Proceedings of the 4th International Conference on Literature, Language, and Culture Development
DOI: 10.54254/2753-7064/2025.20639

Evaluating ChatGPT's Chinese-English Translation Quality of Tender Documents: A Research of MTPE with MQM Scoring Models

Lin Zhou[1,a,*]

[1]*School of Foreign Languages and Literature, Beijing Normal University, 19 Xinjiekouwai Street, Beijing, China*
a. zladd961014@163.com
**corresponding author*

Abstract: Since Google introduced the Transformer model into natural language processing (NLP) in 2017, AI-aided translation has rapidly advanced. At the same time, translation is evolving from a solitary endeavor into a cooperative activity between human translators and machine translation systems, epitomized by the emergency of platforms with the Machine Translation Post Editing (MTPE) function. The advent of new translation modes also leads to increased research evaluating the effectiveness and quality of machine translation, for example, studies on the translation quality under the Multidimensional Quality Metrics (MQM) error typology framework. Involving AI-based translators and MTPE in their translation enables human translators to prepare the engineering documents efficiently. However, researchers notice that it is difficult for most machine translators to figure out the semantic and cultural differences in the source language and generate coherent structural translation in the target language. This research opens up ChatGPT's application in tender document translation under the MQM framework, hoping to cast light on assessment on ChatGPT's translation quality, identification of ChatGPT's errors in translating such documents and suggestions on human translators' performance throughout MTPE.

Keywords: Machine Translation, Translation Error Typology, MTPE, MQM

1. Introduction

As artificial intelligence (AI) technology continues to advance, the application of natural language processing (NLP) as a bridge for human-computer interaction becomes increasingly significant. A famous innovation in NLP is the Transformer model, which has greatly impacted the translation industry, reshaping the landscape and prompting translators to adapt to the era of AI. The Transformer model processes data with attention mechanisms, which is performed in parallel, simultaneously capturing relationships between all elements in a sequence [1]. This parallelism allows the Transformer to train much faster and handle long-range dependencies efficiently. The Transformers model is particularly effective in tasks such as machine translation, document summarization, sentiment analysis, and natural language understanding. They offer interpretability through attention mechanisms, providing insights into which parts of the input sequence are crucial for predictions.

DOI: 10.54254/2753-7064/2025.20639

However, transformers are computationally expensive and require substantial data for practical training.

In response to AI's disruptive influence, translators are utilizing new methodologies, leveraging AI as a tool to assist their translation. Translation is evolving from an individual endeavor into a symbiotic relationship between human translators and machine translation systems, one instance of which is the emerging practice of Machine Translation Post-Editing (MTPE). It has been applied in various industries, including international trade, health, and legal settings, facilitating communication and overcoming communication barriers among people from different cultural and linguistic backgrounds [2]. However, both human and AI-based translations may still contain errors [3], and the performance of machine translation in some domains is far from satisfactory [4].

Machine translation, even though the most popular one, ChatGPT, is not risk-free. Different prompts will cause significant distinctions between the translations regarding faithfulness, fluency, language style, and translation style [5]. Therefore, enhancing ChatGPT's translation quality through prompt engineering is a plausible direction [6]. Particularly, prompts that capture both cross-cultural subtleties and linguistic context can improve the precision and consistency of translation [7]. If their errors are neglected, it can bring about severe consequences such as default, economic, reputational, or health-related outcomes for the translators themselves, clients, and third parties. As globalization continues, more and more Chinese engineering companies choose to develop their business globally and implement overseas engineering projects. This trendy has accelerated since China proposed the Belt and Road Initiative (BRI). In order to attend a tender, contractors are required to submit qualification documents, most original texts of which are in Chinese and shall be translated into a foreign language accurately, correctly, and concisely. Hence, translators play a crucial role in this process. These supporting documents often include registration certificates, letters of award, drawings, contracts, etc. In order to mitigate risks, it is essential that translators understand the text type and features, machine translation's mechanisms and error types, and strengthen their abilities to identify and quantify translation errors.

Therefore, this study aims to study error types of machine translation in the engineering industry and investigate the applicability of machine translation in this realm. By doing this, it is hoped that this study could contribute to helping translators optimize the use of machine translation effectively, especially for those working for Chinese engineering companies overseas.

