Challenges of machine translation application to teaching ESP to construction students

Rusudan Makhachashvili¹, Larysa Mosiyevych² and Tetiana Kurbatova³

¹Borys Grinchenko Kyiv University, 18/2 Bulvarno-Kudriavska Str., Kyiv, 04053, Ukraine
²Zaporizhzhia National University, 66 Zhukovskoho Str., Zaporizhzhia, 69600, Ukraine
³Kryvyi Rih National University, 11 Vitalii Matusevych Str., Kryvyi Rih, 50027, Ukraine

Abstract. The article deals with problems of translation teaching, namely machine translation for ESP students. The study aims to conduct a comparative analysis of machine and human translation of construction terminology, identify causes of errors, provide recommendations for improving quality of students’ translation via post-editing as well as developing their interdisciplinary skills using CAT tools. The main research methods include comparative and contrastive analysis as well as the quantitative method. The research material is presented by titles of construction students’ qualification papers translated from Ukrainian into English. Quality of machine translation is affected by peculiarities of construction terminology due to harmonization of terms according to Eurocodes. The authors prove that application of software to translating construction texts without further proofreading or post-editing by students themselves entails errors including distortion of terminology and, consequently, meaning. There are presented reasons for lexical errors caused by peculiarities of translating multicomponent terms, discrepancies in translating prepositions. The article is intended for a wide range of specialists interested in translating construction texts and teaching ESP. Based on the results obtained, the authors develop recommendations for translating construction texts by using machine translation accompanied by students’ post-editing.

Keywords: translation teaching, machine translation, ESP students, construction terminology, post-editing

1. Introduction

Technical translation issues remain essential for teaching English for specific purposes (ESP) at technical universities of Ukraine. At the age of computer technologies, construction students learning the English language are accustomed to using machine translation, considering the increased volume of information in their professional area due to ongoing intensification of their industry worldwide.

Using translation in ESP teaching and learning is an important issue but it has been very
scarcely researched into so far because of the negative connotations with the grammar-translation method. Since the last decade of the 20th century, revival of interest in the use of translation and use of L1 in foreign language teaching has been noticeable. Translation activities consolidate reading, writing, speaking and listening skills of ESP learners, facilitate communication and comprehension in a specific field, develop analytic skills and teach to focus on accuracy [10]. In fact, translation has never lost importance in practical situations of teaching, especially in the context of ESP, at least as a supplemental method [2]. Translation has started to be seen as a useful tool in teaching ESP [8]; the use of L1 in ESP instruction has also started receiving considerable attention of researchers [11].

The study focuses on CAT (computer-assisted translation) application issues. They imply the use of technology, such as computer software, hardware, networks, and other auxiliary equipment in the translation process [3, 16, 17]. CAT tools do not typically translate the text; they assist the translator in various tasks, such as verification of terminology consistency, source and target text alignment, reuse of previously translated documents, grammar and spelling checking, pre-translation activities, terminology management, proper document formatting, document production, post-editing, etc. [7]. The study deals with developing the CAT competence among ESP students via an indispensable technique – a comparative analysis of machine and human translation with subsequent post-editing which has become an increasingly common practice in the translation field in recent years [9].

Besides general errors of machine translation, its quality is greatly affected by peculiarities of construction terminology due to harmonization of construction terms in compliance with Eurocodes, European standards specifying rules for the design of buildings and other engineering works and products. Among other term systems, construction terms are noted for appearance of new words and word combinations to denote new materials, processes, tools, etc., this making this terminological system extremely dynamic.

The article aims to conduct a comparative analysis of machine and human translation of construction terminology from Ukrainian into English, identify causes of machine translation errors, provide recommendations for improving quality of students’ translation via post-editing as well as developing their interdisciplinary skills using CAT tools. We also suppose that the CAT competence includes all the components of the language competence along with some specific sub-competences: content-knowledge competence, research competence, and monitoring competence.

