerous aspects of human life, particularly communication across different languages and cultures. Since translation serves as a crucial bridge in cross-cultural interactions, it has been influenced by technological progress. Over the past few decades, machine translation (MT), which enables automated language conversion, has undergone significant transformations. The necessity of operating MT in the process of texts has increased due to the dramatic rising of information exchange (Mohamed et al., 2024, P. 34).

The development of AI technology has brought about new chances of communication across different language-speaking communities. The main concern of the present study is to examine the direct impact of new achievements in the area of AI on the automatic translation of Kurdish text. To be more comprehensive in conducting the survey, different text typologies go through examination and two popular platforms of Google Translate and OpenAI GPT are administered to translate texts from English into Kurdish.

II. Literature Review

Recent AI advancements have significantly reshaped MT methodologies. Early translation systems relied on overt grammatical rules and manually provided dictionaries. In rule-based systems (rule-based machine translation [RBMT]), developers structured large sets of linguistic rules and bilingual lexicons to guide the translation of source texts (ST) into target languages (TL) (Rasouli, 2018, p. 4). The process passed through several stages: first, a linguistic analysis of the ST was performed to highlight its morphological and syntactic structure; second, a transfer mechanism mapped those structures to corresponding elements in the TL; then, a generation module produced an output based on pre-

Cihan University-Erbil Journal of Humanities and Social Sciences (CUEJHSS)
Volume X No. 1 (2023) 6 pages
DOI: 10.24086/cuejhss.vol10n1y2026.pp 13-18
Received: 08 April 2025; Accepted: 30 June 2025; Regular research paper: Published 1 January 2026
*Corresponding author's e-mail: fereydoon.rasouli@cihanuniversity.edu.iq

established language rules (Hutchins and Somers, 1992). Although the progressions created in the field of MT were noticeable, it still suffered from linguistic ambiguity, idiomatic expressions, and the structural complexity, particularly in languages with rich morphological feature. While RBMT provided expected translation texts within its limitations, they lacked the flexibility and contextual awareness necessary for accurate translations; making them less effective for complex linguistic structures (Mills, 2023, p. 45).

The introduction of statistical MT (SMT) initiated a new area of MT developments; the SMT was designed based on probabilistic models that enjoyed a large parallel corpus. Using word alignment techniques, early SMT models, such as those developed by Brown et al. (1993), evaluated the similarity between the SL and TL sentences and proposed the result of evaluation as the correct translation of the given source sentence. Later, by focusing on sequences of words rather than individual terms, phrase-based models developed based on these early methods. SMT systems mounted the quality of automatic translation projects in terms of translation accuracy; however, they still face challenges in controlling long-range dependencies and producing naturally flowing sentences (Koehn, 2010).

Generating more coherent and contextually appropriate texts by neural MT (NMT) enunciated the start of a new era in MT studies. Recent systems apply deep learning principles to translate input texts more comprehensibly (Vaswani et al., 2017). Neural MT models consider all sentences as units of sequences and analyze the linguistic relationships between them. Employing an encoder–decoder framework mechanism, these MT systems are able to focus on the most relevant segments of the input text during translation. Consequently, translated projects are generally more fluent and grammatically acceptable in contrast to the previous models (Bahdanau et al., 2015).

The advent of transformers platforms has changed the performance of NMT translation. Before 2017, most translation models used "recurrent" structures (like loops) to process sentences word-by-word, which often struggled with long sentences or complex grammar. Vaswani et al. (2017) created the transformer model, which uses "self-attention" models. The model focuses on all words in a sentence at once, not just one after another. This spotlight helps machines understand connections between distant words like how "it" in a sentence might refer to a noun mentioned five words earlier.

The most up-to-date automatic translation systems, such as OpenAI's GPT (Radford et al., 2018), use the transformer idea to train a model on massive amounts of text. The GPT platform works as a student who reads millions of books and learns to predicate the next word in a sentence. For example, if you write "The cat sat on the…," GPT might predict "mat" because it has seen that phrase often. The ability of prediction helps GPT write translations that sound more natural, like how a human would phrase them. Meanwhile, GPT has a weakness: it only looks at past words (left-to-right) when guessing. Google's BERT (Devlin et al., 2019) fixed this by looking at both sides of a word. Imagine trying to solve a crossword clue where you need letters from before and after the blank. BERT works similarly. For instance, in the sentence "She went to the bank to withdraw money," BERT uses "withdraw money" to confirm "bank" means a financial institution, not a riverbank.

