Filip Kalaš
ORCID: 0000-0003-2554-9465
Ökonomische Universität in Bratislava, Bratislava

Linguistische Treffen in Wrocław, Vol. 27, 2025 (I)
ISSN: 2084–3062, e-ISSN: 2657–5647
https://doi.org/10.23817/lingtreff.27-7
S. 117–129

Evaluation of German-Slovak AI Translation of Stock Market News

The rapid advancement of technology has transformed communication, particularly through innovations in language and translation technologies. These tools have become essential for global interactions and are pivotal in modern linguistic studies. This paper investigates the application of three online statistical machine translation tools, ChatGPT-4, Google Translate and DeepL, for translating specialised German texts into Slovak. The study focuses on ten articles discussing various aspects of the stock exchange, a domain characterised by complex terminology and contextual nuances. By employing both quantitative and qualitative methods, the research evaluates the error rates, translation effectiveness, and the accuracy of these tools in preserving the original context. Specific challenges addressed include handling linguistic intricacies, domain-specific terminologies, and contextual fidelity unique to stock exchange texts. The analysis combines error rate calculations with qualitative assessments, offering a comprehensive evaluation of the tools' capabilities. The findings underscore the limitations and strengths of automated translation systems when applied to specialised text genres, providing critical insights for developers and practitioners in translation technology. The study shows that the tools often struggle with compound terms, anglicisms and jargon words. This study contributes to the growing body of knowledge in language technology, specialised domain translation, and machine translation research, highlighting areas for improvement and potential advancements in automated systems. Its practical implications extend to linguists, translators, and software developers aiming to enhance machine translation tools for specialised applications.

Keywords: machine translation, specialised domain, terminology, stock market language

Bewertung der deutsch-slowakischen KI-Übersetzung von Börsennachrichten

Der rasante Fortschritt der Technologie hat die Kommunikation verändert, insbesondere durch Innovationen in der Sprach- und Übersetzungstechnologie. Diese Werkzeuge, ChatGPT-4, Google Translate und DeepL, sind für globale Interaktionen unverzichtbar geworden und spielen in modernen linguistischen Studien eine zentrale Rolle. In diesem Beitrag wird die Anwendung von drei statistischen Online-Tools für die maschinelle Übersetzung von deutschen Fachtexten ins Slowakische untersucht. Die Studie konzentriert sich auf zehn Artikel, die sich mit verschiedenen Aspekten der Börse befassen, einem Bereich, der durch komplexe Terminologie und kontextuelle Nuancen gekennzeichnet ist. Durch den Einsatz quantitativer und qualitativer Methoden werden die Fehlerquoten, die Effektivität der Übersetzung und die Genauigkeit dieser Tools bei der Bewahrung des ursprünglichen Kontextes bewertet. Zu den besonderen Herausforderungen gehören die Handhabung sprachlicher Feinheiten, bereichsspezifischer Terminologien und die für Börsentexte typische Kontexttreue. Die Analyse kombiniert Fehlerratenberechnungen mit qualitativen Beurteilungen und bietet so eine umfassende Bewertung der Fähigkeiten der Tools. Die Ergebnisse unterstreichen die Grenzen und Stärken automatischer Übersetzungssysteme bei der Anwendung auf spezialisierte Textgattungen und liefern wichtige Erkenntnisse für Entwickler und Praktiker der Übersetzungstechnologie. Die Studie zeigt, dass die Übersetzungstechnologien oft mit Fachzusammensetzungen, Anglizismen und Jargonismen zu kämpfen haben. Diese Studie trägt zum wachsenden

Open Access article, distributed under the terms of the CC licence (BY-SA, https://creativecommons.org/licenses/by-sa/4.0/)

Wissensbestand in den Bereichen Sprachtechnologie, Fachübersetzung und maschinelle Übersetzung bei und zeigt Verbesserungsmöglichkeiten und potenzielle Fortschritte bei automatisierten Systemen auf. Ihre praktischen Auswirkungen erstrecken sich auf Linguisten, Übersetzer und Softwareentwickler, die maschinelle Übersetzungstools für spezielle Anwendungen verbessern wollen.

Schlüsselwörter: maschinelle Übersetzung, Fachgebiet, Fachterminologie, Börsensprache

Autor: Filip Kalaš, University of Economics in Bratislava, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia, e-mail: filip.kalas@euba.sk

Received: 19.11.2024 **Accepted**: 3.3.2025

1. Theoretical background

The surge in technological innovations has accelerated information dissemination and fundamentally transformed communication methods. Language and translation technologies, which have significantly impacted our lives for decades, are no exception to this change. In times of a profound restructuring of society into an information society, the machines and the technologies associated with them have become an indispensable part of humanity (cf. Wrede et al. 2020). Therefore, machine translation (MT) has been playing a significant role since its implementation in the early 1950s, not only among translators but also among scholars, and has been researched from various aspects as indicated by the plethora of related scientific articles.

