

FUNCTIONAL-STYLE VARIABILITY OF NEURAL MACHINE TRANSLATION QUALITY (BASED ON THE RUSSIAN–GERMAN LANGUAGE PAIR)

Seregina M.A., Kozlova S. R.
Rostov-on-Don, Russian Federation
Southern Federal University

Summary. The article presents the results of an experimental study of the quality of neural machine translation from German into Russian using four artificial intelligence systems: DeepSeek, Yandex Translate, PROMT.One, and Alisa AI. Based on 16 authentic texts (scientific, official, journalistic, and literary styles) of 800–1000 words each, systemic errors, advantages, and limitations of each group of services are identified. It is proved that translation quality directly depends on the functional style of the text: large language models (LLMs) consistently outperform specialised NMT translators. It is established that none of the systems is capable of adequately rendering wordplay, authorial irony, poetic form, or complex cultural realities. Practical recommendations are proposed for using AI for draft translation followed by human post-editing.

Keywords: neural machine translation; functional styles; Russian–German language pair; translation adequacy; large language models; NMT; post-editing.

Introduction. The current stage of development of machine translation (MT) is characterised by the dominance of neural networks (NMT) and the active integration of large language models (LLMs) into everyday translation practice. Systems such as DeepSeek, ChatGPT, Yandex Translate, PROMT.One, and Alisa AI have become indispensable tools both for professional translators and for mass users. However, translation quality remains heterogeneous and strongly depends on the text type [3; 4]. Of particular interest is the contrastively complex Russian–German language pair, where differences in morphology (declension, conjugation, gender), syntax (frame constructions, word order), and word formation (long compounds, separable prefixes) create additional obstacles for automated translation.

The aim of this article is to present the results of a comparative analysis of translations performed by DeepSeek, Yandex Translate, PROMT.One, and Alisa AI, and to identify error patterns depending on the functional style of the text. Unlike previous studies that focused either on a single language pair or on one text type [2; 5], our work offers a comprehensive evaluation methodology that simultaneously considers three parameters: adequacy (semantic accuracy), normativity (compliance with language norms), and pragmatic adaptation (rendering of style and persuasive function).

Materials and Methods. The material comprised a corpus of 16 authentic German-language texts, each 800–1000 words in length, representing four functional styles. The scientific style included a medical article on oral hygiene for the elderly and a macroeconomic report on Germany's chemical industry. The official style was represented by the Schwerin cemetery fee schedule, a Bundestag session protocol, and a ruling of the German Federal Court of Justice. The journalistic style included an interview with mountaineer Mollie Hughes, an analytical commentary on a climate station in Hawaii, an article on theatre, and a magazine column fragment. The literary style was represented by fragments from Thomas Mann's novel *Buddenbrooks*, Max Kruse's children's book *Urmel from the Ice*, and Friedrich Schiller's ballad *The Cranes of Ibycus*.

The experimental methodology included a unified translation procedure. For LLMs (DeepSeek, Alisa AI), a single prompt was used: "Translate the following text from German into Russian, preserving the features of the [style]. Do not add your own comments or change the paragraph structure." For NMT translators (Yandex Translate, PROMT.One), standard mode was used without additional settings. Each text was translated only once, without regeneration and without post-editing, to simulate a single user interaction.

Quality assessment was conducted using a three-level system of criteria developed based on the theory of V.N. Komissarov [2, pp. 108–118]. The first level – adequacy (semantic accuracy) – records the completeness and reliability of factual information transfer. The second level – normativity – evaluates compliance with the grammatical, lexical, and stylistic norms of the Russian language. The third level – pragmatic adaptation – determines how successfully the communicative effect of the original (emotions, evaluations, impact on the reader) is conveyed. Each translation received a final score on a 5-point scale: 5 – excellent (no editing required), 4 – good (minor inaccuracies), 3 –

satisfactory (verification required), 2 – poor (significant editor work required), 1 – unsatisfactory (word salad, incomprehensible).

Results and Discussion.