Evolving from early rule-based systems to statistical methods (SMT), then neural networks (NMT), and finally transformer-based models, MT has advanced dramatically over the past decades. However, as Toral et al. (2018) and Prates et al. (2019) state some of the basic problems with the quality of translated texts by MT persist. For instance, cultural differences between SL and TL often lead to the translation problems (e.g., idioms like "break a leg") or specialized terms used in fields like medicine or law.

A. The Assessment Models of MT

To evaluate MT outputs, automatically different systems are proposed by researchers. Meteor is one notable example that was introduced in 2004 to closely mirror human assessments (Lavie and Agarwal, 2007). Meteor works at the word level, scoring translations by directly matching words between the MT output and its reference. When multiple reference translations are available, the system evaluates each separately and selects the highest score (Rasouli et al., 2024, P. 9).

The word error rate, is another common model of MT evaluation that originates from the Levenshtein distance. This model calculates the minimum number of word-level edits – substitutions, omissions, and insertions needed to transform the MT output into the reference sentence (Koehn, 2010; Dobrinkat, 2008).

The most widely adopted automatic evaluation metric for MT outputs is the Bilingual Evaluation Understudy (BLEU) system. This model of evaluation developed at IBM laboratories (Kishore et al., 2001). BLEU provides a quick and cost-effective method for assessing translation quality. In addition, the translation edit rate (TER) by Snover et al. (2006) builds on edit distance concepts but permits block reordering and use extra editing steps to capture word sequence changes. Turian et al. (2003) also presented a model that employs maximum matching strings (MMS). The recent model (MMS) is demonstrated to produce high correlations with human judges. In this study, both metrics of BLEU and TER were selected to assess the quality of the Kurdish outputs of MT systems.

III. Materials and Methodology of the Study

To examine how AI has affected MT performance, this study employed quantitative methods of evaluations. Ensuring that the evaluation reflects diverse contextual challenges, the data of the current study are assembled based on the text typology proposed by Reiss (1971) in different types of informative, expressive, and vocative texts. Sentences translated by Google Translate MT system, OpenAI GPT, and human translator as a reference translation of the study into the Central Kurdish language.

The study utilized several established evaluation metrics. The primary measure was the bilingual understudy

evaluation (BLEU) score (Papineni et al., 2002), which assesses translation accuracy by comparing n-gram overlaps between AI outputs and human reference texts. In addition, to quantify the differences in more details, the TER (Snover et al., 2006) determined the number of modifications required to align AI-generated translations with those produced by humans. Appendix 1 shows the collected data of the study.

IV. Result and Discussion of the Study

To examine the impact of AI on the performance of MT systems, both BLEU and TER metrics were administered on the collected data. As mentioned above, the data of the study were collected based on different text type models proposed by Reiss (1971) to evaluate the performance of the systems in translating of texts from different sorts. Table I shows a brief record of analyzed data that depicts the types of texts in SL and their correspondent translations done by Google Translate MT and OpenAI GPT. The candidate sentences alongside human-translated texts as the reference sentence for the evaluation have been presented in the table as well.

To evaluate the performance of systems, the BLEU score and TER metrics are administrated. BLEU measures N-gram between the MT platforms outputs as the candidates of the study, and related reference translations are done by practiced translators. The scores range from 0 to 1 in a way that the higher score proved the better performance and closer to the reference sentence. The TER test is a means of quantifying the number of necessary edits, such as insertions, omissions, and replacements to convert or assimilate the MT output to a proposed human translation of the same text. Therefore, the lower range of TER equals better performance of the MT.

