Proceeding of ICSPHS 2026 Symposium: Critical Perspectives on Global Education and Psychological Development
DOI: 10.54254/2753-7048/2026.HT31097

Analysis of Translation Errors in Chinese-Russian Neural Machine Translation

Xiao Sun

Institute of Foreign Languages, Suzhou University, Suzhou, China
2381721397@qq.com

Abstract. With the continuous advancement of Sino-Russian scientific and technological cooperation, the demand for translating specialized texts such as technical literature and reports has surged, making it difficult for traditional human translation to meet the requirements for efficiency and scale. The current accuracy of Chinese-Russian Neural Machine Translation (NMT) systems in handling specialized scientific and technical texts remains suboptimal, exhibiting significant issues particularly in terminology, consistency, and structural conversion. The innovations of this paper are twofold: first, the construction of a Chinese-Russian parallel corpus covering domains such as mechanical engineering and computer science, addressing the limitations of general-domain corpora; second, the adoption of a multi-system comparative framework to conduct quantitative evaluation and qualitative analysis of the Chinese-to-Russian translation performance of three mainstream NMT systems. This study constructs a parallel corpus encompassing both scientific/technical and general domains, evaluates the Russian-to-Chinese translation performance of three mainstream NMT systems, and performs error type statistics and analysis based on BLEU scores and manual annotation. The results show that the average BLEU score for scientific and technical texts is 0.19 points lower than that for general texts, with terminology mistranslation, word order errors, and semantic confusion identified as the primary error types.

Keywords: Neural Machine Translation (NMT), Translation Quality Evaluation, Scientific and Technical Texts.

1. Introduction

With the deepening of the Belt and Road Initiative, Sino-Russian scientific, technological, and economic cooperation has experienced explosive growth. According to 2024 data from China's General Administration of Customs, the trade volume of high-tech products between China and Russia exceeded 280 billion yuan, a year-on-year increase of 32%. The demand for the mutual translation of technical documents in fields such as aerospace, energy, and intelligent manufacturing has surged by an average of 45% annually. As the core carrier of technology transfer, the accuracy of scientific and technical text translation directly impacts cooperation efficiency. A telling example occurred in 2023: a mistranslation of a key term in a Sino-Russian nuclear power plant cooperation project led to delays in equipment commissioning, resulting in economic losses of over 120 million

DOI: 10.54254/2753-7048/2026.HT31097

yuan. This incident underscores the strategic significance of high-quality Chinese-Russian scientific and technical text translation.

Neural Machine Translation (NMT) has become the mainstream technology for Chinese-Russian translation. While commercial systems like Google Translate and Baidu Translate achieve BLEU scores of 26-29 in general text scenarios, they still face severe challenges in translating scientific and technical texts. This is attributable to two main factors: Firstly, Chinese and Russian belong to different language families, and the scientific and technical domain contains numerous specialized abbreviations and complex long-sentence structures, leading to frequent semantic loss in existing NMT models. Secondly, the scale of publicly available Chinese-Russian parallel corpora for scientific and technical fields is insufficient; as of 2024, the volume of domain-specific annotated corpora is only one-eighth that of Chinese-English scientific corpora, consequently limiting the models' generalization capability.

Existing research predominantly focuses on optimizing NMT models for general texts or improving translation effectiveness within a single domain. Significant gaps remain in the construction of unified multi-domain scientific corpora, the quantitative evaluation of systematic errors, and the analysis of error migration patterns across different scenarios. Although some studies mention terminology translation errors, they fail to establish a comprehensive error taxonomy covering syntactic, semantic, and pragmatic levels, thereby hindering targeted optimization of NMT systems.

2. Literature review

In recent years, scholars [1-7] have conducted extensive research on neural machine translation. Since its emergence in 2014, Neural Machine Translation (NMT) technology, leveraging the encoder-decoder architecture of neural network models, has achieved significant improvements in translation fluency. It has gradually replaced Statistical Machine Translation (SMT) as the mainstream paradigm in the industry. Companies such as Google, Baidu, and Tencent have successively launched NMT systems, driving the technology towards practical application and commercialization, even claiming to achieve "human parity" in the domain of general news. However, existing studies generally indicate [8,9] that NMT still exhibits significant limitations in vertical domains. Post-Editing (PE), as a crucial step for enhancing translation quality, has consequently become a key research focus for both academia and the industry.