The research aim entails solving the following tasks:

1) to classify common errors in machine translation;
2) to compare the number of errors made after translating titles of students’ master qualification papers by means of Google Translate and Prompt;
3) to perform human translation using specialized dictionaries and compare it with the machine one;
4) to study the problem of construction terms harmonization;
5) to perform a quantitative analysis of errors in machine and human translation;
6) to develop recommendations for translating construction texts by using machine or computer-aided translation;
7) to identify a relationship between specific machine translation problems and students’ post-editing efforts.

**Hypothesis**: On the basis of the above-mentioned tasks, we expect that a decrease in machine translation quality will lead to an increase in human translation quality.

### 2. Methods

In order to achieve the research aims and fulfil the set tasks, the following research methods are applied: comparative analysis of machine and human translation; contrastive analysis to observe changes in translation on the grammatical and lexical levels, selection of correct equivalents and ways of translating multicomponent terms. The quantitative method is used for calculating frequency of terminological and linguistic errors in machine and human translation.

Master degree students of construction specialities of the Engineering Institute of Science and Education, Zaporizhzhia National University, were engaged in the pedagogical experiment. Twenty students translated fifty titles of master qualification papers by using Google Translate and Prompt followed up by manual translation with dictionaries. There are several reasons for selecting the actual material:

1) easy practice of post-editing on individual sentences;
2) high concentration of multicomponent terms within a sentence;
3) a large number of of-phrases adding grammatical difficulties;
4) no possibility of using translation transformations due to conceptual restriction.

At the next stage of our research, terminological and linguistic errors in machine and human translation are compared. Analysis of data obtained enables developing recommendations for teaching ESP to construction students in order to develop their CAT and general digital competences as well as interdisciplinary skills, thus enhancing future specialists’ readiness to apply the acquired knowledge and skills to their professional activities.

### 3. Results and discussion

#### 3.1. Machine translation errors

Looking into previous studies on the use of machine translation for educational purposes, we can distinguish three main concepts:

1) use of machine translation as a bad model [9];
2) use of machine translation as a good model [14];
3) vocational use including translation quality assessment, pre- and post-editing [1, 13].

In our research, we consider the third concept implying that machine translation tools should be combined with post-editing procedures.

At this stage of the research, we identify errors in Ukrainian-English translation of construction terms by means of CAT tools – Google Translate and Prompt. It should be noted that
the peculiarity of Google Translate is the fact that it does not perform direct translation from Ukrainian into English: it first translates into Russian, and then into English. This reduces accuracy and quality of translation.

Analysis of literature sources \cite{4, 12, 18} reveals the following machine translation errors:

- content-based errors (distortion, incompleteness or redundancy of data, fragments left untranslated, loss of consistency);
- terminological errors (violation of terminological integrity, non-conformity with a glossary, dictionaries, industry-related standards of terms);
- linguistic errors (grammatical, lexical, spelling, punctual);
- stylistic errors (stylistic non-conformity of linguistic means, sentence structures, fixed phrases with a given genre, inconsistency of a text).

Error analysis consists in identification and classification of individual errors in a machine translated text \cite{15}. Translation errors are an objective indicator of poor translation quality. To perform the given analysis, we have selected 50 titles of master qualification papers concerning the construction industry. Their lingual features condition only two typical errors in students’ translations – terminological and linguistic. The quantitative analysis reveals that linguistic errors make about 41%, while those in terminology – 59%. Terminological errors prevail due to the following:

1) high concentration of terms within a small context;
2) complexity of translating construction terminology due to Eurocodes. There are fewer linguistic errors because there are no verb phrases in the titles under study.

Let us consider causes of terminological errors made by Prompt and Google Translate programmes. There are also examples of students’ translations by means of Terminological Dictionary for Translation of Eurocodes \cite{5, 6} (table 1).

