ABSTRACT: The purpose of the research is to conduct an experiment to study the impact of translation memory systems (in terms of the use of terminology bases) on the effectiveness of future translators. The article substantiates the relevance of the use of translation memory systems, presents various areas of research on translation automation, considers the possibilities and prospects for the development of CAT systems. Article presents the results of an experimental study that would show how much novice translators rely on translation memory systems and whether they have the vocabulary at the proper level, which is typical for socio-political texts. The obtained results confirmed the hypothesis that when translating the original text by means of a CAT tool using a terminological base with erroneously translated lexical units, the quality of the translation of novice translators will decrease, and it will increase without using a terminological base with erroneously translated lexical units.


RESUMO: O objetivo da pesquisa é realizar um experimento para estudar o impacto dos sistemas de memória de tradução (em termos de uso de bases terminológicas) na eficácia de futuros tradutores. O artigo fundamenta a relevância do uso de sistemas de memória de tradução, apresenta várias áreas de pesquisa sobre automação de tradução, considera as possibilidades e perspectivas para o desenvolvimento de sistemas CAT. O artigo apresenta os resultados de um estudo experimental que mostraria o quanto tradutores iniciantes dependem de sistemas de memória de tradução e se eles possuem o vocabulário no nível adequado, o que é típico de textos sociopolíticos. Os resultados obtidos confirmaram a hipótese de que ao
El propósitos de la investigación es realizar un experimento para estudiar el impacto de los sistemas de memoria de traducción en la efectividad de los futuros traductores. El artículo fundamenta la relevancia del uso de los sistemas de memoria de traducción, presenta varias áreas de investigación sobre la automatización de la traducción. El artículo presenta los resultados de un estudio experimental que mostraría cuánto confían los traductores novatos en los sistemas de memoria de traducción y si tienen el vocabulario al nivel adecuado, que es típico de los textos sociopolíticos. Los resultados obtenidos confirmaron la hipótesis de que al traducir el texto original por medio de una herramienta TAO utilizando una base terminológica con unidades léxicas traducidas erróneamente, la calidad de la traducción de los traductores novatos disminuirá, y aumentará sin utilizar una base terminológica con unidades léxicas mal traducidas.


Introduction

The relevance of the study is due to the increasing importance of ICT in the modern market of translation services. Traditional ideas about translation as a professional activity are changing, and the professional model of a translator's professional activity becomes more complex with them. Scientific and technological progress certainly affects translation activities. More and more translation services are provided and offered via the Internet. Customers and translators may not meet in person to coordinate an order. This way of translating is convenient for freelancers since they can offer services and look for new clients (DOHERTY, 2016).

The latest technologies often determine the types of texts that need to be translated. In addition to the demand for translation of texts in traditional industries, for example, economic or technical ones, new areas are constantly emerging in which the volume of translations is increasing (BUNDGAARD; CHRISTENSEN; SCHJOLDAGER, 2016). A modern translator ceases to be a universal translator, masters a certain industry or several industries in which he/she specializes. He/she must constantly replenish his/her knowledge in the chosen field –
to study not only new terminology but also modern software used in the translation process, which is constantly updated and improved (PYM, 2013).

The translator must be competent in many aspects (POPOVA; BOBROVA; MALTZAGOVA, 2021), and the ability to work with specialized software is extremely important when carrying out translation activities (KUROCHKINA; MAIZIJEVA; MARKOVA, 2021). A special place among ICTs in the professional activity of a translator is occupied by CAT tools, which greatly simplify and accelerate the translation process (KARPIŃSKA, 2017).

**Literature review**

We proceed from the fact that there are (KRÜGER, 2016; ZHANG; YANG, 2012) two areas of research on translation automation: 1) applied (industrial implementation of text translation automation); and 2) theoretical (modeling of human speech activity as one of the methods of its research; development of mathematical formalisms for linguistic descriptions; search for algorithms for processing language objects; research of the relationship between human thinking and machines, etc.).