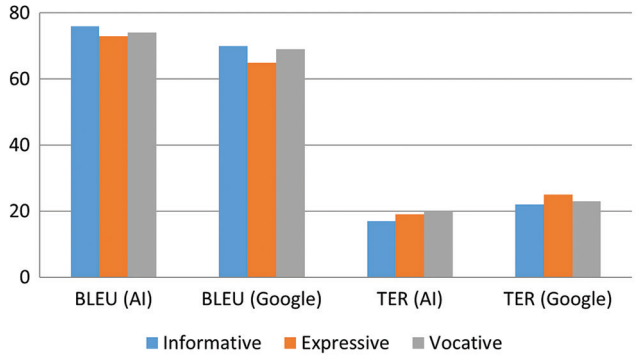
Table II illustrates the overall result of BLEU score and TER classified based on the different text types.

As Table II depicts, the under-study systems perform differently in the translation of different text types. The BLEU scores disclosed the better performance of AI than the Google system in all text types, and the higher differences are in the expressive text, where AI scored 73 in contrast to 65 for Google Translate. In general, the AI-generated translations average a BLEU score of 74 compared to 68 for Google Translate, indicating that the AI translations have a higher degree on n-gram similarity with the reference sentences. In addition, the related data under the TER test shows that translating informative texts by AI needs less correction compared to the translation of the same text type done by Google Translate. Therefore, AI-generated translation of informative texts is more acceptable than that of the same category translated by Google Translate. Based on the average TER of 18.5% versus 23% for Google Translate, it means that fewer edits are needed to translate texts by AI to match the reference sentences. To visualize the comparative findings, chart No.1 highlights the performance differences between systems.

The interpretation of the Bar Chart 1 confirms the data presented in Table II. As presented, across all categories, the AI consistently shows higher values than Google Translate and proves the higher BLEU performance visually. Similarly, AI is consistently lower than Google, indicating that the AI-

TABLE I
Sample of Translation Data

Text Type	Sentence ID	Source (English)	AI Output (Kurdish)	Google Translate (Kurdish)	Human Translation (Kurdish)
Informative	1	The library opens at 9 AM every weekday.	کتێبخانەکە هەموو ڕۆژی کار کردنا لە کاتژمێر ٩ بەیانە دەکرێتەوە.	کتێبخانە هەموو ڕۆژانی هەفتە کاتژمێر 9ی بەیانی دەکرێتەوە.	پەرتووکخانە هەموو ڕۆژێمک کاتژمێر ٩ بەیانی دەکرێتەوە.
Expressive	1	I felt a burst of joy when I saw the sunrise.	کاتێك خۆر هێند بینیم، پر لە شادمانی هەستم.	هەستم بە تەقینەوەی خۆشی کرد کاتێک خۆر هەڵاتنم بینی.	کە دەرکەوتنی خۆرم دیت هەستم بە دڵ خۆشییەکی زۆر کرد.
Vocative	1	Join us today to support our community cleanup!	ئەمڕۆ بەشداربن لە پشتگیری پاککردنەوەی کۆمەڵگایەکەمان	ئەمڕۆ لەگەڵمان بن لە پشتگیریکردنی پاککردنەوەی کۆمەڵگاکەمان	ئەمڕۆ لەگەڵمان کەوە بۆ پشتیوانی کردن لە پاکژ کردنەوەی کۆمەڵگاکەمان.



Bar Chart 1: The Bilingual Evaluation Understudy score and translation edit rate.

TABLE II
Overall BLEU Score and TER

Text Type	BLEU (AI)	BLEU (Google)	TER (AI)	TER (Google)
Informative	76	70	17	22
Expressive	73	65	19	25
Vocative	74	69	20	23
Overall	74	68	18.5	23

BLEU: Bilingual Evaluation Understudy, AI: artificial intelligence, TER: Translation edit rate

generated translations require fewer edits. In other words, the lower TER the more acceptable translated text. Notably, a significant interaction between Translation System and Text Type was identified for both metrics, indicating that the dissimilarity in performance between the two systems depended on the category of analyzed texts. For example, the performance of the AI system in expressive texts is more acceptable than in informative texts. These findings highlight that while the AI translation system generally outperforms Google Translate in quality, the importance of this merit varies systematically with the type of translated text. Appendix 2 shows the analyzed data by python.