2. Literature review

2.1. Machine Translation Post-Editing of Construction Engineering Texts

The translation of construction engineering texts includes, but is not limited to, design and construction drawings, tender documents, technical and commercial proposals, experience certificates, and other construction-related documents. While establishing the construction corpus and studying the lexical frequency, Jinghui Ren researched the features of construction engineering documents and found that construction engineering vocabulary belongs to limited semantic groups, including attributes, meteorology, civil planning, direction and measurement, environment, structures, and applications [8]. Words with a high frequency are usually terms that show the professional features of those documents. Construction engineering documents belong to technical texts, which are usually written in an objective and neutral tone. The text often involves many abbreviations used as subjects, such as RFT (Request for Tender) and PLC (Programmable Logic Controller), and its content is featured by complex logic and nested compound sentences, omitted non-finite verb modifiers, and other similar structures [9]. Due to cultural differences, the use of vocabulary in tender documents varies among countries. For example, in Chinese, the word that refers to the institution/company which initiates the project is "建设单位." When translated into English, "建设"

DOI: 10.54254/2753-7064/2025.20639

is often translated as "construction". In Fiji, most projects use "Employer" as the institution that owns the project, and in some European countries, they also use "Principal."

AI-driven machine translation systems have seen significant improvements in translation quality, which is an output of advancements in deep learning and neural networks [10]. Neural machine translation (NMT) models trained on extensive parallel corpora have become the leading method in the field of machine translation [11]. After reviewing recent studies, Shanshan Wang and Xiaohui Wang argued that these models excel at capturing linguistic patterns of languages, as well as semantic and syntactic nuances, incredibly enhancing the accuracy and fluency of translation [12]. A notable AI-based language model that has gained attention recently is ChatGPT, developed by OpenAI. ChatGPT, powered by a large language model (LLM), works as an AI-based chat that tracks previous prompts and responses, correcting and adapting subsequent answers given the sequence of inputs and outputs. Using ChatGPT, terminology can be efficiently translated. The computer converts linguistic symbols in the text into vectors and processes billions of tokens. Adjusting parameters based on the prediction of the next token according to the context enables MT to generate text that is grammatically correct and logically coherent [13].

Since the machine translation algorithm is not perfect for analyzing the intricacies and subtleties of language, it calls for a human translator's intervention. Against this background, MTPE appears to fix errors or inappropriate content, therefore enhancing the quality of machine translation to an acceptable level [14]. Post-editing is a cognitively demanding task that involves reading the source text, making corrections to the machine-translated output, and creating the final target text [15].

2.2. Error Typology of Translation and MQM

Analyzing errors in machine translation involves many steps, such as identification, categorizing, and understanding of errors incurred by MT algorithms during translation. BLEU (Bilingual Evaluation Understudy) and TER (Translation Error Rate) are two commonly used metrics for assessing translation quality [16]. Regarding machine translation errors, IBM once conducted research on Russian-English translation and identified some common errors made by MT systems, such as transliterated terms, varied meanings, ambiguities, restructured word order, as well as various insertions and adjustments [17]. Popović also proposes a list of common MT errors from a linguistic perspective. For example, at the semantic level, one error is meaning deviation caused by incorrect disambiguation or mistranslation of multi-word expressions [18]. Lommel and several scholars established the Multidimensional Quality Metrics (MQM) framework for analytic Translation Quality Evaluation (TQE) [3]. The advent of MQM provided normalized and formalized methods to evaluate translation quality, which is based chiefly on sample analysis. There are seven high-level dimensions and subordinate error subtypes at various hierarchical levels in the MQM error typology. The seven evaluation angles include terminology, accuracy, linguistic conventions, style, locale conventions, audience appropriateness, design, and markup.

