

Simon Zupan, Zmago Pavličič,
Melanija Larisa Fabčič
University of Maribor, Slovenia

2025, Vol. 22 (1), 171-184(228)
journals.uni-lj.si/elope
<https://doi.org/10.4312/elope.22.1.171-184>
UDC: [8111.111'367.622=163.6:62]:004.89

Machine Translation of Independent Nominal Phrases in Technical Texts

ABSTRACT

This paper deals with machine translations of independent noun phrases in technical texts, which are not part of any sentence structure but function on their own, typically in tables and illustrations. Such nominal structures are common in technical texts because they allow technical writers to increase lexical density and precision in expression. On the other hand, these phrases pose a challenge for machine translation engines, as their meaning depends on the context. Independent noun phrases from a service manual, which were translated from English into Slovene by two different machine translators (DeepL and Google Translate), are considered in this paper. Their comparison with the original showed some limitations of machine translation engines in translating noun phrases, since approximately half of them showed a noticeable change in meaning.

Keywords: technical texts, machine translation, nominal phrases, translation shifts, technical translation

Strojno prevajanje samostojnih samostalniških besednih zvez v tehničnih besedilih

IZVLEČEK

Prispevek obravnava strojne prevode samostojnih samostalniških besednih zvez v tehničnih besedilih, ki niso del stavčnih struktur, temveč se pojavljajo zunaj konteksta, najpogosteje v preglednicah in grafičnih prikazih. Tovrstne besedne zveze se pogosto pojavljajo v tehničnih besedilih, saj piscem omogočajo večjo leksikalno gostoto in konciznost pri izražanju. Po drugi strani predstavljajo izziv za strojne prevajalnike, saj je njihov pomen odvisen od sobesedila. V prispevku so obravnavane samostojne samostalniške besedne zveze iz servisnega priročnika, ki so bile iz angleščine v slovenščino prevedene z dvema različnima strojnima prevajalnikoma (DeepL in Google Translate). Njihova primerjava z izvornikom je pokazala nekatere omejitve strojnih prevajalnikov pri prevajanju samostalniških besednih zvez, saj se je pri približno polovici besednih zvez opazno spremenil njihov pomen.

Gljučne besede: tehnična besedila, strojno prevajanje, samostalniške besedne zveze, prevodni premiki, prevajanje tehničnih besedil

1 Introduction

Technical translation is a specialized branch of translation studies that focuses on conveying technical content across languages. Despite its critical role in global communication, it has historically received less academic attention than other translation domains, even though it accounts for a sizeable portion of worldwide translation output (Kingscott 2002). As is the case with other types of text, many technical texts today are machine-translated. One of the questions this raises is how translation engines deal with the specific characteristics of technical texts such as the use of specialized terminology, lexical density, conciseness, or frequent use of passive voice. The purpose of the present study is to examine how machine translation engines deal with independent nominal phrases, which are common in technical texts, where data is presented in tables or images. The article has two parts: in the first, theoretical part, the major features of technical translation and machine translation are presented. In the second, empirical part, independent nominal phrases from a service manual in English are compared to their translations generated by two machine translation engines and analysed. The article ends by drawing conclusions from the analysis.

2 Technical Translation

Technical translation is a field of translation studies that focuses on texts with technical content. Although it is often referred to together with scientific translation (e.g., Olohan 2016), significant differences do exist between the two areas. The main characteristic of scientific texts is that they “*discuss, analyze and synthesize* information with a view to *explaining* ideas, *proposing* new theories or *evaluating* methods,” while technical texts are “designed to convey information as clearly as possible” (Byrne 2014, 2). Technical texts thus represent an applicative extension of scientific texts. From a research standpoint, it is notable that in comparison with some other fields of translation studies, this field has received little scholarly attention, given that technical translation is estimated to represent as much as 90% of global translation output (Kingscott 2002, 247). On the other hand, according to the BITRA bibliography of translation research, only 9.3% of publications address technical translation (Aixelá 2004).

In practical terms, technical texts refer to a variety of documents with technical content. These range from user manuals and expert technical reports written in narrative linear prose, on the one hand, to data sheets with tables, lists of nominal phrases and little context, on the other (cf. Byrne 2014, 58-73). In turn, technical writing features different textual and linguistic characteristics, depending on its purpose and target readers. One common observation is that the language of technical writing is expected to be clear, simple, and concise (Herman 1993, 11; Byrne 2014, 48). In contrast to literary texts, for example, technical texts thus typically do not abound in elements such as figures of speech, rhyme, or convoluted sentences; instead, technical writing is expected to be clear, objective, and unambiguous. Another characteristic that is directly or indirectly discussed in every treatise on technical translation (e.g., Galinski and Budin 1993; Byrne 2006; 2014; Olohan 2022) is terminology, which refers to a specialized subset of concepts and vocabulary that typify a particular subject area. Indeed, Pinchuck (1977, 19) claims that vocabulary is the most significant linguistic feature

of technical writing. Although Newmark (2008, 151) refuted that, claiming that terminology usually constitutes only 5-10% of the total content of technical texts, terminology remains an essential element of technical writing¹. In contrast to other types of texts, technical texts often also visually distinguish themselves through multimodality, given that they include diagrams, graphs or photographs to complement the verbal text (Byrne 2014, 54).