V. Conclusion

Overall, the analyzed data of the current study shows that the Kurdish translated version done by GPT outperformed the Google Translate with regard to the provided human translation versions of the same texts. The GPT Translation performance achieved an average BLEU score of 74 versus 68 for Google Translate. In addition, the translated text by GPT needs less correction as displayed by the TER test (18.5% vs. 23%). Statistically speaking, these differences in the performance of both platforms are generally and specifically meaningful when the text types come into consideration; this can be more sensitive in expressive and vocative texts, where emotional and persuasive language create more challenges. Although, both GPT and Google Translate systems exhibit some weaknesses in their performance that could be subjected to future analysis and improvement, the combined qualitative and quantitative assessments proposed that GPT conducted the translation project much better by capturing the content and stylistic features of reference translation than the other.

References

Bahdanau, D., Cho, K., & Bengio, Y. (2015). *Neural Machine Translation by Jointly Learning to Align and Translate*. In: International Conference on Learning Representations (ICLR).

Brown, P.F., Pietra, V.J.D., Pietra, S.A.D., & Mercer, R.L. (1993). The mathematics of statistical machine translation: Parameter estimation. *Computational Linguistics*, 19(2), 263-311.

Devlin, J., Chang, M.W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. p4171-4186.

Hutchins, W.J., & Somers, H.L. (1992). *An Introduction to Machine Translation*. United States: Academic Press.

Kishore, P., Salim, R., Todd, W., John, H., & Florence, R. (2001). *Corpus-Based Comprehensive and Diagnostic MT Evaluation: Initial Arabic, Chinese, French, and Spanish Results*. In: Proceedings of Human Language Technology 2002, San Diego, CA.

Koehn, P. (2010). *Statistical Machine Translation*. Cambridge: Cambridge University Press.

Lavie, A., & Agarwal, A. (2007). *METEOR: An Automatic Metric for MT Evaluation with High Levels of Correlation with Human Judgments*. In: Proceedings of the Workshop on Statistical Machine Translation, Prague. p228-231.

Mills, R. (2023). *Machine Translation Quality and Context*. Proceedings of the 2023 International Conference of Translation Technologies. Vol. 5. p88-99.

Mohamed, Y.A., Khanan, A., Bashir, M., Mohamed, A.H.H.M., Adiel, M.A.E., & Elsadig, M.A. (2024). The impact of artificial intelligence on language translation: A review. *IEEE Access*, 12(2024), 25553-25579.

Papineni, K., Roukos, S., Ward, T., & Zhu, W.J. (2002). *BLEU: A Method for Automatic Evaluation of Machine Translation*. In: Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics. p311-318.

Prates, M.O., Avelar, P., & Lamb, L.C. (2019). *Assessing Gender Bias in Machine Translation-a Case Study with Google Translate*. [arXiv Preprint].

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). *Improving Language Understanding by Generative Pre-Training*. Available from: https://openai.com/research/language-unsupervised [Last accessed on 2025 Jan 20].

Rasouli, F. (2018). Assessment of machine translation output: A comparative study between human and automatic models. *Cihan University-Erbil Scientific Journal*, 2(1), 119-141.

Rasouli, F., Soleimanzadeh, S., & Seyyedi, K. (2024). Acceptability of Google translate machine translation system in translation from English into Kurdish. *Cihan University-Erbil Journal of Humanities and Social Sciences*, 8(1), 7-14.

Reiss, K. (1971). *Möglichkeiten und Grenzen der Übersetzungskritik [Translation Criticism: The Potentials and Limitations]*. Germany: Hueber.

Snover, M., Dorr, B., Schwartz, R., Micciulla, L., & Makhoul, J. (2006). *A Study of Translation Edit Rate with Targeted Human Annotation*. In: Proceedings of the Association for Machine Translation in the Americas. Vol. 200. p223.

Toral, A., Wieling, M., & Way, A. (2018). *How Far is Neural Machine Translation from Human Translation*? In: Proceedings of the Third Conference on Machine Translation. p372-379.

Turian, J.P., Shen, L., & Melamed, I.D. (2003). *Evaluation of Machine Translation and its Evaluation*. In: Proceedings of MT Summit IX. New Orleans. p386-393.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., & Polosukhin, I. (2017). *Attention is All You Need*. In: Advances in Neural Information Processing Systems. p5998-6008.