1. Scientific Style: Terminological Accuracy and Syntactic Adaptation

Analysis of the medical text translations showed that LLMs (DeepSeek, Alisa AI) handle highly specialised terminology significantly better than NMT systems. Consider the key example: *Einheitlicher Bewertungsmaßstab für zahnärztliche Leistungen (BEMA)*. DeepSeek and Alisa AI retained the German abbreviation in parentheses, adding an accurate translation: “*Uniform assessment standard for dental services (BEMA)*”. Yandex Translate provided a version without the abbreviation – “*Uniform scale for assessing dental services*” – which reduces the identifiability of the cultural reality. PROMT.One, although semantically accurate, also omitted the abbreviation.

More serious distortions were recorded in translating the term *Versorgungsforschung*. DeepSeek proposed the normative equivalent “*healthcare organisation research*”, while PROMT.One made a gross error, translating literally as “*supply research*”, completely distorting the meaning. This example demonstrates a typical problem of NMT systems: the inability to recognise complex compounds in medical discourse.

An important aspect of the scientific style is the rendering of passive constructions. In German, the passive voice (Passiv) is used much more frequently than in Russian, where active or impersonal constructions are preferred. The phrase “*...konnten keine Richtlinien abgeleitet werden*” was translated by DeepSeek most naturally: “*...it was not possible to develop any recommendations*” (active construction). PROMT.One retained a literal passive: “*...no recommendations were obtained*” – which is grammatically correct but stylistically less elegant for a Russian scientific text. Scores: DeepSeek – 5, Alisa AI – 4, Yandex Translate – 4, PROMT.One – 3.

2. Official Style: Abbreviations, Numerals and Legal Precision

The translation of documents revealed a systemic problem of all AI systems: inconsistency in handling abbreviations. In Alisa AI’s translation, *Kommunalverfassung für das Land Mecklenburg-Vorpommern* (KV M-V) was rendered as “*KU MP*” (Cyrillic abbreviation), but *Gesetz- und Verordnungsblatt für Mecklenburg-Vorpommern* (GVObI. M-V) remained in the original. Such a hybrid strategy disorients the reader. DeepSeek, in contrast, consistently leaves all legal abbreviations untranslated, providing a first explanation in the text, which is more reliable for legal identification.

The most complex fragment was a sentence from a ruling of the German Federal Court of Justice containing an ordinal numeral with a dot: “*...des 16a. Zivilsenats...*”. DeepSeek correctly recognised “16a.” as part of the senate name and translated coherently: “*of the 16a Civil Senate*”. PROMT.One and Yandex Translate perceived the dot as the end of a sentence, leading to absurd constructions: “*Upon the plaintiff’s appeal, a sentence of 16a will be passed.*” This example reveals a fundamental weakness of NMT systems: they cannot distinguish a grammatical dot as a sentence boundary from a dot as part of a numeral or abbreviation.

Another type of error relates to the rendering of gender-marked lexicon. German official documents regularly contain paired forms such as “*Gebührensschuldner bzw. Gebührenschuldnerinnen*”. DeepSeek and Alisa AI omit the feminine reference, using the neutral “*fee debtors*”, which corresponds to Russian legal tradition. Yandex Translate and PROMT.One, on the contrary, try to calque the German construction, creating cumbersome and unnatural phrases: “*fee debtors or female fee debtors*”. Such versions are not only stylistically deficient but also do not conform to the standards of Russian official discourse.

3. Journalistic Style: Expressiveness, Evaluation and Wordplay

Journalistic style requires not only accuracy but also the rendering of emotional and evaluative components. An analysis of the interview with mountaineer Mollie Hughes showed that Yandex Translate and PROMT.One made a gross pragmatic error: they used the masculine gender of past tense verbs (“*I took an interview*”, “*I wanted to understand*”), ignoring the female gender of the speaker, which is obvious from the context. LLMs correctly identified gender and used feminine forms.

More challenging is the handling of figurative meanings and metaphors. In the article about a climate station in Hawaii, the following sentence appeared: “*Entrückt von der Welt, fern von Schmutz und Abgasen, zählt sie zu den wichtigsten Wissenschaftsvorposten der Menschheit.*” The adjective *entrückt* carries connotations of “detachment, distance from worldly bustle”. DeepSeek conveyed expressiveness through “*Detached from the world*”, Alisa AI chose the neutral “*Remote*”,

while PROMT.One made a semantic catastrophe, confusing *entrückt* with the paronym *entzückt* (“enchanted”), and translated as “*Enchanted by the world*” – completely distorting the author’s intention.