With recent advancements in artificial intelligence (AI), there is an obvious potential in harnessing AI's power to revolutionize the translation process (Kunst/Bierwiaczonek 2023). It assists translators in working more efficiently and accurately by providing them tools that automate certain aspects of the translation process, such as detecting and correcting translation errors, suggesting alternative translations, or directly translating spoken or written text. The benefits of this include increased efficiency, consistency, cost savings and accuracy. Štefčík (2015) remarks that the cost and speed of translation are considered to be major factors in ascribing the social, political and economic importance of MT in many other areas of human activity.

MT has evolved over the last decade and has undergone various changes. Thus, types of MT may vary. Reynolds (2017) presents his classification of MT technology, i. e. a) rule-based machine translation using linguistic information about the source and target language such as grammar and dictionaries for MT, b) statistical machine translation (SMT) that determines translation outputs based on statistical models and c) example-based machine translation which operates based on bilingual corpora. Since technological progress cannot be stopped and science is constantly advancing, MT systems have also shifted from a rule-based method to a statistical-based method. Their advantage is that they gain knowledge from bilingual corpora, encompassing texts translated by skilled translators, and then, based on likelihood, they select the most appropriate solution which is applied in the target language. Generally, MT falls withing the realm of computational linguistics, centring on the automated conversion of text or speech from one natural language to another, devoid of human intervention.

Presently, due to ongoing advancements in cognitive science, artificial neural network-powered machine translation has emerged as the dominant method, showcasing robust performance (Hanji/Haiqing 2019). Neural machine translation (NMT) is a radical departure from previous machine translation approaches. It differs from SMT in its utilization of continuous rather than discrete symbolic representations. Furthermore, it employs a singular extensive neural network to manage the complete translation process, eliminating the requirement for extensive feature engineering. Unlike SMT, NMT undergoes end-to-end training instead of separate adjustments to individual components. Tan et al. (2020) stresses its simplicity and the fact that it has achieved state-of-the-art performance on different language pairs, and it becomes the essential technology behind many commercial MT systems. Thus, NMT is an integral part of numerous freely available technologies, which might have turned into the most widely used gadgets among both professional and non-professional society. Although this system can model better and more natural translations, it still has three weaknesses (Wu et al. 2016): 1. its slower training and inference speed, 2. ineffectiveness in dealing with rare words, and 3. failure to sometimes translate all words in the source sentence.

As previously mentioned, NMT is an approach that is used by diverse technologies, online translation services, of which we will deal with Google Translate, ChatGPT-4 and DeepL below. These tools have been chosen because they represent a diverse range of translation technologies: Google Translate excels in accessibility and real-time translation, ChatGPT-4 offers advanced contextual understanding and conversational flexibility, and DeepL is renowned for high-quality and nuanced translations. Other online tools, like Microsoft Translator or Amazon Translate, may not offer the same level of quality, flexibility, or accuracy, and they often lack the advanced contextual understanding or refined translation quality, making them less ideal for a comprehensive study.

1.1 Google Translate

Google Translate (GT) stands as a prominent online translation tool offered by Google, leveraging state-of-the-art machine learning algorithms, including neural machine translation models. With support for an extensive array of languages – over one hundred at last count – GT facilitates the seamless conversion of text, documents, websites, and spoken language between these language combinations. This translation tool has undergone many fundamental changes which have had a profound impact on the accuracy of the translated output, predominantly in more common languages on the Internet. Interestingly, GT typically does not perform direct translations between languages (L1→L2). Instead, it translates initially to English before converting to the intended target language. However, Czech and Slovak are known to take a different path, they are translated directly thanks to their intermediate genealogical closeness (cf. Benjamin 2018). However, in 2016, Google Neural Machine Translation (GNMT) was introduced in what constituted a big step forward, since NMT overrode the previously applied SMT. GNMT represents an improvement in that it is able to handle direct translations between language pairs of no demonstrable genetic relationship (Schuster et al. 2016). By implementing NMT, the translated results in less frequent languages are perceived as less clumsy and more natural.