In addition to the use of passive voice or the prevalence of the present tense, one prominent linguistic feature of technical texts is nominalization (Newmark 2008, 151; Olohan 2022, 329). The frequency of nominal structures in technical texts is not surprising, given that technical writing strives for conciseness, and nominal structures deliver precisely that: lexical density. Regarding the discourse on science and technology, the phenomenon was analysed in detail by functional linguist Michael Halliday (2004). According to him, one result of the evolution of technical writing was that it helped organize grammar as a resource for generating meaning in metaphorical ways. This meant that items such as adjectives and verbs referring to “qualities” and “processes,” were first decoupled from their original lexical realizations, and then both meanings were recoupled through the new grammatical category of noun. One such example is the word *length*, which carries the quality of the adjective “long” (“quality”), but also belongs to the grammatical category of noun, i.e., the nominal meaning of “entity” or “thing.” Given that such words carry two category meanings, Halliday calls the phenomenon “grammatical metaphor.” The advantage of such structures is that it is possible to compress and combine multiple meanings into nominal phrases. On the other hand, this becomes a problem because overly concise and compressed structures sometimes become ambiguous or even incomprehensible (Byrne 2006, 83). One famous example is the phrase *lung cancer death rates*, which can mean anything from the number of deaths from lung cancer, on the one hand, to the amount of time in which patients with lung cancer die, on the other (Halliday 2004, 170). As Halliday’s example also shows, the problem of ambiguity is compounded when such phrases appear with little or no context, which is often the case with technical writing, which abounds in tables and illustrations.

In Slovenia, the field of technical translation in conjunction with machine translation remains under-researched, with most scholars focusing on other types of translation (e.g., Mezeg 2023; Orel Kos 2024).

3 Machine Translation

Machine translation (MT) automates the production of a target-language text from a source-language text. Over the decades, scientists have worked on various approaches to MT (for an overview, see Naveen and Trojovský 2024; Araghi and Palangkaraya 2024). Previously, the two most recognizable ones were Rule-Based MT and Statistical MT. In recent years, however, neural machine translation (NMT) has become the most promising new venue, utilizing models loosely inspired by the human brain, which employ artificial neural networks (see, for example, Zhang and Zong 2020). NMT translation involves two phases: encoding and decoding. During the encoding phase, each word in the source text is given a distinct neural

¹ In Olohan’s (2022) monograph *Scientific and Technical Translation*, for example, the terms *terminology* and *terminological* appear over one hundred times on 250 pages.

representation or embedding. The word embeddings are subsequently combined to form a sentence-level representation. This process modifies the individual representations based on context, resulting in a contextualized interpretation. During the decoding phase, the sentence-level representation is systematically broken down to produce the target sentence one word at a time. These two phases are carried out by interconnected artificial neural networks – the encoder and the decoder – together forming a unified network (Pérez-Ortiz et al. 2022).

NMT can generate non-existent target language words and fluent but inaccurate translations. Fluent output might mask these inaccuracies. Like other significant text dataset technologies, NMT can amplify training data biases. NMT systems require significant training time, computing power, energy, and specialized hardware (GPUs). They also demand massive training datasets unavailable for all language pairs (Kenny 2022). This reliance on automated solutions, however, may divert attention from the critical, in-depth analysis of the source text – a drawback also observed by Koletnik Korošec (2011), who noted that unstructured use of machine translation can undermine thorough source text evaluation.

The present research used two publicly available, multilingual neural machine translation services, DeepL Translator and Google Translate.

DeepL Translator, like most translation systems, employs artificial neural networks for text translation. These networks undergo training on extensive datasets comprising millions of translated texts. DeepL's website (How does DeepL Work 2021) reports numerous enhancements to the underlying neural network methodology. While most publicly available translation systems are direct modifications of the transformer architecture, and DeepL's networks incorporate elements of this architecture, such as attention mechanisms, significant topological differences contribute to a reported substantial improvement in translation quality compared to the public research state of the art. A strong focus is placed on the targeted acquisition of specialized training data to enhance translation quality. This involves the development of specialized crawlers designed to locate and automatically assess the quality of translations available online. While public research typically employs supervised learning for network training, where the network iteratively compares generated translations with training data translations and adjusts weights based on discrepancies, DeepL reportedly utilizes additional techniques from other machine learning domains to achieve notable improvements. Training is conducted on networks with many billions of parameters. Emphasis is placed on efficient parameter utilization, enabling comparable translation quality to be achieved with smaller, faster networks.