Appendix

APPENDIX I: Collected data of the study

Text Type	Sentence ID	Source (English)	AI Output (Kurdish)
Informative	1	The library opens at 9 AM every weekday.	کتێبخانەکە ڕۆژی کارکردندا لە کاتژمێر ٩ بەیانە دەکرێت.
Informative	2	The museum houses over 2,000 historical artifacts.	مۆزەخانەکە زیاتر لە ٢٠٠٠ بەرهەمی مێژوویی پێناسا دەکات.
Informative	3	The city’s population reached 500,000 in the last census.	ژمارەی نیشتەجێبوانی شارەکە لە کۆنسێی ڕابردوو ٥٠٠,٠٠٠ کەس بوو.
Informative	4	Rainfall in the region averages 45 inches per year.	میانگین باران لە هەر ساڵدا لە ناوچەکەدا ٤٥ ئینچەیە.
Informative	5	The new software update will launch next Monday.	نوێترین نوێکردنەوەی نەرمەکاڵا دووشەممەی داهاتوو دەکرێت.
Expressive	1	I felt a burst of joy when I saw the sunrise.	کاتێك خۆرهشید بینیم، پڕ لە شادمانی هەستم.
Expressive	2	The gentle melody filled my heart with warmth.	مەلودی نرمی دڵم پڕ کرد بە گەرمی.
Expressive	3	The beauty of this peaceful garden overwhelms me.	من لە جوانیی ئەم باغەی ئاسایە بە هەستێکی عمیق پڕ بووم.
Expressive	4	The unexpected kindness of a stranger touched me deeply.	خیراتی نەزانراوی کەسێکی نەناس دڵم بە شێوەیەکی عمیق گرت.
Expressive	5	I cannot help but smile when I recall that happy moment.	ناتوانم خەندە بکەم کاتێك ئەو کاتی شادە لەبیرم دێت.
Vocative	1	Join us today to support our community cleanup!	ئەمڕۆ بەشدارین بۆ پشتگیری پاککردنەوەی کۆمەڵگایەکەمان!
Vocative	2	Please donate to help build a new neighborhood library.	تکایە بریاری پێ بدەن بۆ یارمەتیدانی دروستکردنی کتێبخانەیەکی نوێ لە محلی.
Vocative	3	Let’s work together for a greener, cleaner city.	بە یەکگرتوو کار بکەین بۆ شارێکی سەوزتر و پاکتر.
Vocative	4	Share your ideas and make a difference in our local park.	ڕایەکانت بەش بکە و گۆڕانکاری لە پارکی محلییەکەمان بکە.
Vocative	5	Act now and sign the petition to improve public transportation.	ئێستا کردار بکە و دەنگی پێبدە بۆ باشترکردنی گواستنەوەی گشتی.