Let us consider causes of language errors usually associated with translation of multicomponent terms. English multicomponent terms with a typical left-oriented deployment are characterized by dependent components in the left position regarding the main component. On the contrary, Ukrainian multicomponent terms are noted for right-oriented deployment. This fact should be taken into account when translating multicomponent terms.

English grammar rules require not to use the preposition *of* more than three times in the same phrase. Therefore, when translating three- (or more) component terms in the Genitive case, the rule should be observed when the main component in Ukrainian terminology (which is usually the first one) should occupy the final position in the English equivalent. In general, the structural scheme of the terminology group in the English word combination can be presented in the following form:

\[ LPA_n \leftarrow \cdots \leftarrow LPA_2 \leftarrow LPA_1 \leftarrow HW, \]

where *HW* is the head word, *LPA_1*, *LPA_2*, ..., *LPA_n* are one or more left-position attributes that clarify the meaning of the whole term. Examples of machine translation for titles 1 and 2 in table 2 do not harmonize with the translation scheme of multicomponent terms, so the
Table 1
Comparative analysis of some terminological errors in machine and human translation.

<table>
<thead>
<tr>
<th>Paper title (Ukrainian)</th>
<th>Translation by Google Translate</th>
<th>Translation by Prompt</th>
<th>Human Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Удосконалення технології нанесення тюркет-бетону на горизонтальні поверхні.</td>
<td>Improving the technology of applying shotcrete on horizontal surfaces</td>
<td>Improvements of technology of drawing torkret-concrete on horizontal surfaces</td>
<td>Improvement of the technology of applying guinite/shotcrete on a horizontal surface</td>
</tr>
</tbody>
</table>

preposition of makes them clumsy. The third example indicates that the Genitive case in the Ukrainian-language terminology can be translated not only by the preposition of, but also by in if the meaning of the phrase is not lost. Analysis reveals that students are capable of avoiding these errors in human translation.

It is an efficient technique to practice this task starting from analysis of oral English-Ukrainian translation of multicomponent terms. Students should remember that in most cases the final component in an English terminological phrase becomes the first one in Ukrainian translation,
### Table 2
Comparative analysis of some linguistic errors in machine and human translation.

<table>
<thead>
<tr>
<th>Paper title (Ukrainian)</th>
<th>Translation by Google Translate</th>
<th>Translation by Prompt</th>
<th>Human Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Науково-технічний супровід будівництва монолітичних будівель в сучасних умовах.</td>
<td>Scientific and technical support for the construction of monolithic buildings in modern conditions.</td>
<td>Scientific and technical maintenance of construction of monolithic buildings in modern conditions.</td>
<td>Scientific and technical support of monolithic buildings construction in current conditions.</td>
</tr>
<tr>
<td>2. Застосування детермінованої сітьової моделі для оптимізації організаційних процесів будівельно-монтажних робіт на будівництві адміністративної будівлі.</td>
<td>Application of a deterministic network model for optimization of organizational processes of construction and installation works for the construction of an administrative building.</td>
<td>Applications of the model determined network for optimization of organizational processes of installation and construction works on construction of the office building.</td>
<td>Application of a determined network model for organizational processes optimization in construction and installation works to the administrative building construction</td>
</tr>
</tbody>
</table>

*e.g.* gas heating system – система газового опалення, main patching materials – матеріали для якмового ремонту.

It is worth mentioning that students are willingly engaged in the creative process of post-editing. Knowing the structural scheme of English multicomponent terms, they easily replace the *of*-phrase by a multicomponent term.

Let us exemplify machine and human translation by translating the following paper title: “Аналіз іноземного досвіду вибору технологій та матеріалів утеплювання фасадів будівель”:

a) machine translation (Google Translate) – “Analysis of foreign experience in the choice of technologies and materials for insulation of building facades”,
b) human translation – “Foreign experience analysis of the technologies and materials selection for building facades insulation”.

As can be seen in human translation the first components in *of*-phrases become the final ones that enables reducing the number of prepositions *of*.