The most vulnerable point for all AI systems turned out to be wordplay. The theatre name “*Raus Theater*” contains an ambiguity: a reference to founder Milo Rau’s surname and simultaneously the verb *raus* (“to exit”). No service handled this task. DeepSeek translated as “*Rau Theatre*”, Alisa AI retained the original “*Raus Theater*” without explanation, Yandex Translate and PROMT.One interpreted *raus* as a verb (“*Exit Theatre*”), completely destroying the meaning. This fact points to a fundamental limitation of modern AI: it works effectively with standard literary language but is incapable of creative interpretation of non-standard usages.

4. Literary Style: Idiolect, Speech Characteristics and Poetic Form

The greatest difficulties were recorded in the translation of literary prose and poetry. In Thomas Mann’s novel *Buddenbrooks*, long syntactic periods with multiple subordinate clauses lead to literalism and heavy constructions. The fragment: “*Die starken und elastischen Tapeten, die von den Mauern durch einen leeren Raum getrennt waren, zeigten umfangreiche Landschaften...*” All systems used the active voice (“*the wallpaper depicted*”), whereas in Russian the passive or impersonal construction (“*on the wallpaper were depicted*”) is more natural. DeepSeek and Alisa AI retained the complex structure, but it feels unnatural.

Of particular note is the rendering of speech defects in children’s literature. Penguin Ping from Max Kruse’s book *Urmel from the Ice* cannot pronounce the sound “sh”: “*„Ausgepfaffen?“ fragte Ping Pinguin. ...Es klang wie „pf“.*” DeepSeek adapted the defect in Russian, creating the occasionalism “*Vyfpalsja?*” (replacing “s” with “f”), which preserves the playful character of the original. Alisa AI used a less successful variant: “*Vygljadish pfal’shivo?*” – mixing onomatopoeia with lexical distortion. Yandex Translate and PROMT.One completely ignored the speech defect, translating neutrally: “*Did you sleep well?*” This case demonstrates that LLMs at least attempt to find a creative solution, whereas NMT systems do not even recognise the problem.

The translation of Schiller’s ballad *The Cranes of Ibycus* became the most revealing test. DeepSeek was the only service that attempted a rhymed translation (iambic tetrameter with cross rhyme). However, metre was often violated, and rhymes were imprecise. Alisa AI provided a literary prose translation, acceptable in semantics but completely ignoring the poetic form. Yandex Translate and PROMT.One produced a literal word-for-word rendering with gross lexical errors: PROMT translated *Kraniche* as “taps” instead of “cranes” – a classic “false friend” case that the system did not recognise. No AI could convey archaising vocabulary (*der Götterfreund, frommer Schauder*) or cultural realities (*Eumeniden*).

Generalisation and Classification of Errors. Based on the obtained data, a classification of neural machine translation errors has been developed across three levels.

Universal errors characteristic of all systems: lexical literalism (word-for-word translation, ignoring context), syntactic literalism (preserving German word order in Russian sentences), agreement and government errors (especially in cases and prepositions). These errors occur in all functional styles but are most critical in official and scientific texts.

Stylistically conditioned errors: in scientific and official styles – inability to recognise abbreviations, incorrect sentence segmentation due to dots, disruption of logical connections; in journalistic style – loss of emotional-evaluative component, non-distinction of paronyms, inability to render metaphors; in literary style – loss of imagery, authorial irony, idiolect, speech characteristics, and poetic form.

System-dependent differences: NMT translators (Yandex Translate, PROMT.One) demonstrate literalism, gross semantic distortions, and hallucinations significantly more often. LLMs (DeepSeek, Alisa AI) produce smoother and more coherent texts, better understand context and pragmatics, but are incapable of creative interpretation.

Conclusion and Practical Recommendations. The conducted research has confirmed a direct dependence of neural machine translation quality on the functional style of the text. AI handles the scientific style most successfully (average LLM score – 4.8; NMT – 3.2), and the literary style least successfully (LLM – 4.0; NMT – 1.8). Large language models demonstrate a systemic advantage over specialised NMT translators across all styles due to their better understanding of context and pragmatics. However, none of the tested systems is capable of adequately rendering wordplay, authorial irony, poetic form, or complex culturally specific realities.