1.2 GPT-4

Generative Pre-Trained Transformer (GPT), more specifically ChatGPT-4, a recent AI innovation by OpenAI, is a public tool based on GPT language model technology (cf. Kirmany 2022, Biswas 2023). This chatbot exhibits high sophistication, and is adept at handling diverse text-based inquires, ranging from straightforward questions to complex tasks like generating thank you letters or assisting individuals in navigating challenging discussions. Even though it was not specifically designed for translation, it can be fine-tuned for such tasks. GPT involves a dual-phase approach: in the first, pre-training phase, the model assimilates information naturally, akin to a person learning in a new environment or setting, while in the second, fine-tuning phase, it involves more directed and systematic refinement facilitated by its creators (cf. Radford et al. 2018, Lund/Wang 2023). According to Kunst and Bierwiaczonek (2023: 3), ChatGPT-4 benefits from this vast pre-training process that exposes it to diverse language structures and idiomatic expressions.

1.3 DeepL

Another online translation service that utilizes AI and neural network technology is DeepL. It is known for its accuracy and natural-sounding translations, and it provides translations between multiple language pairs. Compared to GT, this tool seems to still lag regarding its popularity, however, its translation accuracy is far greater than GT's (cf. Polaková/Klimová 2023). Additionally, the results of authors' surveys (cf. Yulianto/Supriatnaningsih 2021) demonstrate that DeepL offers varied synonyms and verb forms during translation, and it translates complex texts with enhanced comprehensibility. Interestingly, phraseology in the broadest sense remains a major challenge for both MT tools (cf. Bacquelaine 2023).

MT tools will likely enhance accuracy, naturalness and contextual awareness through AI advancements. Proper use of technological tools may improve the quality of translation, and we agree with Varela Salinas and Burbat (2023: 255) who claim that translators do not need to worry about the disappearance of their profession, but instead they should learn to use MT to increase their productivity.

2. Translation assessment

Quality standards for translation services such as EN 15038 (2006) and ISO 17100 (2015) require that translation be revised by a second translator. Revision is essential to ascertain the adequacy of a translation for its intended purpose, and to effectuate requisite adjustments concerning aspects such as terminological consistency, register, style and linguistic conventions. This meticulous process aims to uphold and ensure the quality and fidelity of the translated content (van Rensburg 2017).

There have been many discussions about the assessment of translation quality, there is nevertheless still no universal definition of translation quality or even generally accepted methods of assessment. Concerning this matter, Štefčík (2015: 143) emphasizes that national and international standards of translation (ATA, Sical) are available, but they are not widely accepted as objective criteria for evaluating the quality of translation. Generally, criteria such as linguistic correctness, fidelity to the source text, readability of the target text, equivalence and transfer of the meaning are taken into consideration when evaluating the quality of a translated text.

3. Reaserch aim

The primary objective of this study, as previously outlined, is to ascertain the error rate in translated specialised texts utilizing three distinct online MT tools. This error rate will subsequently be quantitatively presented and further elucidated qualitatively through the examination of specific examples. The obtained findings will help us to refute or corroborate the posited hypotheses:

H_1: higher error rate in target translation indicates its reduced lexical diversity,

H_2: errors in the translation of specialised texts will be primarily manifested on the level of lexical semantics,

H_3: the highest error rate will be detected mainly in the terminological category.

4. Data and methods

For the needs of this research, we excerpted specialised journalistic texts from the German weekly business news magazine "WirtschaftsWoche". The dataset consisted of the 10 latest texts from WirtschaftsWoche's section "BörsenWoche" 'stock market week', that is, the weekly investor newsletter authored by Jan-Lukas Schmitt and his team. The time span of the articles is October 2023 – December 2023. These contain specific investment tips, in-depth analyses and updates on our model portfolios. This means that the researched articles convey information from the stock market environment, whose language is perceived as relatively isolated subsystem. Its most obvious and striking characteristics are manifested on the lexical level, i. e., there is a plethora of various terms, slang expressions, anglicisms as well as idiomatic expressions (Kalaš 2021).

We translated the original German inventory version into Slovak using both Google Translate, DeepL and ChatGPT-4. All text was translated once on January 4, 2024. For CHATGPT-4, we employed the following prompt suggested by the chatbot itself:

"Could you please translate this text into Slovak?". Each text underwent translation through separate API requests.

In the subsequent phase of processing the machine-translated texts, we conducted an analysis of the errors that arose during the MT process. Here, we implemented Vaňko's (2017) error framework designed for Slovak. It is crucial to emphasize that this framework had already been applied by Slovak scholars (cf. Petráš/Munková 2023, Benková et al. 2021, Welnitzová et al. 2020, Wrede et al. 2020) who attempt to demonstrate the quality of MT via product analysis, more specifically the translated accuracy and representativeness of the content of the source text as well as the fluency and linguistic (grammatical) correctness in terms of the target text. At this point it is worth highlighting that none of them focused on stock exchange language.