The other type of linguistic errors is associated with synonyms. Comparative analysis of Prompt and Google Translate programmes reveals that sometimes they give synonyms for one and the same word. It is worth mentioning that synonyms mainly relate not to construction terms, but to general English words. Unfortunately, students are not used to consulting monolingual dictionaries which are indispensable while distinguishing synonyms.

For example, the paper title “Особливості реконструкції існуючої будівлі гуртожитку по вулиці М. Гончаренко в м. Запоріжжі” is translated by Google Translate as “Features of
The reconstruction of the existing dormitory building on M. Honcharenko Street in Zaporizhzhia” and as “Peculiarities of the reconstruction of the existing dormitory building on M. Honcharenko Street in Zaporizhzhia” by Prompt. Taking into account the meanings of nouns features (one’s structure or make-up; form, shape, bodily proportions) and peculiarities (the quality or state of being peculiar; individuality; singularity) the decision is in favour of the word features.

The paper title “Дослідження усунення кренів будівель і споруд в Україні” is translated by Google Translate as “Research of elimination of rolls of buildings and constructions in Ukraine” and as “Investigation of the removal of cranes of buildings and structures in Ukraine” by Prompt. Research is diligent inquiry or examination to seek or revise facts, principles, theories, applications, etc. Investigation is the act or process of examining a crime, problem, statement, etc. carefully, especially to discover the truth. So, the word research is more relevant in that context. Besides the problems with synonyms, we can see the wrong translation of the word крен: roll and crane are inappropriate. It should be translated as lurch. One more error can be seen here: the word споруди is translated incorrectly by Google Translate.

It is worth noting that the context plays an essential role in translating technical texts since it helps differentiate synonyms. In paper titles, the context is incomplete, this fact causing some challenges.

Thus, machine translation errors can be both terminological and linguistic. Terminological errors are associated with Eurocodes, while linguistic ones are due to multicomponent terms and synonymy.

3.2. Specific features of translating construction terms

In 2014, the mechanism for simultaneous implementation of construction codes based on national technological traditions and building norms was harmonized with regulatory documents of the European Union (Eurocodes) and came into force in Ukraine. Eurocodes are European standards for designing structures of buildings developed by the European Committee for Standardization on the basis of the agreement with the European Community Commission. It is because of the complex and legally responsible nature of the construction industry that Eurocodes are to be considered by translators when working with construction terms.

Yet, not only students, but also many ESP teachers are not familiar with the Eurocodes dictionary designated to be guided by when translating construction terms.

Our research analyzes the terminological dictionary of Eurocodes by Hordeiev, Mushchanov and Perelmuter [5] published in 2013. The dictionary is aimed at unifying English-Ukrainian-Russian translation of Eurocodes. The fact that the Eurocodes dictionary has two editions confirms complexity of harmonizing construction terms.

The Ukrainian translation in the Eurocodes dictionary of the first edition differs from that in the second edition [6]. For example, membrane theory analysis in the first edition is translated as розрахунок за безмоментною теорією, while in the second edition it is translated as розрахунок за мембранною теорією. One more example: critical temperature of structural steel element is translated as критична температура конструктивного сталевого елементу, while in the second edition it is translated as критична температура конструкційного сталевого елемента. In our opinion, the option конструкційний is more relevant and accurate as it means suitable for constructing buildings or their parts and the word конструктивний means
associated with a structure and design.

The problem of providing Ukrainian equivalents of construction terms also occurs when translating the term integral abutment: the first edition translates it as об’єднаний стоян, while the second edition – об’єднана берегова опора.

After analyzing the given examples, we can conclude that the Ukrainian translation of some terms in the second edition is more relevant.

The mentioned discrepancies in translation cause a problem of term unification and standardization that results from the fact that translators working at a single document can suggest different options for translating the same terms. Because of this very discrepancy between dictionaries, machine translation without post-editing becomes inefficient when teaching ESP to construction students.