Text Type	Sentence ID	Source (English)	Google Translate (Kurdish)
Informative	1	The library opens at 9 AM every weekday.	کتێبخانە هەموو ڕۆژانی هەفتە کاتژمێر 9ی بەیانی دەکرێتەوە.
Informative	2	The museum houses over 2,000 historical artifacts.	مۆزەخانەکە زیاتر لە 2000 شوێنەواری مێژوویی تێدایە.
Informative	3	The city’s population reached 500,000 in the last census.	ژمارەی دانیشتوانی شارەکە لە دوایین سەرژمێریدا گەیشتە 500 هەزار کەس.
Informative	4	Rainfall in the region averages 45 inches per year.	بارانبارین لە هەرێم بە تێکڕا 45 ئینچ لە ساڵێکدا.
Informative	5	The new software update will launch next Monday.	نوێکردنەوەی نەرمەکاڵا نوێیەکە ڕۆژی دووشەممەی داهاتوو دەخرێتە بازارەوە.
Expressive	1	I felt a burst of joy when I saw the sunrise.	هەستم بە تەقینەوەی خۆشی کرد کاتێک خۆرهەڵاتنم بینی.
Expressive	2	The gentle melody filled my heart with warmth.	ئاوازە نەرمەکە دڵی پڕ کرد لە گەرمی.
Expressive	3	The beauty of this peaceful garden overwhelms me.	جوانی ئەم باخچە ئارامە سەرم دەکێشێت.
Expressive	4	The unexpected kindness of a stranger touched me deeply.	میهرەبانی چاوەڕوان نەکراوی کەسێکی نامۆ بە قوولّی دەستی لێدام.
Expressive	5	I cannot help but smile when I recall that happy moment.	ناتوانم خۆم بپارێزم لە زەردەخەنەکردن کاتێک ئەو ساتەوەختە خۆشەم بیر دێتەوە.
Vocative	1	Join us today to support our community cleanup!	ئەمرۆ لەگەڵمان بن بۆ پشتگیریکردنی پاککردنەوەی کۆمەڵگاکەمان!
Vocative	2	Please donate to help build a new neighborhood library.	تکایە بەخشین بکەن بۆ یارمەتیدان لە دروستکردنی کتێبخانەی نوێی گەڕەک.
Vocative	3	Let’s work together for a greener, cleaner city.	با پێکەوە کار بکەین بۆ شارێکی سەوزتر و پاکتر.
Vocative	4	Share your ideas and make a difference in our local park.	بیرۆکەکانتان بەیەکەوە باس بکەن و جیاوازی دروست بکەن لە پارکە ناوخۆییەکەمان.
Vocative	5	Act now and sign the petition to improve public transportation.	لە ئێستاوە کاربکە و واژۆ لەسەر داواکارییەکە بۆ باشترکردنی هاتوچۆی گشتی.

Text Type	Sentence ID	Source (English)	Human Translation (Kurdish)
Informative	1	The library opens at 9 AM every weekday.	پەرتووکخانە هەموو ڕۆژێک کاتژمێر ٩ بەیانی دەکرێتەوە.
Informative	2	The museum houses over 2,000 historical artifacts.	مۆزەخانەکە زیاتر لە ٢٠٠٠ کاری هونەری لەخۆگرتووە.
Informative	3	The city’s population reached 500,000 in the last census.	لە دوایین سەرژمێریدا ژمارەی دانیشتوانی شارەکە گەیشتوتە ٥٠٠٠٠٠ کەس.
Informative	4	Rainfall in the region averages 45 inches per year.	ڕێژەی مامناوەندی باران بارین لەم هەرێمەدا سالانە ٤٥ ئینچە.
Informative	5	The new software update will launch next Monday.	ئاپدەیتە نوێیەکەی ئەم بەرنامەیە دووشەممەی داهاتوو دەست پێدەکات.
Expressive	1	I felt a burst of joy when I saw the sunrise.	کە دەرکەوتنی خۆرم دیت هەستم بە دڵ خۆشییەکی زۆر کرد.
Expressive	2	The gentle melody filled my heart with warmth.	مۆسیقا ئارامەکە دڵی پڕ کردم لە گەرمی.
Expressive	3	The beauty of this peaceful garden overwhelms me.	جوانی ئەم باخە هێمنە نقوومی خۆی کردم.
Expressive	4	The unexpected kindness of a stranger touched me deeply.	لە دەروونمەوە هەستم بە میهرەبانیە چاوەڕوان نەکراوەکەی ئەم کەسە بیانیە کرد.
Expressive	5	I cannot help but smile when I recall that happy moment.	کە ئەم کاتە دڵخۆشییەم بیردێتەوە ناتوانم زەردەخەنە نەکەم.
Vocative	1	Join us today to support our community cleanup!	بۆ پشتیوانی کردن لە پاکژ کردنەوەی کۆمەڵگاکەمان ئەمڕۆ لەگەڵمان کەوە.
Vocative	2	Please donate to help build a new neighborhood library.	تکایە بۆ دروست کردنی پەرتووکخانەی نوێی گەڕەکەکەمان هاوکاری مادی بکە.
Vocative	3	Let’s work together for a greener, cleaner city.	با پێکەوە کار بکەین بۆ دروستکردنی شارێکی سەوزتر و خاوێنتر.
Vocative	4	Share your ideas and make a difference in our local park.	بۆ دروستکردنی جیاوازی لە پارکی ناوچەکەمان ڕا بۆچوونتانمان لەگەڵ هاوبەش بکەن.
Vocative	5	Act now and sign the petition to improve public transportation.	بۆ باشتر کردنی دۆخی هاتوچۆی گشتی هەر ئێستا هەڵوێست وەرگرەو داواکارییەکە واژۆ بکە.