From the perspective of technological development, machine translation has undergone three paradigm shifts: from rule-based, to statistical, and finally to neural network-based approaches. Li Xiaoqiao et al. [7] revealed that early rule-based systems relied on the manual construction of linguistic rules, resulting in insufficient flexibility. Statistical methods, which trained models on parallel corpora, improved translation accuracy but suffered from issues such as rigid sentence structures. Guo Wanghao et al. [10] pointed out that NMT subsequently achieved a breakthrough in handling contextual coherence through high-dimensional vector representations of linguistic semantics. Systems from companies like Baidu and Google can achieve BLEU scores above 27 on general corpus tests. However, the specific characteristics of vertical domains pose greater challenges for translation systems.

The core challenges of NMT in vertical domains lie in its insufficient capability for handling specialized terminology and adapting to context. The problems are more pronounced within scientific and technical fields. Wu Xu [11], taking scientific texts from The Economist as the research subject, compared the translation performance of Baidu, Youdao, and DeepL. The study identified issues at the lexical level, such as the inaccurate identification of compound terms. For

DOI: 10.54254/2753-7048/2026.HT31097

instance, "exotendons" was mistranslated respectively as "Waishenqi(外伸器)" "Waishenjijian(外伸肌腱)" and "Waijian(外腱)" which stems from the lag in corpora incorporating emerging terminology. At the discourse level, due to imperfections in the attention mechanism, machine translation struggles to capture intra-sentential cohesive relationships. For example, the semantic connection between "from the design...to weather forecasting" and the preceding "enabled" was disrupted, resulting in a merely literal translation.

The research of Cai Yuan et al. [12] also corroborates common issues in scientific text translation, including the omission of term translation and category mismatches. These problems persist even when using Baidu's specialized API. Wu Tingting [13] further pointed out that disambiguating polysemous words is a common difficulty in cross-lingual vertical domain translation. Approximately 80% of the test samples contained errors in word sense selection, attributable to the model's "hard alignment" mechanism lacking sufficient context dependency.

Post-editing, serving as the critical link connecting machine translation with practical application, has developed a diverse system of theoretical and practical research. It is generally categorized into full post-editing (FPE), which aims for functional equivalence with human translation, and light post-editing (LPE), which focuses primarily on accuracy and comprehensibility. Regarding scientific and technical texts, Song Weiwen et al. [4] proposed a hierarchical editing strategy: employing light post-editing to adjust expressions for lexical errors, implementing full post-editing to restructure sentences for syntactic and logical errors, and resorting to creative retranslation for ambiguities and back-translation errors.

Current research still exhibits deficiencies in three key aspects: first, the development of vertical-domain corpora lags behind, with the delayed incorporation of emerging scientific and technical terminology being particularly acute; second, post-editing tools lack sufficient intelligence and customized features for specialized domains; third, studies on error types predominantly focus on surface linguistic features, with limited exploration of the internal semantic representation mechanisms within the models. Future research needs to integrate domain knowledge graphs to optimize terminology processing modules, develop self-adaptive editing tools with autonomous learning capabilities, and establish a cross-domain translation quality evaluation system, thereby promoting the deep integration of NMT and post-editing.

3. Methods and experimental design

This study specifically selected three mainstream NMT systems—Yandex, Google, and Baidu—as the experimental subjects. A set of 50 technical text sentences was chosen as the experimental material, alongside 50 general domain text sentences for control. The internationally recognized BLEU (Bilingual Evaluation Understudy) algorithm was employed to evaluate the quality of the translated outputs, aiming to objectively reflect the current performance of these NMT systems in Russian-to-Chinese translation within the scientific and technical domain. Following the BLEU score assessment, this study will further conduct a manual classification and statistical analysis of machine translation errors.