5. Results and discussion

5.1 Error distribution

ChatGPT-4 as a modern machine translation tool produced the largest error rate in areas III (syntactic-semantic agreement and V (lexical semantics), with almost six times the error rate in area V compared to III. Surprisingly, the error distribution in areas I (predicativeness) and IV (syntax of complex sentences) is more or less proportional, whereas area II (modality) does not depict any translation error.

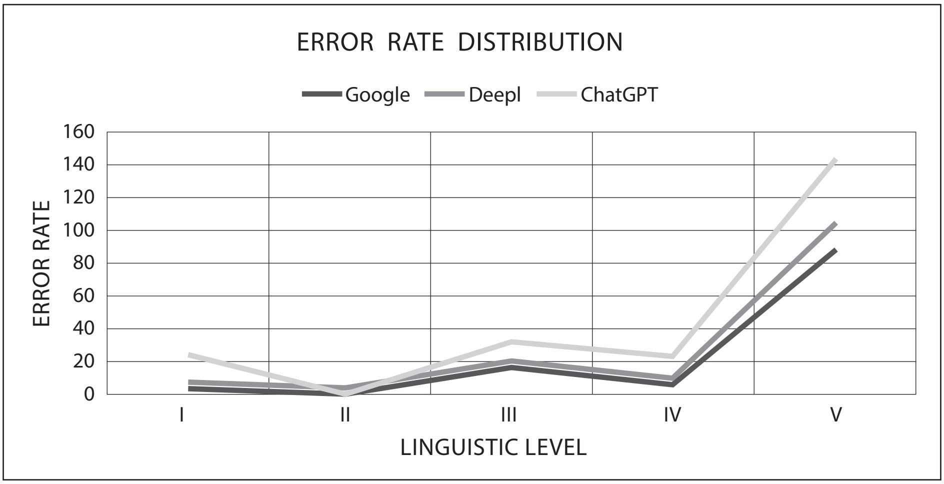

Fig. 1. SEQ Figure * ARABIC 1: An overall error rate distribution across five linguistic levels

As a result, no machine translation tool is error-free. Unanticipatedly, linguistic levels show diverse or even significantly diverging error rates. From a frequency point of view, the lowest error rate (17 in total) was detected in translated texts using the DeepL tool. Slightly higher error values (145 in total) were produced by Google Translate.

Unexpectedly, the highest error rate was detected in translations using the state-of-the-art artificial intelligence tool ChatGPT-4 with a total of 222 errors.

Given the expansive scope of errors spanning various linguistic levels within this framework and considering the sheer volume of data, the quantitative and qualitative analyses will exclusively address the aspect of lexical semantics.

5.2 Evaluation of texts translated by Google Translate

The total number of errors varies from text to text. The reason for this frequency dispersion cannot be its lexical diversity, since the latter, as has been shown, is textually proportional. Therefore, we assume that text complexity plays a key role. When selecting the text, we have ensured that all of them come from the stock exchange domain; however, the individual topics in the articles may differ from each other in terms of their complexity. Average error rate: 10.5 errors/text.

Within lexical semantics, the translation of terms followed by adequate transfer of meaning appears to be the most problematic in the qualitative analysis. Considering the distribution of errors in these two domains, the other domains appear to be irrelevant, and we will not deal with them further.

(1)

SL[1]: *So haben sich die **Musterdepots** der WiWo geschlagen.*

MT: *Takto sa darilo **modelových skladom** WiWo.*

PE: *Takto sa darilo **virtuálnym portfóliám** týždenníka WiWo.*

In example Nr. 1 we notice an inappropriate translation of the compound term *Musterdepots* 'model securities account depot' as *modelový sklad* 'model store'. It stands for a virtual securities account that has the same functions as a real securities account. Moreover, GT failed to decipher the abbreviation WiWo which stands for the German business news magazine Wirtschaftswoche and accordingly it should have been labelled with its attribute *týždenník* 'weekly magazine'.

(2)

SL: *Wie sich die **Musterdepots** geschlagen haben und warum unser Autor lieber **an** den **Weihnachtsmann** als an Analystenprognosen glauben würde.*

MT: *Ako sa darilo **modelovým portfóliám** a prečo by náš autor radšej veril **v Santa Clausa** ako v prognózy analytikov.*

PE: *Ako sa darilo **virtuálnym portfóliám** a prečo by náš autor veril radšej **na** na **Ježiška** ako na prognózy analytikov.*

In example 2, the term *Musterdepot* appears again. We would have expected the machine translation tool to deal with its translation similarly, but it offered us an alternative

[1] SL stands for text in source language, MT stands machine translation into target language and PE represents postediting effected by a posteditor.

equivalent. In a way, it offered us the correct alternative since this German term has two synonymous equivalents in Slovak, *modelové portfólio* or *virtuálne portfólio*.