3. Methodology

3.1. Research Material

Construction engineering documents analyzed in this research were selected from tender documents submitted by a subsidiary (hereinafter referred to as S) of a Chinese state-owned construction company (hereinafter referred to as C) during a competitive bidding process. S has been established in Fiji for 15 years and has implemented more than 40 projects across the South Pacific area, each exceeding a million Fijian dollars. As C's regional representative, S usually leverages C's experience and certifications to strengthen its proposals. In standardized bidding activities in Fiji, interested bidders are required to submit evidence of similar experience, such as completed contracts, drawings,

DOI: 10.54254/2753-7064/2025.20639

letters of award, completion certificates, pre-sale certificates, and technical proposals, to meet qualification criteria. The materials submitted by S are inaccessible to external parties, and additionally, employees of S and C rarely use AI-assisted translation tools to translate such experience documents; therefore, it can be assured that the materials used for this study were not included in ChatGPT's pre-training data before.

To ensure the randomness and effectiveness of this study, the author selected three documents from one project conducted by C in China within the past five years and recently used as part of S's experience documentation for a tender in Fiji. The selected documents include a Letter of Award (479 Chinese characters), a Construction Contract (3,057 Chinese characters), and a Completion Record (1,195 Chinese characters). These three documents were translated from Chinese to English via MTPE, and after their submission, S successfully passed the qualification test.

3.2. Translation Platform and Translator

This study took Twinslator (https://transpace.iol8.com/personal) as the objective translation platform; it is an AI system built on a self-developed multilingual large language model, aiming to create a "digital twin" of human translators through deep interaction. Registered translators on the platform can access different machine translation tools, such as ChatGPT, Google Translator, and DeepL Translator. Human translators can leverage these tools for pre-processing translations and then apply MTPE. Since ChatGPT is a generative transformer, its translation results will vary due to different prompts. Using one translation platform can avoid the interference of different prompts. Additionally, Twinslator provides detailed metrics, such as total word count, revised and unrevised word counts, and the revision percentage, ensuring clarity and traceability in the translation process. In this research, all documents were initially translated using ChatGPT within the Twinslator platform and subsequently refined through human post-editing.

The author of this article has passed the C-E NAATI examination, the national standards and certifying authority for translators and interpreters in Australia, and has been working for S company for more than two years, mainly in charge of translation of tender documents. The documents translated by ChatGPT in the Twinslator platform were edited by the author and then evaluated with the MQM2.0 Evaluation Scorecard.

3.3. Translation Quality Evaluation and Participant

MQM scoring models evaluate translation quality from the seven high-level dimensions, including terminology, accuracy, linguistic conventions, style, locale conventions, audience appropriateness, design and markup. Under most high-level dimensions, there are several subcategories to identify errors accurately. These dimensions have a specific and defined meaning, which should be distinguished from their general and common definitions in a dictionary. Errors are divided into neutral error, minor error, major error and critical error, depending on the extent to which they hinder correct and thorough comprehension of the translation. The corresponding Error Type Weights (ETWs) are neutral error (0 point), minor error (1 point), major error (5 points) and critical error (25 points). Unlike BLUE and other automatic evaluation metrics, MQM does not require reference translations, allowing greater flexibility in assessing translation quality. Absolute Penalty Total (APT) is the sum of all Error Type Penalty Totals. It can be calculated by the formula:

$$\sum_{i,j}(\text{Error Countij} \times \text{Severity Multiplierj} \times \text{Error Type Weighti}) \qquad (1)$$

Where: i = index for Error Types; j = index for Severity Level.

There are other key parameters in the MQM framework. Per-Word Penalty Total (PWPT) is determined by dividing the APT by the Evaluation Word Count (EWC). The EWC of this research is

DOI: 10.54254/2753-7064/2025.20639

the total word count in each translation document. The Normed Penalty Total (NPT), representing the Per-Word Error Penalty (PWEP) total relative to the Reference Word Count (RWC), is obtained by multiplying the PWPT by the RWC, representing the PWEP total relative to the RWC, which is always set up as 1000. A satisfactory translation should pass the threshold set in Passing Threshold Calibrated QS. Otherwise, the Quality Rating could be failed. In the General Comment column, a detailed comment regarding the translation quality will be provided to summarize the result.

Other than assessing from the seven dimensions, modifying word counts and percentages will also be analyzed. The result was recorded in Table 1. As shown in Table 1, the percentages of modified words in original text in three documents all exceeded 55%, highlighting the need for further improvement in the quality of the machine translation used in this study.

Table 1: Modifying Word Count Calculation Result.