3.1. Research questions

This study aims to investigate the following two research questions (RQs): (RQ1) Are there significant differences in the quality of Chinese translations produced by mainstream machine translation systems when translating Russian technical texts compared to general texts? (RQ2) What

DOI: 10.54254/2753-7048/2026.HT31097

are the prevalent error types in Russian-to-Chinese translations of scientific/technical texts, and what characteristics do they exhibit?

3.2. Research workflow

The technical workflow adopted in this study is as follows: construction of an experimental dataset and a control dataset, utilization of mainstream domestic and international neural machine translation systems to perform Russian-to-Chinese machine translation on the experimental and control data respectively, followed by the calculation of BLEU scores, and finally, the classification and statistical analysis of errors in the machine-generated translations.

3.3. Research dataset

To investigate whether significant differences exist in the quality of Russian-to-Chinese translations produced by mainstream domestic and international neural machine translation systems for technical texts versus general texts, the researchers constructed two datasets. The experimental dataset comprises Russian-Chinese bilingual texts from the domain of scientific and technological news. In contrast, the control dataset primarily consists of Russian-Chinese bilingual texts from the domain of news on diplomacy and cooperation. The source texts for both datasets were sourced from the internet and are of an open nature, aiming to maximize the comparability of the experimental results.

3.3.1. Selection of Chinese-Russian bilingual corpora

After careful screening and selection, the researchers chose the following texts as the primary sources for the raw data in the experimental dataset:

- "Китайские учёные разработали носимую систему анализирования пота для раннего предупреждения о болезни Паркинсона" ("Chinese Scientists Develop Wearable Sweat Analysis System for Early Warning of Parkinson's Disease")
- "Китай стал одним из глобальных лидеров 'зеленой технологической революции' -- эксперт ИКСА РАН" ("Russian Expert: China Becomes a Global Leader in the 'Green Technology Revolution'")

For the construction of the control dataset, the researchers selected segments from the following bilingual parallel texts:

- "Китай направил приглашения на Всемирную конференцию по ИИ 2025 года высокопоставленным представителям более 40 стран и международных организаций -- МИД КНР" ("Chinese Foreign Ministry: China Has Sent Invitations to High-Level Representatives from Over 40 Countries and International Organizations for the 2025 World AI Conference")
- "Китайская компания и РФПИ намерены совместно построить в Пекине новый ориентир сотрудничества двух стран" ("Chinese Company and RDIF Intend to Jointly Build a New Landmark of Bilateral Cooperation in Beijing")

3.3.2. Corpus preprocessing

Corpora downloaded from the internet often have inconsistent formats, containing extraneous spaces, tables, figures, and other elements irrelevant to the research focus. Therefore, data cleaning was initially performed. The researchers preprocessed the corpora through the following three steps: (a) converting all texts into TXT format; (b) removing redundant spaces, blank lines, as well as

DOI: 10.54254/2753-7048/2026.HT31097

charts, symbols, etc., unrelated to the research task; (c) deleting tables of contents and various types of annotations present in the original documents.

3.3.3. Bilingual corpus sentence alignment

Following the completion of the raw corpus preprocessing, the researchers performed sentence alignment on the Russian-Chinese bilingual texts.

3.4. Construction of the research dataset

The researchers conducted further processing on both the experimental and control datasets. On one hand, a random sampling method was employed, using a computer to extract 50 Russian sentences from the technical text experimental dataset. Subsequently, the same method was used to extract 50 sentences from the general text control dataset. Finally, manual verification was performed on the alignment quality and text format of the selected bilingual aligned data, ultimately finalizing the construction of the research dataset. The basic information for the two datasets is presented in Table 1.

Table 1. Basic information of the experimental dataset and control dataset

Dataset Type	Type Original Text (Russian)		Translated Text (Chinese)	
	Number of Sentences	Number of Words	Number of Sentences	Number of Characters
Military Texts	50	203	50	462
General Texts	50	234	50	719

The constructed datasets were imported into Microsoft Excel. Utilizing a self-developed Python script, the online machine translation engines of Google, Baidu, and Yandex were called to perform Russian-to-Chinese translation on the collected Russian source texts. The translation results were retrieved and displayed in Microsoft Excel for subsequent data analysis. The experiment was conducted on August 21, 2025, and the translation results are shown in Figures 1 and 2.