In this sentence we find a culturally specific term *Weihnachtsmann*, which represents a folklore figure giving gifts to children and adults on Christmas Eve in German-speaking countries. Needless to say, the name of this figure varies across countries and cultures, hence it output the term *Santa Claus*. In Slovakia, there is the Christ Child as a deep-rooted Christmas figure. In this case the MT tool did translate the term, but inadequately to the cultural background of the target language. Interestingly, in another text passage there occurs an alternative translation of *Weihnachtsmann* as *Mikuláš* 'Saint Nicolas' who is, in Slovakia, traditionally considered to be a gifts bearer. The children are bestowed gifts in his honour on 6 December rather than 24 December.

The MT sentence contains a mistake in syntactic-semantical correlation. The verb *glauben an* can be translated into Slovak as *veriť na* or *veriť v*. The latter tends to be linked with collocations *veriť v Boha* 'to believe in God', whereas in this sentence the preposition *na* is semantically more appropriate.

5.3 Evaluation of texts translated by DeepL

Fig. 2 demonstrates a directly proportional distribution of error rate also in terms of DeepL. The highest and at the same time equal error rate has been discovered in the area of adequate transfer of meaning and terms. These two areas are followed by stylistic compatibility. Average error rate: 8.8 errors/text.

(3)

SL: *In Zeiten steigender **Notierung** sind Absicherungen gegen fallende Kurse günstig.*

MT: *V čase **rastu cien** je výhodné zabezpečenie sa proti poklesu cien.*

PE: *V čase rastúcej **kotácie** je výhodné zabezpečiť sa proti poklesu cien.*

In example Nr. 3 we encounter the term *Notierung*, which stands for 'quotation', i. e. statement of price. In this statement the price has been increasing and thus the MT tools translated it as *rast cien* 'price increase'. The meaning of *Notierung* implicitly conveys it is a process of a price statement of stocks. Therefore, its translation into *rast cien akcií* would be definitely more appropriate that the suggested equivalent. Another way of translating this term is to use the Slovak financial equivalent *kotácia*. The term *Absicherung*, highlighted in italics, has been translated into Slovak as *zabezpečenie sa*, which is correct indeed, however, such deverbal nouns are not frequently used in Slovak in the reflexive form. For this reason, the version without the clitic *sa* sounds more natural. To avoid such imprecision, we suggested replacing the noun with a verb.

(4)

SL: *Handeln Sie nach Sprichwörtern? „Sei gierig, wenn andere ängstlich sind", sagt etwa Warren Buffett – die alte Version von „**Buy the dip**".*

MT: *Obchodujete podľa príslovi? „Buď chamtivý, keď sa ostatní boja", hovorí napríklad Warren Buffett – stará verzia „**buy the dip**".*

PE: *Obchodujete podľa príslovi? „Buď chamtivý, keď sú ostatní vystrašení", hovorí Warren Buffett, čo je staršia verzia príslovia „špekulujte na pokles".*

At the beginning of sentence 4 is a polysemantic verb *handeln,* which means either to trade or to act. Both equivalents could be used in this sentence, yet the equivalent *obchodovať* 'to trade' is far better considering its context.

Surprisingly, the sentence includes the original English phrase *buy the dip*, which has not been translated into Slovak at all. It may have been done on purpose, working on the assumption that specialists in the stock exchange should understand this phrase, even in its original version. Nonetheless, in Slovak as well as in German there exist localised versions which could have been applied. Therefore, in the postediting phase we have used its Slovak equivalent, since not everyone may be familiar with the English phrase.

5.4 Evaluation of texts translated by ChatGPT-4

Fig. 2 illustrates that texts translated via ChatGPT-4 have higher error rates considering an adequate transfer of meaning and it leads terminological errors with 12 points. Generally, both fields show stupendously high error rates. The error rate in stylistic compatibility reaches "only" 18 points. Average error rate: 14.2 errors/text.