Document	**Translator**	**Stage**	**Original Total Word Count**	**Translation Total Word Count**	**Modified Translation Word Count**	**Unmodified Translation Word Count**	**Percentage of Modified Words in Original Text**
Construction Contract	Jolin	MTPE	3057	1905	1207	698	63.36%
Letter of Award	Jolin	MTPE	479	275	175	100	63.64%
Completion Record	Jolin	MTPE	1195	646	362	286	55.86%
In Total			4731	2828	1744	1084	--

4. Results and Discussion

Scoring of the translation documents. The machine translation of the selected documents was analyzed based on the seven angles. After the human translator's assessment, the scoring results for the three documents are presented in Tables 2, 3, and 4. In general, the ChatGPT translations achieved scores exceeding the threshold of 90, standing for overall satisfactory performance in translating Chinese-English engineering tender documents.

Table 2: Scoring of the translation document of Letter of Award.

Evaluation Results	**Quality Rating (QR)**	**Raw Quality Score (RQS)**	**Absolute Penalty Total (APT)**	**Per-Word Penalty Total (PWPT)**	**Normed Penalty Total (NPT)**	**Calibrated Quality Score (CQS)**
	PASS	90.9	25	0.091	90.91	90.91

Table 3: Scoring of the translation document of Construction Contract.

Evaluation Results	**Quality Rating (QR)**	**Raw Quality Score (RQS)**	**Absolute Penalty Total (APT)**	**Per-Word Penalty Total (PWPT)**	**Normed Penalty Total (NPT)**	**Calibrated Quality Score (CQS)**
	PASS	96.0	84	0.044	44	92.00

DOI: 10.54254/2753-7064/2025.20639

Table 4: Scoring of the translation document of Completion Record.

Evaluation Results	Quality Rating (QR)	Raw Quality Score (RQS)	Absolute Penalty Total (APT)	Per-Word Penalty Total (PWPT)	Normed Penalty Total (NPT)	Calibrated Quality Score (CQS)
	PASS	96.0	26	0.040	40	95.99

It is noted that there are two prominent error types commonly appearing in the AI-assisted translation: terminology and locale conventions (e.g., name format, address format, number format). As shown in Table 5, in the translation of the Letter of Award, there are 11 translation errors of terminology; in the Construction Contract, there are 16; and in the Completion Record, there are 9. Terminology translation errors frequently appear when translating a company's name, contractor/employer's name, and engineering terms.

Table 5: Sum of Translation Errors of Terminology.

Document	Neutral Errors Count	Minor Errors Count	Critical Errors Count	Sum
Letter of Acceptance	**2**	**5**	**4**	**11**
Construction Contract		**9**	**7**	**16**
Completion Record		**7**	**2**	**9**

The subsequent examples are listed to promote a further understanding of machine translation's performance.

Table 6: Comparison of ChatGPT Translation and Human Edited Translation (Example 1).

Original text	ChatGPT translation	Human edited translation
建设单位名称 (或代建单位): 佛山市汇之源城北污水处理有限公司	Construction Unit Name (or Proxy Construction Unit): Foshan Huizhiyuan North City Sewage Treatment Co., Ltd.	Name of Employer (or Employer's Agent): Foshan Huizhiyuan Chengbei Sewage Treatment Co., Ltd.

In Table 6, the Chinese word "建设" is usually translated as "construction" or "construct" in English. However, in this context, it refers to the person or institution which initiates the project. According to the General Conditions of the FIDIC Red Book Conditions of Contract for Construction, "Employer" means the person named as the employer in the Award of Tender and the legal successors in the title to this person. "Contractor" means the person(s) named as the contractor in the Letter of Tender accepted by the employer and the legal successors in title to this person(s). Given the definition, one project is given by the employer to the contractor. Therefore, "建设单位" here should be Employer rather than Contractor. The wrong translation would misguide readers, resulting in disqualification for the tender, so it can be regarded as a critical error.

Besides, the ChatGPT translation of the employer's name following the original order is also problematic. It adopted literal translation, translating "城北" into "North City." However, after the translator visited Qichacha, the leading corporate information provider in China, and searched online,

DOI: 10.54254/2753-7064/2025.20639

it was noticed that only "Foshan Huizhiyuan Chengbei Sewage Treatment Co., Ltd." was used in a patent as the company's English name. Therefore, the translator revised it as shown in the patent.