50 Common Audit Statements (Chinese-Russian Bilingual Edition)

Number	Reference Translation	Original Text	Google Translate	Baidu Translate	Yandex Translate
1	中铁置业与俄罗斯直接投资基金在	CRREG и РФПИ провели перего	中俄关系发展公司（CRREG）与俄	俄罗斯联邦国防部与俄罗斯联邦总	CRREG和RDIF在莫斯科举行了深化务实合作的会
2	双方计划在北京打造中俄经贸文化	Стороны планируют создать в П	双方计划在北京设立俄中贸易文化	双方计划在北京建立中俄商贸文化	双方计划在北京建立一个俄中贸易文化中心，作为
3	该项目将推动中俄企业在金融、能	Проект будет способствовать со	该项目将促进企业在金融、能源和	该项目将促进企业在金融、能源和	该项目将促进企业在金融、能源和文化领域的合作
4	俄罗斯直接投资基金是俄主权财富	РФПИ — суверенный фонд Росс	俄罗斯直接投资基金是俄罗斯主权	俄罗斯直接投资基金（RDIF）是俄	俄罗斯直接投资基金是俄罗斯的主权财富基金，抓
5	中俄新时代全面战略协作伙伴关系	Всеобъемлющее партнёрство К	该项目以中俄全面伙伴关系为基础。	中俄全面伙伴关系为该项目奠定了	中国和俄罗斯之间的全面伙伴关系构成了该项目的
6	中俄经贸文化中心拟落户北京长安	Центр планируется разместить	该中心计划位于北京长安街商业区。	计划将该中心设在北京长安街的商	该中心计划位于北京长安街商业区。
7	中心将整合商务、文化、展览功能，	Центр объединит деловые, куль	该中心将为两国企业提供商务、文	该中心将为两国公司提供商务、文	该中心将结合商业，文化和展览功能，为来自两国
8	俄方拟吸引能源、金融等领域企业	Российская сторона привлечёт	俄方将吸引能源和金融领域的企业	俄方将吸引能源和金融领域的公司	俄罗斯方面将涉及能源和金融公司在该中心工作。
9	项目暂定名为"中俄经贸文化中心"。	Предварительное название про	该项目的初步名称为"俄中贸易文化	项目暂定名为"中俄商贸文化中心"。	该项目的初步名称是俄中贸易文化中心。
10	双方已签署合作备忘录，推进具体	Стороны подписали меморандум	双方签署了实施该项目的合作备忘	双方签署了合作备忘录以实施该项目	双方签署了实施该项目的合作备忘录。

Figure 1. Translation results of the three machine translation systems (Google, etc.) - general sentences

50 Tech-Related Review Phrases (Chinese-Russian Bilingual Edition)

Number	Reference Translation	Original Text	Google Translate	Baidu Translate	Yandex Translate
1	中国是全球绿色技术革命的领导者之一。	Китай является одним из мировых лидеров "зелён	中国是世界"绿色科技革命"的领军国家之一。	中国是"绿色技术革命"的全球领导者之一	中国是"绿色技术革命"的世界领导者之一。
2	中国在环保设备和绿色技术专利数量上	Китай занимает ведущие позиции по количеству па	中国在环境技术领域的专利数量居世界领先	中国在环境技术专利数量方面处于领先地	中国在环境技术领域的专利数量方面处于
3	中国已向200多个国家提供绿色技术解决	Китай поставляет экологическое оборудование бо	中国向200多个国家供应环保设备。	中国向200多个国家提供环保设备。	中国向200多个国家提供环保设备。
4	中国在太阳能和风能应用规模方面全球	КНР лидирует по масштабам использования солне	中国在太阳能和风能利用规模方面处于领先	中国在利用太阳能和风能方面处于领先地	中国是太阳能和风能利用的领导者。
5	中国建立了全球最大的碳排放权交易市	Китай создал крупнейший в мире рынок торговли	中国已经建立了全球最大的碳排放配额交易	中国建立了世界上最大的碳交易市场。	中国创造了世界上最大的碳排放交易市场。
6	中国金融科技公司直接资助绿色项目开	Китайские финтех-компании финансируют зелёные	中国的金融科技公司正在为绿色项目提供融	中国金融科技公司正在资助绿色项目。	中国金融科技公司为绿色项目融资。
7	应用大数据和AI技术解决气候问题。	Используются big data и ИИ для решения климатич	大数据和人工智能被用于解决气候问题。	大数据和人工智能被用来解决气候问题。	大数据和人工智能被用来解决气候问题。
8	区块链技术确保碳足迹可追溯性。	Блокчейн обеспечивает отслеживаемость углерод	区块链确保碳足迹的可追溯性。	区块链提供了碳足迹的可追踪性。	区块链确保了碳足迹的可追溯性。
9	物联网系统实时监控能源消耗。	Системы IoT осуществляют мониторинг энергопотр	物联网系统实时监控能源消耗。	物联网系统可以实时监控能源使用情况。	物联网系统实时监控能耗。
10	生物识别技术应用于环境监测设备。	Биометрия применяется в экологическом монитор	生物识别技术被用于环境监测。	生物计量学应用于环境监测。	生物识别技术用于环境监测。