Referring to the peculiarities of using CAT programs in the work of a translator, researchers (PYM, 2011) considered the advantages that a translator receives by seeking help from such tools. First of all, the skills and abilities to use CAT programs have already become a mandatory requirement for the participation of a translator in large translation projects (ABU DAYYEH, 2020).

Despite the successes in the development of translation programs (OLOHAN, 2011), fully automated translation of texts with extensive topics is still an impossible task, and machine translation of texts with a narrow focus and limited vocabulary is performed with mandatory human participation (WANG, 2020). The paper (BOWKER, 2015) also notes that today the results of automated translation are still far from perfect, and professional correction of errors in the translated text is absolutely necessary.

The translator spends a lot of time working on the translation independently and performing it manually (SHCHEBUNYAEVA et al., 2021). Meanwhile, work in the same field of translation (for example, translation of medical texts or legal documents) implies a certain generality of terms and stable expressions that are often found in texts (KOPYTINA et al., 2021). Working manually, the translator must spend time retranslating those elements of
the text that CAT systems are able to optimize, significantly reducing the time required for translation (STARLANDER, 2013).

It was determined (BAAR, 2012; LIU, 2017) that CAT systems will acquire even greater development in the future, will receive more, and can be used with even greater efficiency.

The study (AL-RUMAIH, 2021) notes that future translators should be able to use translation memory systems, for which special courses are introduced in vocational education programs that introduce students to the necessary tools.

According to researchers (MARCHENKOVA; KOLOBKOVA, 2020), it is advisable to offer a course on mastering modern translation technologies to senior students, since they must already have developed translation skills and abilities to work with the CAT system. It is advisable to improve these skills and abilities and the formation of information and computer competence through the implementation of exercises, the main task of which should be to translate authentic texts using CAT programs (GRIDASOVA et al., 2021).

Outdated formats of working with printed text are being replaced by a more modern version of the translation in electronic format (YAO, 2017). CAT systems do not act as an end in themselves of training and professional activity but as a means of implementing these types of activities at a high-quality level (SIN-WAI, 2010). Translation technologies in the course of mastering them by students should be positioned as one of many, albeit very important, components of professional competence.

However, the ability to use CAT tools is not enough to perform a high-quality translation. The translator must know the vocabulary and terms specific to a particular industry, which will significantly improve the quality of the final result. Editing a text translated using a CAT program will become much more complicated in the absence of the necessary knowledge since the translator will be somewhat disoriented (MAFULAH et al., 2018).

In this regard, we consider it necessary to conduct an experiment that would show how much novice translators rely on translation memory systems and whether they have the vocabulary that is typical of socio-political texts at the proper level. According to scientific research (JUAN; YAHAYA, 2019), students in translation mainly focus on the surface structure of the original text, while professional translators, first of all, analyze the meaning of the original message and try to recreate the essence of the original message.
That is why the purpose of the study is to conduct an experiment to study the impact of translation memory systems (in terms of the use of terminology bases) on the effectiveness of future translators.

**Research objectives**

- To select the original text and create a terminology base containing errors;
- To organize the translation of the original text and divide the participants of the experiment so that one group (experimental) uses the proposed terminology base, and the other (control) does not;
- To determine the principles by which the translation texts will be analyzed, and analyze the translation texts based on these principles;
- To carry out quantitative processing of the obtained research results and to analyze and interpret the results obtained;
- To formulate conclusions and prospects for further research based on the results obtained.

Research hypothesis: the quality of the translation of novice translators will decrease when translating the original text by means of a CAT tool using a terminological base with erroneously translated lexical units, and it will increase without using a terminological base with erroneously translated lexical units.

The article consists of an introduction, a literature review, methods, results, discussion, and conclusion.