3.3. Analysis of data obtained

At the final stage of our research, machine translation errors are calculated and compared with those in human translation. In machine translation, term-related errors make 59% and linguistic/language errors – 41%. The number of errors made by Google Translate and Prompt programmes is nearly equal. When translating manually, these figures are lower: 45% and 30% respectively. Thus, the number of terminological errors decreases by 14%, while language errors – by 11% (figure 1).

![Figure 1: Comparison of machine and human translation errors.](https://example.com/figure1)

Our hypothesis has been proven: a decrease in machine translation quality causes an increase in human translation quality. The results of the study have been shown to the students. The positive outcomes have increased their interest and motivation to study ESP. Post-editing training is based on analysis of terminological and linguistic errors, which develops students’ analytical skills by forcing them to evaluate the difference in meanings of terms, find discrepancies with Eurocodes and select a more appropriate term.

Thus, reduction of terminological errors in human translation is the most significant achievement because incorrect translation of construction terms can lead to a complete distortion of meaning and technological errors in production.
4. Conclusions

The data obtained prove the advantage of human translation over machine one and substantiate the need for post-editing of machine translation.

When comparing quality of translation by Google Translate and Prompt programmes, it appears to be equally low. Since the titles of master qualification works should provide maximum information and reflect the essence of research, they are noted for long structures and multicomponent terms. As can be seen from the analysis, almost all 50 titles translated by machine translators contain terminological and linguistic errors.

Thus, the research has proven that a decrease in machine translation quality results in an increase in human translation quality. When translating, both programmes ignore requirements of the Eurocodes, which is quite natural for them. They also do not take into account the rules for translating multicomponent terms and offer options with numerous repetitions of the preposition of, which is unacceptable for English grammar. In addition, Prompt and Google Translate programmes may provide synonyms for one and the same word. We confirm that poor machine translation quality makes its post-editing obligatory.

The data analysis enables the following recommendations for teaching English for Specific Purposes for construction students:

1) making students acquainted with such linguistic phenomena as homonymy and synonymy of terms, “translator’s false friends” in order to prevent possible errors when using machine translation;
2) providing examples of possible translation options for multicomponent terms, taking into account the ways of avoiding excessive of-phrases;
3) introducing all kinds of specialized translation dictionaries of construction terminology – both printed (the Eurocodes dictionary) and online ones (e.g., Multitran) to students;
4) selecting appropriate samples of inconsistencies of terms in translation dictionaries and their equivalents in the Eurocodes dictionary;
5) reminding students about mandatory post-editing of machine translation;
6) introducing a special course Application of CAT tools and Machine Translation to Translating Technical Texts at non-linguistic universities of Ukraine.

The results of the study enable developing students’ CAT and general digital competences. We have proved that CAT translation competence includes all the components of language acquisition, along with some specific sub-competences: content-knowledge competence, research competence, and monitoring competence. Besides, the results help to enhance students’ interdisciplinary skills of combining knowledge, analyzing and thinking critically. The latter is essential, as adequate translation is impossible without critical thinking skills.

While conducting the comparative analysis, the students apply unconsciously the top levels of Bloom’s taxonomy to achieve effective learning outcomes: analyzing, evaluating, and creating. Post-editing translation learning, the results of which are presented in the paper, generates motivation (it makes students prove themselves that they are better than artificial intelligence), knowledge and creativity. Essential personal qualities are generated in the process of learning a foreign language when the emphasis is shifted to motivational, cognitive and creative, aspects
of future specialists’ personality. Search for terminological errors in machine translation contributes to developing construction students’ professional competences, namely the ability to analyze information collected in documents and draw appropriate conclusions.

Further research prospects involve analysis of CAT tools applied to English construction texts in order to select them for in/off-class activities when teaching ESP. Percentage of multi-component units, lexical density, complexity of materials, etc. are the criteria for their selection.

References