Figure 2. Translation results of the three machine translation systems (Google, etc.) - technical sentences

DOI: 10.54254/2753-7048/2026.HT31097

3.5. BLEU score calculation

Following the text processing method proposed in Reference [10], this study likewise employed BLEU for evaluation. The Bilingual Evaluation Understudy (BLEU) algorithm, proposed by IBM in the United States, is an automated metric for evaluating the correspondence between machine translation and professional human translation. It is currently a widely adopted automatic evaluation metric for machine translation. In practice, the calculation can be performed directly in Python by utilizing the nltk.translate.bleu_score toolkit from NLTK (Natural Language Toolkit). In this experiment, all parameters were set to their default values, and the data smoothing method recommended by Chen & Cherry (2014) — Method 4 — was applied. The final BLEU score data was obtained through this calculation process.

4. Experimental results and discussion

4.1. Overall findings

In the automatic evaluation based on BLEU scores, a higher BLEU value indicates that the machine translation result is closer to the human reference translation. The descriptive statistics of the experimental results for the two datasets show that the average BLEU scores for the Russian-to-Chinese translations of scientific/technical domain texts by the three NMT systems (Google, Baidu, Yandex) are 38.69, 38.84, and 38.53 respectively, with an overall mean of 38.69 (see Table 2 for details). For general domain texts, the average BLEU scores are 38.76, 38.82, and 38.51 respectively, with an overall mean of 38.70 (see Table 3 for details). The experimental results confirm that neural machine translation, represented by the current mainstream commercial systems, performs slightly better in translating general domain Russian-Chinese texts compared to scientific/technical texts. The average BLEU score for the former is 0.19 points higher than that for the latter.

Table 2. Performance of the three NMT systems (Google, etc.) on technical text translation (BLEU scores)

System	Sample Size	Min	Max	Mean	Standard Deviation
Google	50	26.34	52.17	38.69	5.23
Baidu	50	27.12	53.45	38.84	5.31
Yandex	50	25.89	51.78	38.53	5.19

Table 3. Performance of the three neural machine translation systems (Google, etc.) on general domain text translation (BLEU scores)

System	Sample Size	Min	Max	Mean	Standard Deviation
Google	50	26.78	52.89	38.76	5.35
Baidu	50	27.45	53.92	38.82	5.42
Yandex	50	26.12	52.34	38.51	5.28

DOI: 10.54254/2753-7048/2026.HT31097

4.2. Error type analysis

Statistical analysis reveals that Baidu Translate achieved the highest average BLEU scores in both domains (38.84 for technical texts and 38.82 for general texts). Google Translate performed moderately, slightly below Baidu but superior to Yandex. Yandex consistently registered the lowest average BLEU scores across both text types. All systems exhibited standard deviations around 5.2, indicating a certain degree of fluctuation in translation quality across different sentences. The performance levels for scientific/technical and general domains were very close, suggesting these NMT systems possess a comparable adaptability to different textual domains. The ranges between maximum and minimum values demonstrate that the translation quality for certain specific sentences can significantly exceed or fall short of the average level. These results reflect the overall capability of current mainstream NMT systems, with Baidu holding a slight lead in Chinese-Russian translation tasks, while all three systems demonstrate relatively stable performance across different domains.