Two distinct language models are currently offered by DeepL for the translation of specific language pairs: a classic model and a next-generation model. The classic language model uses DeepL's established AI neural network architecture for translation and is available for all supported languages. Over 800 language combinations are currently possible, including Slovene. DeepL Translator also supports translations into British English and American English. The next-generation language model is powered by a large language model (LLM) infrastructure. This LLM leverages extensive multilingual text corpora to address complex problems and is specifically trained for translation. Using proprietary LLMs within the next-generation model improves translation quality, particularly for longer texts. Specialized

LLM infrastructure, uniquely tuned for language processing, facilitates more human-like translations, and reduces the risk of hallucinations and misinformation. Furthermore, unlike general-purpose models trained on publicly sourced internet data, DeepL's next-generation model benefits from over seven years of proprietary data curated for translation and content creation. Currently, however, the next-generation language model does not support Slovene ("About the Next," n. d.).

Google Translate is the second publicly available multilingual neural machine translation service used in the present research. Like DeepL, it offers a website interface, mobile applications for Android and iOS, and an application programming interface (API). As of February 2025, it supports 249 languages and language varieties at various levels. Launched in April 2006 as a statistical machine translation service, it gathered initial linguistic data from United Nations and European Parliament documents and transcripts. Rather than direct translation, texts were initially translated to English and pivoted to the target language for most supported language combinations. In September 2016, Google's research team announced the development of the Google Neural Machine Translation system (GNMT) to enhance fluency and accuracy. In November of the same year, Google Translate transitioned to GNMT. This system employed an extensive end-to-end artificial neural network utilizing deep learning. GNMT improved translation quality compared to statistical machine translation by employing an example-based machine translation (EBMT) method, learning from millions of examples. Whole sentences were translated at once rather than piecemeal. This broader context facilitated the identification of more relevant translations, subsequently rearranged and adjusted for improved grammatical accuracy and human-like fluency. Since 2020, GNMT has been phased out, and deep learning networks based on transformers have been implemented.

Despite advancements in automated translation, Google's engineers acknowledge that its quality remains imperfect, especially for low-resource languages. Even the latest models are susceptible to common machine translation errors, such as "poor performance on particular genres of subject matter (domains), conflating different dialects of a language, producing overly literal translations, and poor performance on informal and spoken language" (Caswell and Liang 2020).

4 Empirical Study

To evaluate machine translation, independent nominal phrases were compared with their respective machine translations. "Independent nominal phrases" in this paper refers to phrases that meet the following two criteria: 1) they have nouns as their heads; and 2) they are not an integral part of any sentence but instead appear on their own, outside any (explicit) syntactic structure, in technical texts typically in tables and illustrations. The text used in the analysis was a service repair manual for the diesel and gasoline Caterpillar forklifts of the GP and DP 15K, 18K, 20K, 25K, 30K, 35K series (Pub. No. 99719-60120), which were produced between the mid-1990s and 2007 (Caterpillar LPG n. d.)². The original text was in English and was available in electronic form as a readable pdf document. The source

² The authors want to thank Darko Rihard and Marko Fajfar from Vilfis d.o.o. for their help with forklift truck-related terminology.

texts were not additionally pre-formatted before translation. The complete manual comprised 384 pages. For the study, the first thirty pages of the manual were machine-translated into Slovene using Google Translate (GT) and the professional (subscription-based) version of DeepL (DL). Next, the first one hundred subsequent independent units comprising nominal phrases from an illustration and a table on pages 1-2 to 1-5 were extracted and aligned with their two machine translations. Excluded were repetitions of identical phrases with identical translations. Most translation units were simple nominal phrases with single noun heads (e.g., *front axle*), including single-word phrases (e.g. *mast*), while a small number of other units comprised sets of (appositional) nominal phrases separated by parentheses, slash or colon (e.g., *Kg/mm (lb/in.)*; *tread (front/double tires)*; *applicable truck model designation 35: 3.3 ton class*). These examples were treated as single translation units because they functioned as one unit of meaning. The original text included a few typos and grammatical errors, which remained uncorrected because the idea was to see how translation engines would deal with these. In total, the corpus of one hundred source units in English comprised a total of 305 words and 139 lemmas (e.g., *truck*, *trucks* are two words, the base form of which (*truck*) corresponds to one lemma)³. Some phrases reappeared in identical form in the corpus several times: the most frequent, for example, was *serial number*, which recurred six times; four phrases (e.g., *simplex mast*; *duplex mast*) appeared four times; the rest had fewer recurrences. All one hundred source units were compared to their corresponding translations generated by the two translation engines and evaluated qualitatively and quantitatively.