Appendix II

Analyzed data by python

```
# 1) Prepare your data
# Each list below has 15 items, one per sentence.
# (In your actual script, copy the exact strings from your table.)
AI_outputs = [
# Informative
“کتێبخانەکە هەموو ڕۆژی کارکردندا لە کاتژمێر ٩ بەیانە دەکرێت.”,
“مۆزەخانەکە زیاتر لە ٢٠٠٠ بەرهەمی مێژوویی پێناسا دەکات.”,
“وودربار یێسنۆک ەل ەکەراش یناوبێجەتشین یەرامژ
٥٠٠,٠٠٠ سەک ووب.”,
“میانگین باران لە هەر ساڵدا لە ناوچەکەدا ٤٥ ئینچەیە.”,
“نوێترین نوێکردنەوەی نەرمەکاڵا دووشەممەی داهاتوو دەکرێت.”,
# Expressive
“کاتێك خۆرەشید بینیم، پڕ لە شادمانی هەستم.”,
“مەلودی نرمی دڵم پڕ کرد بە گەرمی.”,
“من لە جوانیی ئەم باغەی ئاسایە بە هەستێکی عمیق پڕ بووم.”,
“خیراتی نەزانراوی کەسێکی نەناس دڵم بە شێوەیەکی عمیق گرت.”,
“ناتوانم خەندە بکەم کاتێك ئەو کاتی شادە لەبیرم دێت.”,
# Vocative
“ئەمڕۆ بەشداربن بۆ پشتگیری پاککردنەوەی کۆمەڵگایەکەمان!”,
“تکایە بڕیاری پێ بدەن بۆ یارمەتیدانی دروستکردنی کتێبخانەیەکی نوێ لە
محلی.”,
“بە یەکگرتوو کار بکەین بۆ شارێکی سەوزتر و پاکتر.”,
“ڕایەکانت بەش بکە و گۆڕانکاری لە پارکی محلییەکەمان بکە.”,
“ئێستا کردار بکە و دەنگی پێبدە بۆ باشترکردنی گواستنەوەی گشتی.”
]
Google_outputs = [
# Informative
“کتێبخانە هەموو ڕۆژانی هەفتە کاتژمێر 9ی بەیانی دەکرێتەوە.”,
“مۆزەخانەکە زیاتر لە 2000 شوێنەواری مێژوویی تێدایە.”,
“ادیرێمژرەس نییاود ەل ەکەراش یناوتشیناد یەرامژ
گەیشتە 500 هەزار کەس.”,
“بارانبارین لە هەرێم بە تێکڕا 45 ئینچ لە ساڵێکدا.”,
“نوێکردنەوەی نەرمەکاڵ ەکەیێون یژۆڕ یەممەشوود یاداهاتوو
دەخرێتە بازاڕەوە.”,
# Expressive
“هەستم بە تەقینەوەی خۆشی کرد کاتێک خۆرهەڵاتنم بینی.”,
“ئاوازە نەرمەکە دڵی پڕ کرد لە گەرمی.”,
“جوانی ئەم باخچە ئارامە سەرم دەکێشێت.”,
“میهرەبانی چاوەڕوان نەکراوی کەسێکی نامۆ بە قووڵی دەستی لێدام.”,
“ناتوانم خۆم بپارێزم لە زەردەخەنەکردن کاتێک ەوی
ساتەوەختە خۆشەم بیر دێتەوە.”,
# Vocative
“ئەمرۆ لەگەڵمان بن بۆ پشتگیریکردنی پاککردنەوەی کۆمەڵگاکەمان!”,
“تکایە بەخشین بکەن بۆ یارمەتیدان لە دروستکردنی کتێبخانەی نوێی
گەڕەک.”,
“با پێکەوە کار بکەین بۆ شارێکی سەوزتر و پاکتر.”,
“بیرۆکەکانتان بەیەکەوە باس بکەن و جیاوازی دروست بکەن لە پارکە
ناوخۆییەکەمان.”,
“لە ئێستاوە کاربکە و واژۆ لەسەر داواکارییەکە بۆ باشترکردنی هاتوچۆی
گشتی.”
]
Human_references = [
# Informative
“پەرتووکخانە هەموو ڕۆژێەک کاتژمێر ٩ بەیانی دەکرێتەوە.”,