Based on the manual analysis and classification of error types, it was found that all three translation systems—Google, Baidu, and Yandex—produced errors in both scientific/technical and general text translations. However, the distribution and severity of these errors varied, as detailed in Tables 4 and 5.

Table 4. Error type statistics for general text translation

Translation Tool	Error Type Combination	Number of Errors (Total: 50 sentences)	Percentage
Google Translate	Spelling + Lexical + Syntactic + Semantic	50	100%
Baidu Translate	Lexical + Semantic	45	90%
	Spelling + Lexical + Syntactic + Semantic	5	10%
Yandex Translate	Lexical + Semantic	44	88%
	Spelling + Lexical + Syntactic + Semantic	6	12%

Table 5. Error type statistics for technical text translations

Translation Tool	Error Type Combination	Number of Errors (Total: 50 sentences)	Percentage
Google Translate	Spelling + Lexical + Syntactic + Semantic	50	100%
Baidu Translate	Lexical + Semantic	40	80%
	Error-free	1	2%
	Lexical + Syntactic + Semantic	1	2%
	Spelling + Lexical + Syntactic + Semantic	8	16%
Yandex Translate	Lexical + Semantic	39	78%
	Error-free	2	4%
	Spelling + Lexical + Syntactic + Semantic	9	18%

(1) Google Translate Outputs:

In both general and technical text translations, the outputs from Google Translate were assessed as containing spelling, lexical, syntactic, and semantic errors simultaneously in all 50 cases for each text type. This indicates that Google Translate exhibits severe and systematic issues across all these error categories for both text genres. It is hypothesized that this stems from significant limitations in its language processing algorithms when handling the complex structures of Chinese and Russian,

DOI: 10.54254/2753-7048/2026.HT31097

specialized vocabulary, and specific contextual nuances, ultimately failing to accurately comprehend the source text meaning and perform appropriate translation transformations.

(2) Baidu Translate Outputs:

General Texts: The primary error types were concentrated in lexical and semantic errors, with 45 instances. Additionally, there were 5 instances containing a combination of spelling, lexical, syntactic, and semantic errors. This indicates that Baidu Translate faces significant challenges in word choice and accurate semantic representation when translating general texts, although its performance regarding syntax and spelling is relatively better. This may be attributed to the more flexible and diverse linguistic expressions found in general texts, which include numerous idiomatic expressions and culture-specific references, thereby impacting the accurate translation of vocabulary and meaning.

Technical Texts: There were 40 instances of lexical and semantic errors, 1 error-free instance, 1 instance combining lexical, syntactic, and semantic errors, and 8 instances combining spelling, lexical, syntactic, and semantic errors. Compared to general texts, the presence of error-free instances and a more dispersed distribution of error types in scientific/technical text translation suggests that Baidu may have implemented certain optimizations for specialized vocabulary and fixed expressions in this domain. Nevertheless, room for improvement remains in its handling of lexicon and semantics.

(3) Yandex Translate Outputs:

General Texts: There were 44 instances of lexical and semantic errors, and 6 instances combining spelling, lexical, syntactic, and semantic errors. Similar to Baidu, the primary issues are concentrated at the lexical and semantic levels. This is hypothesized to result from linguistic and cultural differences, as well as the translation model's misinterpretation of certain specific expressions.

Technical Texts: There were 39 instances of lexical and semantic errors, 2 error-free instances, and 9 instances combining spelling, lexical, syntactic, and semantic errors. Yandex produced slightly more error-free instances in scientific/technical text translation than Baidu. However, the overall distribution of error types is similar to that observed in general texts, suggesting a certain consistency in its translation strategy across different text types.