Table 7: Comparison of ChatGPT Translation and Human Edited Translation (Example 2).

Original text	**ChatGPT translation**	**Human edited translation**
总建筑面积:53096.88m² ,其中地下室建筑面积:47023.79m² ,建筑层数:地上 3 层,地下 2 层,计容面积:5998.3m² ,基底面积:3225.96m²	Total construction area: 53096.88m², including basement construction area: 47023.79m², number of floors: 3 floors above ground, 2 floors underground, calculated capacity area: 5998.3m², base area: 3225.96m²	Total construction area: 53096.88m², including basement construction area: 47023.79m², number of floors: 3 floors above ground, 2 floors underground, calculated capacity area: 5998.3m², base area: 3225.96m²

In Table 7, "总建筑面积" and "计容面积" are two engineering terms calling for revision. "总建筑面积" means the sum of the construction areas of all floors, including the total usable area of all units and the total shared construction area. It should be translated as "Construction Gross Area." The ChatGPT translation of the term is "total construction area," which is not a professional translation, although it can be understood. When calculating the floor area ratio (FAR), not all of the total construction area of a community or plot is necessarily included. The portion included in the FAR calculation is called the building area (construction area) used in the FAR calculation. Namely, the translation of "计容面积". The ChatGPT translation of this term is the same as the former term's translation, which is not the formal description of engineering projects.

Besides translation errors of terminology, errors in locale conventions are also notable. The following examples demonstrate how ChatGPT translation fails to identify some elements in the original language, especially when translating names and addresses.

Table 8: Comparison of ChatGPT Translation and Human Edited Translation (Example 3).

Original text	**ChatGPT translation**	**Human edited translation**
承包人项目经理(项目总负责人):刘彦华 项目施工负责人:孙新安 项目设计负责人:王阳 项目试运行负责人:王玉梅	Contractor's project manager (project general leader): Liu Yanhua Project construction leader: Sun Xin'an Project design leader: Wang Yang Project trial operation leader: Wang Yumei	Contractor's project manager (project leader): Yanhua Liu Project construction manager: Xin'an Sun Project design manager: Yang Wang Project trial operation manager: Yumei Wang.

When translating a Chinese name, the machine translation puts the family name before the given name. This was regarded as a minor issue in this analysis because it wouldn't impede any understanding of the contractor's experience. However, in English, it is conventional to write the given name first, so the translator adjusted all the translations following the Western name convention.

In the example as shown in the Table 8, the translation of "项目总负责人" is also worth analyzing from the perspective of linguistic conventions. In this context, both "项目经理" and "项目总负责人" mean the person who is in charge of the project. The Chinese character "总" means "general," but it is not necessary to be translated as such here because, in engineering, a project manager or project

DOI: 10.54254/2753-7064/2025.20639

leader is the person who will be responsible for the whole project. There is no need to emphasize the meaning of "general." Moreover, using "general leader" in the engineering industry is weird. Therefore, the translator deleted the word "general" in the final translation. In Fiji, people who manage construction, design, or trial operations are managers or engineers, such as construction manager/engineer and design manager/engineer. Therefore, the translator also revised the word "leader" to "manager".

Another error, though not commonly observed in the translation of all three documents, is still critical, the error type of which is identified as locale conventions (number format). In the selected construction contract, there were many numbers. In the Chinese contract, the contract amount was presented in "大写" and "小写." "大写," a unique and complicated writing method of Chinese characters, is mostly used in accounts and documents to prevent the purposed alteration of price. In contrast, "小写" follows the standard Arabic numeral system. Ensuring consistency between these two formats is essential for accurately documenting contract amounts. Table 9 shows how ChatGPT made mistakes in translating the two Chinese numeral writing styles.

Table 9: Comparison of ChatGPT Translation and Human Edited Translation (Example 4).