In the absence of a specific model for describing translation shifts in acontextual nominal phrases, descriptors were adapted from other translational models and theories, such as those by Leuven-Zwart (1989; 1990); Klaudy and Károly (2005); Toury (2012); and Krüger (2015). Following a preliminary comparison and analysis of various types of nominal phrases, the following descriptors were used to describe the relationship between the source and target translation units generated by the same translation engine:

No shift. Source and target phrases have equal or near-equal semantic, formal, and functional properties. Examples include phrases such as *general information*, whose Slovene translation *splošne informacije* is considered both a formal and functional equivalent of the English phrase. Other examples include target phrases that have several possible lexical varieties (e.g., *vrsta motorja* or *tip motorja* for *engine type*), all of which are considered adequate.

1. **Semantic shift.** The semantic gap is too large to infer the meaning of the source phrase based on the translation. A typical example is the phrase *duplex mast*, which was translated as *dvostranski jambor* in Slovene. Although the head noun *jambor* corresponds to the English noun *mast*, it does so only in the context of sailboats; in the context of heavy machinery, however, the correct technical term in Slovene is *jarem* or *teleskop*. In addition, the adjective *duplex*, referring to the two stacks or sections of the mast that can be extended vertically, is translated as *dvostranski*. i.e., as two-sided, which likewise is a mismatch with the original meaning. The category also includes examples of made-up translations, which in the context of artificial intelligence are popularly referred to

³ The corpus was analysed with Sketch Engine (<http://www.sketchengine.eu>).

as hallucinations. One such interesting example is the word *underclearnace* (sic). As can be seen, the original technical term is misspelled and should have been spelled as *underclearance*, referring to the physical distance between the frame of the forklift and the ground below it. The translation engine, however, “translated” the original phrase as *podnaprava* in Slovene, which practically is an nonexistent noun in the Slovene lexicon, given that only one or two references could be found in relation to the *sub-installation*, but even those were from an unrelated domain of emission allowances.

2. **Terminological shift.** Source and target phrases overlap semantically to the extent that the meaning of the original phrase can be inferred; however, the term used is a general or non-standard expression and not an established or standard technical term. An example of this type of shift is the phrase *single wheels*, referring to number of parallel wheels at the same end of the forklift truck axis. While GT used the correct technical equivalent *enojna kolesa* in Slovene, DL translated the phrase as *posamezna kolesa*, whose back-translation is *individual wheels*. Although *posamezna kolesa* could apply in other contexts and makes it possible to decipher its meaning, the correct technical term to be used in this context is *enojna kolesa*. Another case in point involves examples of poor style, although the adverb *nazaj* is used. The category also includes examples of poor style such as *poimenovanje*, which appeared as the DL equivalent of the English phrase *designation*, referring to the type of forklift truck; a stylistically and terminologically better translation in Slovene would have been *oznaka*.
3. **Grammatical shift.** Source and target phrases have different grammatical features. Given that English is an analytic language and Slovene as the target language a synthetic one with several inflectional morphemes, target phrases are expected to deviate from the source ones grammatically; also possible are grammatical disagreements (e.g., number, gender) within target phrases. One such example is the English phrase *minimum intersecting isle*, which was translated as *najmanjša otok*, where the feminine suffix -a in the attributive adjective disagrees with the masculine head noun *otok* – the morphologically correct version of the phrase would be *najmanjši otok*.
4. **Orthographic shift.** Target text features orthographic shifts such as incorrect hyphenation, capitalization etc. One example is the abbreviation *Ref. No.*, where both abbreviated words are capitalized in English. The first letters in the corresponding Slovene translation *Ref. Št.* likewise are capitalized; however, this conflicts with the rules of Standard Slovene, according to which lower case should have been used in the second abbreviation.
5. **Terminological inconsistency.** The same source phrase is translated in various ways in the target texts. Included is every first iteration of a different translation. One such example in English is the phrase *simplex mast*, which in Slovene appears in three different translations: *enostavni jambor*; *dvostranski* (sic) *drog*; and as *simplex jambor*.
6. **No translation.** In a small number of examples, no translation was provided, and the original source text phrase was reproduced in the target text, e.g.: the phrase *[Mast]* (square brackets used in the original text) appears as *[Mast]* also in Slovene (where the corresponding technical term is *teleskop*).

The descriptors were not discrete categories excluding one another. In a small number of units, two descriptors were used for the same translation. The target phrase *najmanjša otok*, for example, included a semantic shift as well as a grammatical error because of a gender disagreement between the adjective and the noun in Slovene translation. Similarly, the phrase [*podvozje*] was an example of an adequate translation and thus marked as “no shift;” however, it was marked for terminological inconsistency, because the previous iteration of the same original phrase (*chassis*) was machine-translated by the same engine as *šasija*.