The most rational strategy for using AI in translation practice is the “pre-translation + post-editing” model: automatic draft translation followed by mandatory human editing. For scientific and official texts, the volume of post-editing is minimal (10–20% of the text), for journalistic texts – moderate (30–40%), for literary texts – maximal (80–100%). Poetic translation and texts containing wordplay should currently be performed entirely by humans without the use of AI.

References

1. Atroshchenko, A. S., Ashikhmina, T. A. (2021). Subektnaya kategoriya perevoda kak kriteriy otsenki deyatel'nosti perevodchika [The subjective category of translation as a criterion for evaluating translator's activity]. *Nauchnye issledovaniya i innovatsii*, (1), 110–116.
2. Komissarov, V. N. (2011). *Sovremennoye perevodovedeniye* [Modern translation studies]. Moscow: R. Valent. 408 p.
3. Rarenko, M. B. (2021a). Mashinnyy perevod kak vyzov [Machine translation as a challenge]. *Vestnik Moskovskogo universiteta. Seriya 22. Teoriya perevoda*, (2), 117–125.
4. Rarenko, M. B. (2021b). Mashinnyy perevod: ot perevoda “po pravilam” k neyronnomu perevodu [Machine translation: from rule-based to neural translation]. *Sotsialnyye i gumanitarnyye nauki. Otechestvennaya i zarubezhnaya literatura. Ser. 6. Yazykoznanie*, 70–79.
5. Seregina, M. A. (2024). Voprosy obucheniya yazyku v kontekste razvitiya iskusstvennogo intellekta [Issues of language teaching in the context of artificial intelligence development]. In **Aktualnyye problemy nauki i tekhniki. 2024: Materialy Vserossiyskoy (natsionalnoy) nauchno-prakticheskoy konferentsii** (pp. 404–406). Rostov-on-Don: DSTU.
6. Khromykh, A. A. (2017). Klassifikatsiya funktsionalnykh stiley rechi (teoreticheskiy aspekt) [Classification of functional speech styles (theoretical aspect)]. In *Filologicheskiye nauki v Rossii i za rubezhom* (pp. 71–74). St. Petersburg: Svoye izdatel'stvo.
7. Sheyko, A. M. (2024). Opredeleniye kriteriyev otsenki kachestva pismennogo perevoda [Defining criteria for assessing the quality of written translation]. *Verkhnevolzhskiy filologicheskii vestnik*, (1/36), 149–155.
8. DeepSeek-V3: AI repository. Available at: <https://www.promt.ru/technology/neural-machine-translation/> (accessed: 02.02.2026).
9. A Neural Network for Machine Translation, at Production Scale: Google Research. Available at: <https://research.google/blog/a-neural-network-for-machine-translation-at-production-scale/> (accessed: 02.02.2026).

УДК 372.881.111.1

ИНСТРУМЕНТ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА PERPLEXITY ДЛЯ ПРЕОДОЛЕНИЯ ИНТЕРФЕРЕНЦИИ ПРИ ИЗУЧЕНИИ ФРАНЦУЗСКОГО ЯЗЫКА КАК ВТОРОГО ИНОСТРАННОГО

Сорокова Д. А.
Тамбов, Российская Федерация
ФГБОУ ВО «ТГУ им. Г. Р. Державина»

Научный руководитель:
Брыксина И. Е.
Тамбов, Российская Федерация
ФГБОУ ВО «ТГУ им. Г. Р. Державина»

Аннотация. Статья исследует межъязыковую интерференцию в контексте изучения французского языка на базе английского. Представлена идея использования Perplexity AI для исправления этих ошибок. Предлагается внедрение персонализированных промптов для обучения.

Ключевые слова: межъязыковая интерференция; прагматическая интерференция; Perplexity; французский язык как второй иностранный; искусственный интеллект.

В современной практике преподавания иностранных языков наблюдается тенденция, при которой английский язык становится первым изучаемым иностранным, а французский вторым. В процессе обучения, учитывая родство этих индоевропейских языков, неизбежно возникает явление интерференции, заключающееся в неосознанном переносе языковых норм с одного языка на другой [1].