Original text	勘察费合同价:人民币(大写):叁佰柒拾肆万贰仟贰佰玖拾元 (小写): ¥3,742,290.00 元 勘察费综合单价:人民币(大写):每米壹佰玖拾叁元肆角 (小写): ¥193.40 元/米;
ChatGPT translation	Survey fee contract price: RMB (in capitals): three million seven hundred and forty-two thousand two hundred and ninety yuan (in number) ¥3,742,290.00 Comprehensive unit price of survey fee: RMB (in capitals): one hundred and ninety-three yuan and forty cents per meter (in number) ¥193.40/meter;
Human edited translation	Contract price for investigation: RMB (in words): Three Million Seven Hundred Forty-Two Thousand Two Hundred Ninety Yuan, (in figures): ¥3,742,290.00, Comprehensive unit price of investigation fee: RMB (in words): One Hundred Ninety-Three Yuan and Forty Cents per meter (in figures): ¥193.40 per meter;

In the ChatGPT translation, the machine translator failed to identify the correct meaning of "大写" in this context. Therefore, it translated the Chinese words as "in capitals." In other cases, machine translators could not recognize the complicated Chinese numeral systems, so they missed the translation. The missing texts in this study were all related to the Chinese numeral system. Besides, the figure translated by machine translation was not perfect, so human translators were requested to edit it.

5. ChatGPT's Potential in Tender Document Translation and MTPE Strategies Development

This study integrated ChatGPT into engineering document translation, demonstrating frequent MT errors. Overall, the two error types that commonly appeared are related to terminology and locale conventions. Since Twinslator is the translation platform in this study and the translation result of ChatGPT varies with the distinct prompts, this study is confined to the Twinslator platform.

DOI: 10.54254/2753-7064/2025.20639

Despite the errors in ChatGPT's translation of tender documents, ChatGPT's overall performance in translating Chinese engineering documents into English is more efficient than that of traditional human translation modes. ChatGPT translation is a robust product of mass data training, so it can quickly analyze the source language's grammar, semantic patterns, etc., and imitate a human-like translation style. However, this study notices that human translators are still essential for ChatGPT translation to guarantee the information's essence and coherence. In engineering document translation, human translators could focus on editing errors such as terminology and locale conventions, especially unifying the name and address formats and correcting number translations.

Different language pairs usually apply distinct language rules. Therefore, it is crucial to examine each language pair to identify MT errors and tailor-make MTPE strategies accordingly [19]. Within the same language pair, the text type associated with the industry and realm in which they are classified should be considered to develop unique sub-strategies. After recognizing the similarities of those sub-strategies, a comprehensive MTPE strategy system could be established to guide the translation for both learners and professionals, particularly assisting them in identifying the MT errors.

6. Conclusion

This study was conducted to investigate how ChatGPT performs in translating engineering tender documents and identify errors in machine translation to assist human translators' post-editing under the MQM scoring model, corroborating that ChatGPT is powerful enough to produce human-like translation. The translation quality of three targeted documents passed the threshold. The findings also suggested that ChatGPT translation still needs to be improved in engineering document translation, calling for human translators' post-editing efforts to review, amend, and localize the translation. It has been noticed that the error types were concentrated on terminologies and local conventions. Most of them were minor errors, representing that the translation could be understandable during the contexts but not like jargon. However, there were frequent errors in number translation, revealing that ChatGPT is partly problematic in identifying the format of numbers and transforming the numbers between the traditional and current formats. More data could be gathered to improve the model from the areas analyzed in this study.

However, it is significant to point out that the research results should be interpreted with caution due to certain limitations. First of all, this limitation may give rise to concerns regarding the applicability and generalizability of the results. This study was based on the machine translation model used by the Twinslator platform, and the research objects were three engineering documents. Therefore, this study might promote research on machine translation, especially on ChatGPT translation of engineering documents. However, a large quantity of sources would enhance the research values. Secondly, the local conventions referred to in this thesis were Fijian-style English and in construction engineering scenarios since the tender documents were submitted for bidding activities in Fiji. The usage of English also varied in different countries, regions, and scenarios; therefore, this research could be regarded as a valuable practice in similar research in the South Pacific region.

Taking all things into account, it is obvious that machine translation, represented by ChatGPT translation, has accelerated human translation, given its strong data analysis and processing capabilities. However, it needs amendments in terminology translation, format adjustment, and language localization, which shows that machine translation is a crucial tool in translation rather than a 100-percent replacement of human translators. Therefore, it is advised that translators should realize that the application of machine translation will enhance their working efficiency. They should also be open and positive to get more know-how about machine translation, including their error detection ability, editing and localization strategies, and necessary MT knowledge. Nevertheless, this study

DOI: 10.54254/2753-7064/2025.20639

looks forward to inspiring human translators and MT engineers to delve into the research and improve the MT quality together in the near future.