5 Results

Overall, the results showed that in most categories, both engines produced translations of comparable quality. No translation shifts were observed in 46% of translations generated by DL, while GT performed slightly better with 49% of units with no translation shifts. On the other hand, both engines generated a similar proportion of translations with semantic shifts: 41% in DL vs. 42% in GT. GT performed slightly better in terms of terminological shifts, which were observed in 8% of all translations, while in DL translations, that proportion was 22%. In contrast, no grammatical shifts were observed in DL generated phrases, while 4% of GT contained grammatical errors. Orthographic shifts were observed in 4% of both DL and GT translations. The only category with more noticeable discrepancies was terminological inconsistency, where no inconsistencies were observed in DL, whereas in GT, in 3% of the units the same technical term was translated in two or more ways. Only one percent of units remained untranslated by both translation engines. The distribution is presented in Table 1.

TABLE 1. Distribution of translation shifts.

	No shift (%)	Semantic shift (%)	Terminological shift (%)	Grammatical shift (%)	Orthographic shift (%)	Terminological inconsistency (%)	No translation (%)
DeepL	46	41	12	0	4	0	1
Google Translate	49	42	8	4	4	3	1

As is the case with all quantitative data, the numbers show only part of the picture. The following is a detailed qualitative analysis of examples included in each of the seven categories.

5.1 No Shift

As indicated by the relative values, just under half of all translation units in the corpus showed no shifts, meaning that the phrases were considered functional equivalents of their source phrases. The analysis showed that this group of translation units could be divided into two subgroups. The first featured phrases which are considered common and are widely used in other more general contexts. A typical example is the phrase *General information*, which appeared four times in the source English text. In both translations, it was consistently

translated as *Splošne informacije*. The absence of shifts comes as no surprise, given that this is the standard version of the phrase in Slovene, which appears in a variety of contexts even beyond technical writing, as indicated by over 600 instances of the phrase in the largest corpus of Written Standard Slovene, Gigafida 2.0. The same goes for other common phrases that also appear in non-technical writing such as *serial number* or *dimensions*, whose respective Slovene equivalents *serijska številka* and *dimenzije* also relatively frequently appear in non-technical texts (cf. Gigafida 2.0). The engines also successfully dealt with a few terms that were considered more technical, such as *output shaft*, a mechanical part that connects the drive wheels and the gearbox, which was translated as *izhodna gred* by both DeepL and GT. However, it should be noted that neither of the translation engines translated the phrase as *odgonska gred*, which is another technical equivalent for the same mechanical part in Slovene.

5.2 Semantic Shifts

As evident from Table 1, over 40% of units in both translations included target text phrases whose meaning deviated from the source phrase to the extent that it made their understanding practically impossible. The analysis showed that most of these radical shifts were the result of two factors: 1) the ambiguity of phrases whose meaning is context-dependent; 2) the properties of technical language and terminology that typically are not part of the general vocabulary. In most cases, the translation engines struggled with the same units--however, not always.

A typical example of a unit whose meaning is context-dependent is the source noun *truck*. In the source manual, the noun is consistently used in the shorter, elliptical form of the longer phrase *forklift truck*, a standard expression for this type of industrial vehicle. However, in Slovene, the elliptical form poses a challenge for translation engines, given that the equivalent of *truck* in Slovene is *tovornjak*, which in turn is unrelated to forklift trucks; instead, *tovornjak* is a standard general Slovene term for a specialized vehicle for transporting freight. In translation, the meaning of the source unit thus changed radically. It is also interesting that in those sections of the English manual, where the complete phrase *forklift truck* was used, neither of the two translation engines had difficulties and consistently translated it as *viličar*, which is the Slovene equivalent of the phrase *forklift truck*; however, once the elliptical form appears, both translation engines struggled with its interpretation, regardless of the fact that the phrase appeared in longer full form elsewhere in the text. Lack of context also posed a problem in the one-word phrase *reverse*, which in this case referred to the travel speed for driving backwards. Although *obratno*, used in GT is a lexical meaning of the adverb, it does not fit the context; instead, the adverb used should have been *vzurnatno*, as correctly identified by DL.