**Methods**

An experimental study on the influence of translation memory systems (in terms of the use of terminology bases) on the quality of translations by novice translators was conducted at the RUDN University in 2021. Within the framework of three stages: preparatory, during which a hypothesis was formulated and an experiment plan was developed; the main one, which consisted in the practical implementation of the experiment; and the final one, which provided for the analysis and interpretation of experimental data.

To test the hypotheses, we selected 26 1st-year students of the Master's degree program of the Translation Department of the Faculty of Foreign Languages, divided into two
experimental groups (EG-1 – 12 people and EG-2 – 14 people), including six boys and 20 girls aged 21-22 years.

The experimental study was conducted while teaching the basics of working with CAT tools in the framework of the "Translation Practice" discipline. Students get acquainted with the basics of working in translation memory systems: they translate texts from English into Russian and vice versa, using a machine translation module, translation memory bases, and terminology bases.

The research material was 26 translation texts of articles on socio-political topics, translated using CAT tools by novice translators – student-members of experimental groups.

An excerpt of an article in English on socio-political topics from "The Independent" British online magazine was borrowed as the original text. The choice of the primary source is due to the fact that students get acquainted with the peculiarities of the translation of socio-political texts within the framework of translation practice classes during their studies at the university. They translate political and economic news, expand their knowledge of the political structure of different countries and acquire background knowledge necessary for translation. The volume of the text is 21 sentences, 473 words, 3,288 printed characters (with spaces). The time to complete the translation is 80 minutes.

Students used the Memsource cloud-based translation memory system during the translation, as well as a terminology database compiled specifically for the experiment, which contained errors in the translation of 28 lexical units.

All students have fully translated the proposed text into Russian. The students of the first experimental group (EG-1) used the terminology base with errors during the translation, while the students of the second experimental group (EG-2) were self-reliant.

When checking the translations made by students, we relied on an assessment system that allowed us to take into account the features of our research:

1) Errors due to which the content of the original text is significantly distorted and/or the wrong translation option is selected, which is offered in the terminology database (1 penalty point is awarded);

2) Errors due to which the content of the original text is partially lost and/or the translation proposed in the database is partially selected (0.5 penalty points are awarded);

3) Errors that do not significantly or almost do not affect the content of the original text (0.1 penalty points are awarded).

The reliability of the differences in the results of students from different experimental groups was assessed by Student's t-test for unrelated samples.
Results

The results of the experimental section of the EG-1 are shown in Table 1.

Table 1 – The results of the experimental cut performed by the EG-1 using the terminology base with errors (in penalty points)

<table>
<thead>
<tr>
<th>Student</th>
<th>Type 1 errors</th>
<th>Type 2 errors</th>
<th>Type 3 errors</th>
<th>Total number of penalty points</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8.0</td>
<td>1.5</td>
<td>0.1</td>
<td>9.6</td>
</tr>
<tr>
<td>2</td>
<td>9.0</td>
<td>0.5</td>
<td>0.2</td>
<td>9.7</td>
</tr>
<tr>
<td>3</td>
<td>12.0</td>
<td>1.5</td>
<td>0.3</td>
<td>15.8</td>
</tr>
<tr>
<td>4</td>
<td>6.0</td>
<td>0.0</td>
<td>0.2</td>
<td>6.2</td>
</tr>
<tr>
<td>5</td>
<td>5.0</td>
<td>1.0</td>
<td>0.2</td>
<td>6.2</td>
</tr>
<tr>
<td>6</td>
<td>6.0</td>
<td>0.5</td>
<td>0.2</td>
<td>6.7</td>
</tr>
<tr>
<td>7</td>
<td>4.0</td>
<td>2.0</td>
<td>0.2</td>
<td>6.2</td>
</tr>
<tr>
<td>8</td>
<td>7.0</td>
<td>1.5</td>
<td>0.1</td>
<td>8.6</td>
</tr>
<tr>
<td>9</td>
<td>9.0</td>
<td>1.0</td>
<td>0.3</td>
<td>10.3</td>
</tr>
<tr>
<td>10</td>
<td>8.0</td>
<td>1.5</td>
<td>0.1</td>
<td>9.6</td>
</tr>
<tr>
<td>11</td>
<td>5.0</td>
<td>0.5</td>
<td>0.2</td>
<td>5.7</td>
</tr>
<tr>
<td>12</td>
<td>7.0</td>
<td>0.5</td>
<td>0.1</td>
<td>7.6</td>
</tr>
<tr>
<td>Average value</td>
<td>7.17</td>
<td>1.0</td>
<td>0.17</td>
<td>8.52</td>
</tr>
</tbody>
</table>