(5)

SL: *Seit 440 **Ausgaben** begleitet die BörsenWoche Kapitalmärkte-Trends. Wie **sich** die Musterdepots **geschlagen haben** und warum unser Autor lieber **an** den **Weihnachtsmann** als an Analystenprognosen **glauben** würde.*

MT: *440 **vydání** časopisu BörsenWoche sledovalo trendy na kapitálových trhoch. Ako sa modelové portfóliá **darili** a prečo by náš autor radšej veril **vianočnému mužovi** ako prognózam analytikov.*

PE: *Už v 440 vydaniach rubriky BörsenWoche sledujeme trendy na kapitálových trhoch. Ako sa darilo modelovým portfóliám a prečo by náš autor radšej veril na Ježiška než prognózam analyzikov?*

In example Nr. 5 we can observe several mistranslated passages. The first is the polysemantic term *Ausgabe*, which the MT tool translated correctly at first glance, yet into Czech. Its inflectional structure is slightly different in Slovak. For the term *Musterdepot* an adequate equivalent has been chosen, but the case reflected in this term and governed by the verb *darili* is not preserved. This verb governs the dative, hence the accurate form *modelovým portfóliám*. In addition, we can perceive the already analysed problem of the incorrectly translated cultural term *Weihnachtsmann*, for which the literal translation has been chosen. Interestingly, the MT tool set this word into the dative which alters the whole meaning of the phrase. Instead of saying "believe in Santa Claus", the MT version proposed "believe Santa Claus", as if he was the one to give investors advice.

(6)

SL: *Die Bewertungen von* ***Wachstumswerten*** *werden meist von Prognosen getragen, die wegen der höheren Inflationsraten* ***abdiskontiert*** *werden müssen.*

MT: *Hodnotenia* ***rastových spoločností*** *sú zvyčajne podporené prognózami, ktoré* ***je potrebné odsotovať*** *kvôli inflačným sadzbám.*

PE: *Ohodnotenia* ***rastových akcií*** sú *zvyčajne podporené prognózami, ktoré* ***treba diskontovať*** *pre inflačné sadzby.*

In example 6 we register four mistakes caused in the translation process. The first refers to the compound word *Wachstumswerte* 'growth stocks'. The first part has been translated adequately, however, the second part faces a translation shift. None of the meanings of the word *Werte* stands for *spoločnosti* 'companies'.

Another example is the verb *abdiskontieren* which is represented by its past participle form in the sentence. MT tools translated it as *odsotovať*, which is a word that does not exist in Slovak whatsoever, and we have replaced it by the Slovak equivalent *diskontovať* 'discount'.

The last revised translation is a copula phrase *je potrebné* 'is necessary' that has been replaced by sentence adverb *treba* 'is necessary'.

To sum up, all three tools exhibit varying performance across the ten texts, with fluctuating error rates (Fig. 2). ChatGPT-4's performance fluctuates significantly, with noticeable peaks in complex texts (like text 3 and text 5). This suggests that it may struggle to maintain consistency when processing nuanced lexical semantics in challenging contexts. While error rates of Google Translate are more stable than ChatGPT-4, an increase in errors is evident for some complex texts. However, it performs better than ChatGPT-4 in texts 4 through 6. Deepl exhibits the least fluctuation and has consistently lower error rates, particularly in complex texts. Its stability and accuracy suggest it might be better equipped to handle lexical nuances in complex language structures.

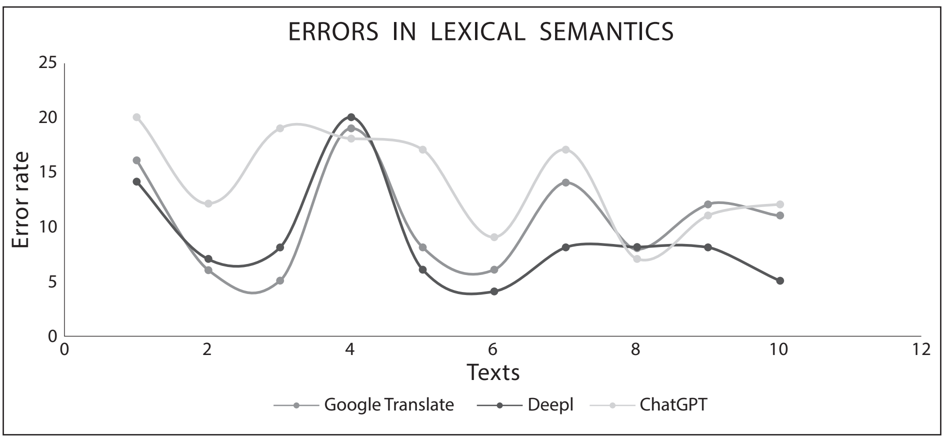

Fig. 2. Figure SEQ Figure * ARABIC 2: Error rate in lexical semantics across ten texts

6. Conclusion

The provided evaluation of translated specialised texts serves as the foundation for answering the posited hypotheses. H_1 can be refuted since the analysis suggests that error rates in the target translation are not directly linked to lexical diversity but rather to text complexity and contextual challenges. The lexical diversity was found to be proportional. H_2 is corroborated because the majority of errors occurred in lexical semantics, particularly in the translation of terms and maintaining an adequate transfer of meaning. This aligns with the hypothesis, as lexical semantics was identified as a critical area where errors predominantly manifested across all three machine translation tools. H_3 is also corroborated because terminological errors were specifically identifies as having the highest frequency, particularly in translations involving compound terms and domain-specific jargon.