A similar problem appears with the source phrase *free lift*, referring to “the distance a forklift operator can raise the forks without extending the mast” (“What is a free lift,” n. d.). In both machine translations, however, the phrase turned into *brezplačno dvigalo*, which could be backtranslated as a *free elevator* and obviously bears no relation to the source phrase. The problem arose because the engines seemingly built the translation based on the headword *lift* in the source text. One of its lexical meanings is elevator; in turn, this most likely led to the use of an incorrect premodifier *brezplačno*, the Slovene lexical equivalent of the adjective

free, i.e. something requiring no monetary compensation. In the given context, of course, the phrase is out of place. Similarly, *work performance*, describing the properties of the truck, was translated incorrectly by both engines as *delovna uspešnost*; the latter is an established phrase in Slovene; however, not in the context of machinery but rather in labour relations. A similar problem appeared in relation to the English phrase *transmission serial number*, which seems straightforward as a designation of the serial number of the assembly connecting the engine and wheels. Neither of the translation engines struggled with the common phrase *serial number* and adequately translated it as *serijska številka*. However, both misinterpreted the premodifying noun *transmission*. It is true that one of its lexical meanings is that of *prenos* (in the sense of a *transfer*), which is used by both translation engines, but in this case, the meaning becomes misplaced, given that the phrase relates to the engine part.

5.3 Terminological Shifts

Terminological shifts were the third most common category of translation shifts. In contrast to semantic shifts, this category included translations that can be understood by readers but are terminologically inadequate because of failure to comply with standard technical terminology. Both translation engines had comparable results in this category, with DL outperforming GT by a small margin. As was the case with semantic shifts, terminological shifts in both translation engines commonly appeared in one-word phrases whose meaning was highly context dependent. An example of this the word *items*, which appeared on top of a column referring to the technical specifications of the forklift truck that are presented in the table. In DL, the noun was translated as *elementi* and in GT as *predmeti*. Although both are lexical equivalents of *items*, neither of the two translations fits the context; *postavke* would have been better. Similarly, *poimenovanje* (DL) and *imenovanje* (GT) are both close to one of the source phrase lexical meanings of *designation*; however, *oznaka* is considered a more adequate technical translation.

But multi-word units also posed a challenge. With some, the discrepancy was less noticeable than with others. A case in point is the phrase *disassembly diagram*, referring to the diagrams in the manual that show the order or relationship in which parts are disassembled. Both engines translated the phrase as *diagram razstavljanja*, whose meaning is likely to be clear to most speakers of Slovene, although the established technical term in Slovene is *shema razstavljanja*. Another example of a phrase that the engines struggled with was *travel speed*. As the translation *hitrost potovanja* shows, the confusion likely arose from the noun *travel*, whose basic lexical meaning in Slovene is that of *potovanje*; however, in this context, the resulting phrase, in conjunction with the head noun *hitrost*, refers more to the pace at which tourists enjoy their travels.

5.4 Grammatical Shifts

Unsurprisingly, there were few grammatical shifts. None were observed in the DL translations, and only four in the GT translations. One of those was an example of gender disagreement between the headword and its premodifier (*najmanjša otok*), while two displayed a grammatical case mismatch (*obremenitev porazdelitev* instead of *porazdelitev obremenitve* and *Powershift menjalnik modeli* instead of *modeli z menjalnikom Powershift*). The last shift featured

the longer phrase *overall height (to top of mast lowered)*, where GT failed to incorporate the participle *lowered* into the translation; in turn, the resulting phrase featured an incorrect use of the participle in the postmodifying position: *skupna višina (do vrha jambora spuščen)*.

5.5 Terminological Inconsistencies

Terminological inconsistencies were also infrequent. In DL, none were observed, while GT featured three units where the same term was rendered in various ways in the translation. The first involved the noun *disassembly*, which appeared in three different phrases. In the first two, *disassembly diagram* and *disassembly sequence*, the noun was translated as *razstavljanje*; however, in the third iteration, *suggestions for disassembly*, the same noun appeared as *demontaža*, which typically is the standard technical term for the procedure described. The second term was *chassis*. When this noun appeared as part of a phrase, it was translated as *šasija*, which is an established technical term for *chassis* in Slovene. It is notable that in the third iteration, the noun appeared on its own as a single-word phrase and was translated as *podvozje*, which is a synonym for *šasija* in Slovene.

The third and most notable example was the noun *mast*, referring to one of the main forklift parts, the mechanical implement for lifting or lowering the load at the front of the vehicle. In the original, it appeared as part of fifteen different phrases. The first of those iterations was the *mast serial number*, in which *mast* is translated as *jambor* by both engines. Although *jambor* is a Slovene lexical equivalent of *mast*, it only applies in the context of sailboats; in forklifts, the corresponding technical term is *jarm* or *teleskop*. On the same page in the manual, *mast* also appears in the phrase *chassis and mast model identification*. In this instance, GT translated the noun as *drog*. In the remaining iterations in the GT translation, the noun varied again between *drog* and *jambor*; in DL, it was consistently translated as *jambor*.

5.6 No Translation

Both translations included only one unit that remained untranslated by both translation engines: the word [*mast*]. One plausible reason for this was the square brackets, which may have confused the engines and had them mistake the brackets for part of machine language.