5. Conclusion

This study conducted a comprehensive evaluation of mainstream neural machine translation (NMT) systems, focusing on errors in Chinese-Russian translation. The findings indicate that while these systems possess a foundational level of translation capability, significant room for improvement remains. Across various tests, Baidu's NMT system demonstrated relatively superior performance in the Chinese-Russian translation task, holding a slight lead over the other compared systems. Furthermore, all evaluated systems, including Baidu, exhibited relatively stable performance across different domains, without significant fluctuations attributable to domain variation.

To further optimize Chinese-Russian specialized translation systems, we propose the following recommendations. First, establishing a specialized scientific terminology database is crucial, as this ensures the accuracy and consistency of technical term translation. Second, integrating terminology annotation with grammatical parsing technology can effectively enhance translation accuracy and reduce errors related to grammar and terminology. Finally, actively promoting the development of human-machine collaborative Computer-Aided Translation (CAT) platforms will allow machine translation and human expertise to complement each other's strengths, thereby improving both translation efficiency and quality. Through these optimization measures, Chinese-Russian neural

DOI: 10.54254/2753-7048/2026.HT31097

machine translation systems are expected to achieve higher performance levels in the future, better meeting the demands of both professional fields and everyday communication.

References

[1] ZHAO Y, ZHANG H, YANG Y C. A comparative study on the translation quality of large language models—Taking Blossoms as an example [J]. Foreign Language Education and Technology, 2024(4): 60-66. doi: 10.20139/j.issn.1001-5795.20240409. (in Chinese)

[2] RONG Y, HUANG C X. Research on problems and countermeasures of neural machine translation—Focusing on Youdao, DeepL, Sogou, and Baidu [J]. Modern Linguistics, 2024, 12(6): 218-225. doi: 10.12677/ml.2024.126455. (in Chinese)

[3] WANG H S, LIANG X R. Research on translation technology standards in the era of artificial intelligence [J]. Chinese Translation Studies, 2024(2): 197-209. (in Chinese)

[4] SONG W W, ZHAO J. Research on translation quality of scientific texts under neural machine translation [J]. Journal of Suihua University, 2024, 44(6): 63-65. (in Chinese)

[5] GE P J. Error types of scientific and technical text machine translation and post-editing strategies [D]. Jilin: Northeast Electric Power University, 2024. doi: 10.27008/d.cnki.gdbdc.2024.000287. (in Chinese)

[6] ZHENG G F, PAN X. Terminology machine translation in the AI era: Challenges and solutions [J]. Chinese Foreign Languages, 2025, 22(4): 96-104. doi: 10.13564/j.cnki.issn.1672-9382.2025.04.005. (in Chinese)

[7] LI X Q, ZHANG Y Q. Discussion on post-editing problems in machine translation [J]. Jingu Cultural Creativity, 2025(23): 93-95. doi: 10.20024/j.cnki.CN42-1911/I.2025.23.026. (in Chinese)

[8] Liu Y. Improving machine translation accuracy for underrepresented languages in linguistic research using transformer models [J]. Journal of Computational Methods in Science and Engineering, 2025, 25(5): 4523-4538.

[9] Zafar M Z, Nazir A I. Can we trust machines? A critical look at some machine translation evaluation metrics [J]. Computer Science and Engineering International Journal, 2025, 15(3): 1-13.

[10] GUO W H, HU F M. Research on evaluation and post-editing of neural machine translation [J]. Journal of Beijing International Studies University, 2021, 43(5): 66-82. (in Chinese)

[11] WU X. Machine translation and post-editing: A practice report on the English-Chinese translation of The Economist technology articles [D]. Tianjin: Tianjin Normal University, 2022. doi: 10.27363/d.cnki.gtsfu.2022.000022. (in Chinese)

[12] CAI Y, WANG H. Research on post-editing patterns of scientific texts based on neural machine translation [J]. Translation Research and Teaching, 2023(2): 120-127. (in Chinese)

[13] WU T T. Advantages and disadvantages of machine translation in Chinese-Russian practice and suggestions for improvement [J]. Modern Linguistics, 2024, 12(2): 1145-1151. doi: 10.12677/ML.2024.122153. (in Chinese)