“مۆزەخانەکە زیاتر لە ٢٠٠٠ کاری هونەری لەخۆگرتووە.”,
“ەکەراش یناوتشیناد یەرامژ ادیرێمژرەس نییاود ەل
گەیشتۆتە ٥٠٠٠٠٠ کەس.”,
“ڕێژەی مامناوەندی باران بارین لەم هەرێمەدا سالانە ٤٥ ئینچە.”,
“ئاپدەیتە نوێیەکەی ئەم بەرنامەیە دووشەمەی داهاتوو دەست پێدەکات.”,
# Expressive
“کە دەرکەوتنی خۆرم دیت هەستم بە دڵ خۆشێیەکی زۆر کرد.”,
“مۆسیقا ئارامەکە دڵی پڕ کردم لە گەرمی.”,
“جوانی ئەم باخە هێمنە نقوومی خۆی کردم.”,
“لە دەروونمەوە هەستم بە میهرەبانیە چاوەڕوان نەکراوەکەی ئەم کەسە
بیانیە کرد.”,
“کە ئەم کاتە دڵخۆشییەم بیردێتەوە ناتوانم زەردەخەنە نەکەم.”,
# Vocative
“بۆ پشتیوانی کردن لە پاکژ کردنەوەی کۆمەلگاکەمان ئەمرۆ لەگەڵمان
کەوە.”,
“تکایە بۆ درووست کردنی پەرتووکخانەی نوێی گەرەکەکەمان هاوکاری
مادی بکە.”,
“با پێکەوە کار بکەین بۆ درووستکردنی شارێکی سەوزتر و خاوێنتر.”,
“بۆ دروستکردنی جیاوازی لە پارکی ناوچەکەمان ڕاو
بۆچوونەکانمان لەگەڵ هاوبەش بکەن.”,
“بۆ باشترکردنی دۆخی هاتووچۆی گشتی هەر ئێستا
ئەلیستی واژۆ داواکارییەکە بکە.”
]
# 2) Compute BLEU
# A. Corpus-level BLEU for AI outputs
bleu_ai = sacrebleu.corpus_bleu(AI_outputs, [Human_references])
print(“AI BLEU Score (Corpus-Level):”, bleu_ai.score)
# B. Corpus-level BLEU for Google outputs
bleu_google = sacrebleu.corpus_bleu(Google_outputs, [Human_references])
print(“Google Translate BLEU Score (Corpus-Level):”, bleu_google.score)
# 3) Compute TER
# We’ll do sentence-level TER and then average.
# The pyter.ter function expects tokenized text (list of tokens).
# For simplicity, let’s do a quick split on whitespace.
def tokenize_kurdish(text):
return text.strip().split()
ter_ai_list = [ ]
ter_google_list = [ ]
for ai_sent, google_sent, ref_sent in zip(AI_outputs, Google_outputs, Human_references):
# Tokenize
ai_tokens = tokenize_kurdish(ai_sent)
google_tokens = tokenize_kurdish(google_sent)
ref_tokens = tokenize_kurdish(ref_sent)
# Compute TER
ai_ter = pyter.ter(ai_tokens, ref_tokens)
google_ter = pyter.ter(google_tokens, ref_tokens)
ter_ai_list.append(ai_ter)
ter_google_list.append(google_ter)
# Average TER
avg_ter_ai = sum(ter_ai_list)/len(ter_ai_list) * 100 # convert to %
avg_ter_google = sum(ter_google_list)/len(ter_google_list) * 100
print(“AI TER (Average, %):”, avg_ter_ai)
print(“Google Translate TER (Average, %):”, avg_ter_goo gle)
```