It was revealed that MT tools exhibited distinct proficiency in handling terminology (Fig. 3). The highest error rate was identified in translations conducted by ChatGPT-4 (56 points), followed by slightly fewer errors in texts translated via Google Translate (50 points), whereas the lowest error rate was observed in translations carried out by DeepL. Surprisingly, despite its design for interactive communication and its diverse range of applications, ChatGPT-4 performed unsatisfactorily compared to the other MT tools, not only in the category of terminology, but also across the remaining subcategories.

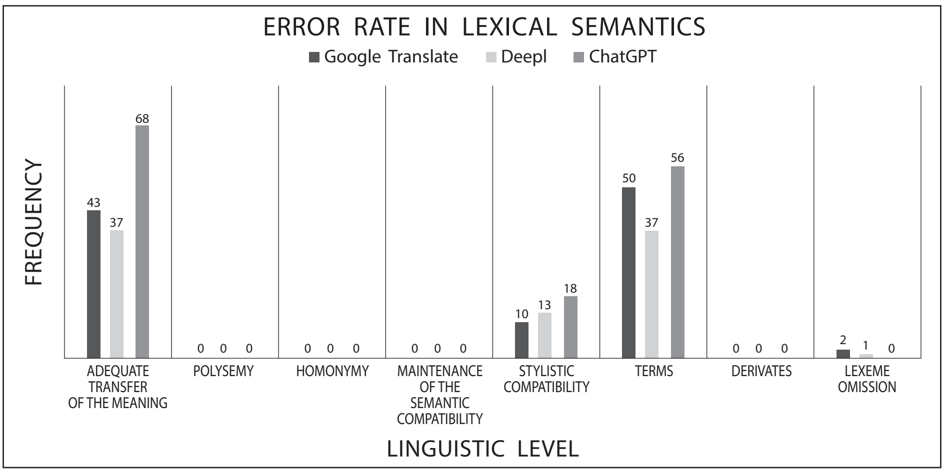

Fig. 3. Figure SEQ Figure * ARABIC 3: Error rate in lexical semantics across semantic levels texts

When translating terms, particularly in the specialised stock exchange language, a multitude of compound terms arises, which are characteristic for the German language. MT tools often encounter difficulties in handling these compounds, either translating both parts literally (e. g., *Musterdepots*) or correctly translating only one part (e. g., *Frühindikatoren*). Another challenge lies in the presence of English terms within stock exchange texts. Depending on the capabilities of the MT tool, these Anglicisms may either be translated into the target language or deemed unnecessary for translation. Errors were also detected in the translation of jargon such as *Werte*, *Wachstumswerte* and *Kursraketen*, which are highly interdisciplinary and frequently misunderstood. All tools handle polysemy, homonymy, and derivates well, with no errors. Lexeme omission is rare, occurring only slightly in ChatGPT-4 and Google Translate. In general, the trend suggests that Deepl is the most effective for complex texts, while ChatGPT-4 faces the most challenges, particularly in maintaining semantic and stylistic accuracy.

Acknowledgements

The research has been elaborated within the project "A-25-106/3020-18 Artificial intelligence: challenges for linguistics and marketing and the project. I-24-113-00 Artificial Intelligence in humanities and socioeconomic field".

References

Bacquelaine, Françoise. *DeepL and Google Translate Translating Portuguese Multi-Word Units into French: Progress, Decline and Remaining Challenges (2019–2023)*. https://repositorio-aberto.up.pt/bitstream/10216/149418/2/624249.pdf. 10.10.2024.

Benková, Lucia, Dasa Munková, Ľubomír Benko and Michal Munk. "Evaluation of English-Slovak Neural and Statistical Machine Translation". *Appl. Sci.* 11 (2948) (2021): 1–17. Print.

Biswas, Som. *Role of Chat GPT in Education*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4369981. 12.10.2024.

Benjamin, Martin. *Teach you backwards*. https://www.teachyoubackwards.com/extras/pivot/ 10.10.2024.

En 15038. *Translation services – Service requirements. European Committee for Standardization*. Brussels, 2006.

Iso 17100. *Translation services – Requirements for translation services. International Organization for Standardization*. Geneva, 2015.

Hanji Li and Chen Haiqing. "Human vs. AI: An Assessment of the Translation Quality Between Translators and Machine Translation". *International Journal of Translation, Interpretation and Applied Linguistics* 1.1 (2019): 43–54. Print.