Another group of items that remained untranslated were imperial units of measurement, which accompanied metric units of measurement in brackets, e.g., *mm (in.)*. Given that metric units are standard in Slovene, the use of imperial alongside them was acceptable.

5.7 Miscellaneous

An interesting example includes translation of words with typos, which is not an uncommon phenomenon in texts. The case in point is the phrase *underclearnace (at frame)*, referring to the distance between the chassis and the ground. DL translated it as *podnaprava (v okvirju)*. Although the word sounds feasible in Slovene in terms of its form and morphological characteristics, it is not a common word and is hence an example of hallucination. Misinterpretation by DL manifests itself in the prepositional phrase *at frame*, which indicates the point at which the distance from the ground is measured, whereas Slovene translation

places that same point *inside* the frame, suggesting that the engine misinterpreted it. What is also notable that the same word remained untranslated by GT.

6 Conclusions

This study highlights the challenges of machine translation in handling independent nominal phrases in technical text. The comparison of Google Translate and DeepL into Slovene revealed both their strengths and their limitations in dealing with specialized terminology.

Nearly half the translated units showed no shift, indicating that common phrases were adequately rendered. However, semantic shifts were prevalent (over 40%), often due to ambiguity and lack of contextual information. Key issues included mistranslation of elliptical forms (e.g., *truck* instead of *forklift truck*) and misinterpretation of industry-specific terms like *free lift* and *transmission serial number*.

Terminological shifts affected precision, with general expressions replacing technical terms. While these translations were understandable, they lacked standard industry accuracy. Grammatical and orthographic shifts were minimal, with DeepL producing no grammatical errors and Google Translate showing minor inconsistencies. However, terminological inconsistencies in Google Translate indicated weaker consistency mechanisms compared to DeepL.

A small number of untranslated units, such as *mast*, suggest formatting-related processing issues in machine translation engines. The study underscores the importance of integrating domain-specific resources and human post-editing, as well as pre-editing and pre-formatting of texts to enhance translation reliability. This observation aligns with the findings of Hazemali et al. (2024), whose evaluation of chatbot performance in reading digitized texts showed that while the chatbot used in the study exhibited some success in handling typos and minor language errors, it produced only a 20% success rate in tasks demanding deeper language comprehension and struggled with complex sentence structure and domain-specific terminology in Slovene. Experienced human translators typically do not miss phenomena such as repetition in the immediate textual vicinity; in addition, humans can also process graphic representations of information. While both translation engines performed comparably, future improvements should focus on context recognition and the handling of specialized terminology. These findings, alongside evidence from studies in other domains – such as Mohar, Orthaber and Onič (2020), who demonstrated that machine translation quality deteriorates with increasing sentence complexity in literary texts – underscore the need for ongoing refinement of MT systems to better handle both technical and stylistically rich content. It should be noted that technical translators, in contrast, for example, with literary translators, strive above all towards precision and comprehensibility “since the consequences of lexical error, however slight, are more serious: a poor literary translation leads to a dissatisfied reader, whereas a misleading technical translation could result in a hazard to human life” (Hann 1992, 7).

Further research should explore the impact of context on machine translation accuracy and investigate AI-driven enhancements for better translation consistency and precision.