References

[1] *Soydaner, Derya. "Attention mechanism in neural networks: where it comes and where it goes." Neural Computing and Applications 34.16 (2022): 13371–85.*

[2] *Vieira, Lucas Nunes, Minako O'Hagan, and Carol O'Sullivan. "Understanding the societal impacts of machine translation: a critical review of the literature on medical and legal use cases." Information, Communication & Society 24.11 (2021): 1515–32.*

[3] *Lommel, Arle, et al. "The Multi-Range Theory of Translation Quality Measurement: MQM scoring models and Statistical Quality Control." arXiv preprint arXiv:2405.16969, 2024.*

[4] *Yang, Yanxia, et al. "Performance and perception: machine translation post-editing in Chinese-English news translation by novice translators." Humanities and Social Sciences Communications 10.1 (2023): 1–8.*

[5] *Gao, Yuan, Ruili Wang, and Feng Hou. "How to design translation prompts for ChatGPT: An empirical study." arXiv preprint arXiv:2304.02182, 2023.*

[6] *Gao, Ruiyao, et al. "Machine translation of Chinese classical poetry: a comparison among ChatGPT, Google Translate, and DeepL Translator." Humanities and Social Sciences Communications 11.1 (2024): 1–10.*

[7] *Liu, Huey-Ing, and Wei-Lin Chen. "X-transformer: a machine translation model enhanced by the self-attention mechanism." Applied Sciences 12.9 (2022): 4502.*

[8] *Ren, Jinghui. "Mining Lexical Co-occurrence Features Based Research On Construction Engineering English." IOP Conference Series: Earth and Environmental Science. Vol. 267. No. 6. IOP Publishing, 2019.*

[9] *Gou, Miao. "Syntactic Features and Translation Strategies of English for Construction: A Review of English for Construction." Building Structure, vol. 51, no. 14, 2021, pp. Appendix 4–5 (in Chinese).*

[10] *Wang, Haifeng, et al. "Progress in machine translation." Engineering 18 (2022): 143–53.*

[11] *Bahdanau, Dzmitry. "Neural machine translation by jointly learning to align and translate." arXiv preprint arXiv:1409.0473, 2014.*

[12] *Wang, Shanshan, and Xiaohui Wang. "Cross-Language Translation and Comparative Study of English Literature Using Machine Translation Algorithms." Journal of Electrical Systems 20.6s (2024): 2221–29.*

[13] *Feng, Zhiwei. "Term Translation in Era of Large Language Models." China Terminology, vol. 26, no. 3, 2024, p. 93 (in Chinese).*

[14] *Mossop, Brian. "Lynne Bowker and Jairo Buitrago Ciro. Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community. Bingley, Emerald Publishing, 2019, 111 p." TTR: traduction, terminologie, rédaction 32.2 (2019): 250–55.*

[15] *Yang, Yanxia, and Xiangling Wang. "Predicting student translators' performance in machine translation post-editing: interplay of self-regulation, critical thinking, and motivation." Interactive Learning Environments 31.1 (2023): 340–54.*

[16] *Pang, Yaxin, and Xiaoyan Wang. "A Study on the Translation Quality of ChatGPT in the Context of Large Language Model." Modern English, vol. 22, 2023 (in Chinese).*

[17] *Al-Khalifa, Hend, Khaloud Al-Khalefah, and Hesham Haroon. "Error Analysis of Pretrained Language Models (PLMs) in English-to-Arabic Machine Translation." Human-Centric Intelligent Systems, 2024: 1–14.*

[18] *Popović, Maja. "Error classification and analysis for machine translation quality assessment." Translation Quality Assessment: From Principles to Practice, 2018: 129–58.*

[19] *Tsuji, Kayo. "Identifying MT Errors for Higher-Quality Second Language Text." International Journal of Translation, Interpretation, and Applied Linguistics (IJTIAL) 6.1 (2024): 1–12.*