Lund, Brady and Wang Ting. "Chatting About ChatGPT: How May AI and GPT Impact Academia and Libraries?". *SSRN* 2023. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4333415. 3.11.2024.

Kačala, Ján. *Syntagmatický slovosled v slovenčine.* Martin: Vydavateľstvo Matice slovenskej, 2013. Print.

Kalaš, Filip. *Kollokationen in der Börsensprache. Eine kontrastive Fallstudie.* Hamburg: Verlag Dr. Kovač, 2020. Print.

Kirmani, Ahmand R. "Artificial Intelligence-Enabled Science Poetry". *ACS Energy Letters* 8 (2022): 574–576. Print.

Kunst, Jonas R. and Kinga Bierwiaczonek. "Utilizing AI Questionnaire Translations in Cross-Cultural and Intercultural Research: Insights and Recommendations". *International Journal of Intercultural Relations* 97 (2023): 1–11. Print.

OpenAI. ChatGPT-4. https://chat.openai.com/chat.

Petráš, Patrik and Dana Munková. "Machine Translation Based on Neural Networks – A Promising Way to Translate from Analytic Languages into Flective Slovak?". *Slovenská reč* 88.1 (2023): 74–89. Print.

Polaková, Petra and Blanka Klimová. "Using DeepL Translator in Learning English as an Applied Foreign Language – An Empirical Pilot Study". *Heliyon* 9 (2023): 1–7. Print.

Radford, Alec, Karthik Narasimhan, Tim Salimans and Ilya Sutskever. "Improving Language Understanding by Generative Pre-Training". *OpenAI*, 2018. https://s3-us-west-2.amazonaws.com/openai-assets/research-covers/language-unsupervised/language_understanding_paper.pdf. 5.11.2024.

Reynolds, Peter. "Machine Translation, Translation Memory and Terminology Management". *Handbook of Terminology* 1 (2017): 276–290. Print.

Schuster, Mike, Melvin Johnson and Nikhil Thorat. *Google Research Blog*. 22.11.2016. https://blog.research.google/2016/11/zero-shot-translation-with-googles.html. 5.11.2024.

Štefčík, Jozef. "Evaluating Machine Translation Quality: A Case Study of a Translation of a Verbatim Transcription from Slovak into German". *Vertimo Studijos* 8 (2015): 139–153. Print.

Tan, Zhixing, Shuo Wang, Zonghan Yang, Gang Chen, Xuancheng Huang, Maosong Sun and Yang Liu. "Neural Machine Translation: A Review of Methods, Resources, and Tools". *AI Open* 1 (2020): 5–21. Print.

Van Rensburg, Alta. *Developing Assessment Instruments: The Effect of a Reviser's Profile on the Quality of the Revision Product*. Self-published, 2017. Print.

Vaňko, Juraj. "Kategoriálny rámec pre analýzu chýb strojového prekladu". *Mýliť sa je ľudské (ale aj strojové)*. Ed. Daša Munková, Juraj Vaňko and Jakub Absolon. Nitra: Univerzita Konštantína Filozofa v Nitre, 2017, 83–100. Print.

Varela Salinas, María-José and Ruth Burbat. "Google Translate and DeepL: Breaking Taboos in Translator Training. Observational Study and Analysis". *Iberica* 45 (2023): 243–266. Print.

Welnitzová, Katarína, Daša Munková and Oľga Wrede. "Chybovosť v predikatívnosti a kvalite strojového prekladu". *Jazyk a kultúra* 41–42 (2020): 160–172. Print.

Wirtschaftswoche. BörsenWoche Editorial. Wirtschaftswoche. https://www.wiwo.de/boersenwoche/editorial/. 4.10.2024.

Wrede, Oľga, Daša Munková and Renáta Pavlová. „Bewertung der Qualität maschineller Übersetzung journalistischer Texte aus dem Deutschen ins Slowakische". *Jazyk a kultúra* 41–42 (2020): 173–185. Print.

Wu, Yonghui, Mike Schuster, Zhifeng Chen, Quoc V. Le and Mohammed Norouzi. "Google's Neural Machine Translation System: Bridging the Gap Between Human and Machine Translation". *Ithaca* 2 (2016): 1–23. Print.

Yulianto, Ahmad and Rina Supriatnaningsih. "Google Translate vs. DeepL: A Quantitative Evaluation of Close-Language Pair Translation (French to English)". *AJELP* 9.2 (2021): 109–127. Print.

ZITIERNACHWEIS:

Kalaš, Filip. „Evaluation of German-Slovak AI Translation of Stock Market News", *Linguistische Treffen in Wrocław* 27, 2025 (I): 117–129. DOI: 10.23817/lingtreff.27-7.