References

- “About the next-generation language model.” n. d. DeepL Help Center. <https://support.deepl.com/hc/en-us/articles/14241705319580-About-the-next-generation-language-model>.
- Aixelá, Franco, J. 2004. “The study of technical and scientific translation: An examination of its historical development.” *Journal of Specialised Translation* 1. https://jostrans.soap2.ch/issue01/art_axela.php.
- Araghi, Sahar, and Alfons Palangkaraya. 2024. “The link between translation difficulty and the quality of machine translation: A literature review and empirical investigation.” *Language Resources & Evaluation* 58: 1093–1114. <https://doi.org/10.1007/s10579-024-09735-x>.
- Byrne, Jody. 2006. *Technical Translation: Usability Strategies for Translating Technical Documentation*. Springer.
- . 2014. *Scientific and Technical Translation Explained*. Routledge.
- Caswell, Isaac, and Bowen Liang. 2020. “Recent advances in Google Translate.” *Google Research Blog*, June 8. <https://ai.googleblog.com/2020/06/recent-advances-in-google-translate.html>.
- “Caterpillar LPG forklifts specifications.” n. d. *Lectura specs*. <https://www.lectura-specs.com/en/specs/forklifts/lpg-forklifts-caterpillar>.
- Galinski, Christian, and Gerhard Budin. 1993. “New trends in translation-oriented terminology management.” In *Scientific and Technical Translation*, edited by Sue Ellen Wright and Leland D. Wright, Jr. 209–16. John Benjamins.
- Gigafida 2.0: Corpus of Written Standard Slovene. <https://viri.cjvt.si/gigafida>.
- Halliday, M.A.K. 2004. *The Language of Science*, edited by J. J. Webster. Continuum.
- Hann, Michael. 1992. *The Key to Technical Translation. Volume 2: Terminology/Lexicography*. John Benjamins.
- Hazemali, David, Janez Osojnik, Tomaž Onič, Tadej Todorović, and Mladen Borovič. 2024. “Evaluating chatbot assistance in historical document analysis.” *Moderna arhivistika* 7 (2): 53–83. <https://doi.org/10.54356/ma/2024/biub3010>.
- Herman, Mark. 1993. “Technical translation style: Clarity, concision, correctness.” In *Scientific and Technical Translation*, edited by Sue Ellen Wright and Leland D. Wright, Jr., 11–20. John Benjamins.
- “How Does DeepL Work? #Network Architecture.” 2021. *DeepL Blog*, November 1. https://www.deepl.com/en/blog/how-does-deepl-work#network_architecture.
- Kenny, Dorothy. 2022. “Human and machine translation.” In *Machine Translation for Everyone: Empowering Users in the Age of Artificial Intelligence*, edited by Dorothy Kenny, 24–46. Language Science Press.
- Kingscott, Geoffrey. 2002. “Technical translation and related disciplines.” *Perspectives* 10 (4): 247–55. <https://doi.org/10.1080/0907676X.2002.9961449>.
- Klaudy, Kinga, and Krisztina Károly. 2005. “Implication in translation: Empirical evidence for operational asymmetry in translation.” *Across Languages and Cultures* 6 (1): 13–28. <https://doi.org/10.1556/Acr.6.2005.1.2>.
- Koletnik Korošec, Melita. 2011. “Applicability and challenges of using machine translation in translator training.” *ELOPE: English Language Overseas Perspectives and Enquiries* 8 (2): 7–18. <https://doi.org/10.4312/elope.8.2.7-18>.
- Krüger, Ralph. 2015. *The Interface between Scientific and Technical Translation Studies and Cognitive Linguistics with Particular Emphasis on Explicitation and Implication as Indicators of Translational Text-Context Interaction*. Frank & Timme.
- Leuven-Zwart, K. M. v. 1989. “Translation and original: Similarities and dissimilarities I.” *Target* 1 (2): 151–81.
- . 1990. “Translation and original: Similarities and dissimilarities II.” *Target* 2 (1): 69–95.
- Mezeg, Adriana. 2023. “Ali sploh še potrebujemo prevajalce? Strojno prevajanje iz francoščine v slovenščino.” *Ars & Humanitas* 17 (1): 139–54. <https://doi.org/10.4312/ars.17.1.139-154>.
- Mohar, Tjaša, Sara Orthaber, and Tomaž Onič. 2020. “Machine translated Atwood: Utopia or dystopia?” *ELOPE: English Language Overseas Perspectives and Enquiries* 17 (1): 125–41. <https://doi.org/10.4312/elope.17.1.125-141>.

- Naveen, Palanichamy, and Pavel Trojovský. 2024. "Overview and challenges of machine translation for contextually appropriate translations." *iScience* 27 (10): 110878. <https://doi.org/10.1016/j.isci.2024.110878>.
- Newmark, Peter. 2008. *A Textbook of Translation*. Twelfth impression. Longman.
- Olohan, Maeve. 2016. *Scientific and Technical Translation*. Routledge.
- . 2022. "Translating technical texts." In *The Cambridge Handbook of Translation*, edited by Kirsten Malmkjær, 321–39. Cambridge University Press.
- Orel Kos, Silvana. 2024. "Introduction of machine translation into audiovisual translation teaching." *ELOPE: English Language Overseas Perspectives and Enquiries* 21 (1): 185–208. <https://doi.org/10.4312/elope.21.1.185-208>.
- Pérez-Ortiz, Juan Antonio, Mikel L. Forcada, and Felipe Sánchez-Martínez. 2022. "How neural machine translation works." In *Machine Translation for Everyone: Empowering Users in the Age of Artificial Intelligence*, edited by Dorothy Kenny, 141–64. Language Science Press.
- Pinchuck, Isadore. 1977. *Scientific and Technical Translation*. André Deutsch.
- Toury, Gideon. 2012. *Descriptive Translation Studies and Beyond*. Benjamins.
- "What is a free lift on a forklift?" n. d. *American Forklifts*. <https://americanforklifts.org/what-is-a-free-lift-on-a-forklift/>.
- Zhang, JiaJun, and Chengqing Zong. 2020. "Neural machine translation: Challenges, progress and future." *Science China Technological Sciences* 63: 2028–50. <https://doi.org/10.1007/s11431-020-1632-x>.