Source: Prepared by the authors

As follows from Table 1, the average value of the penalty point is 8.52 in EG-1, while the average value of errors of the first type, which include a significant distortion of the content or the choice of a translation of a certain word or phrase proposed by the terminology base, is 7.17. The error rate of the first type varies from 4.0 to 12.0 penalty points. The high score for errors of the first type in EG-1 is explained by the fact that students were mainly guided by the base rather than by their knowledge and context.

The average penalty score for errors of the second type is 1.0, which can be explained by the fact that the lexical units, the content of which was slightly changed, were generally familiar to students. Most of the mistakes of the second type were made by student 7 (the penalty score is 2.0), which can be explained by the fact that some lexical units were unknown to him, unlike other subjects.

The average penalty score for errors of the third type, which includes those that do not affect the content of the message is 0.17 in EG-1 (with the highest indicator being 0.3 penalty points).
Table 2 presents the results of an experimental cross-section of students of the EG-2, which did not use the terminology base with errors.

**Table 2** – The results of an experimental cut performed by the EG-2 without the use of a terminology base with errors (in penalty points)

<table>
<thead>
<tr>
<th>Student</th>
<th>Type 1 errors</th>
<th>Type 2 errors</th>
<th>Type 3 errors</th>
<th>Total number of penalty points</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.0</td>
<td>2.0</td>
<td>0.0</td>
<td>5.0</td>
</tr>
<tr>
<td>2</td>
<td>4.0</td>
<td>1.0</td>
<td>0.1</td>
<td>5.1</td>
</tr>
<tr>
<td>3</td>
<td>5.0</td>
<td>2.0</td>
<td>0.1</td>
<td>7.1</td>
</tr>
<tr>
<td>4</td>
<td>4.0</td>
<td>2.0</td>
<td>0.2</td>
<td>6.2</td>
</tr>
<tr>
<td>5</td>
<td>2.0</td>
<td>1.5</td>
<td>0.1</td>
<td>3.6</td>
</tr>
<tr>
<td>6</td>
<td>4.0</td>
<td>0.5</td>
<td>0.3</td>
<td>4.8</td>
</tr>
<tr>
<td>7</td>
<td>7.0</td>
<td>0.5</td>
<td>0.1</td>
<td>7.6</td>
</tr>
<tr>
<td>8</td>
<td>2.0</td>
<td>2.0</td>
<td>0.1</td>
<td>4.1</td>
</tr>
<tr>
<td>9</td>
<td>4.0</td>
<td>0.5</td>
<td>0.3</td>
<td>4.8</td>
</tr>
<tr>
<td>10</td>
<td>7.0</td>
<td>1.5</td>
<td>0.1</td>
<td>8.6</td>
</tr>
<tr>
<td>11</td>
<td>10.0</td>
<td>1.0</td>
<td>0.1</td>
<td>11.1</td>
</tr>
<tr>
<td>12</td>
<td>6.0</td>
<td>2.5</td>
<td>0.2</td>
<td>8.7</td>
</tr>
<tr>
<td>13</td>
<td>5.0</td>
<td>1.5</td>
<td>0.1</td>
<td>6.6</td>
</tr>
<tr>
<td>14</td>
<td>2.0</td>
<td>2.5</td>
<td>0.2</td>
<td>4.7</td>
</tr>
<tr>
<td>Average value</td>
<td>4.64</td>
<td>1.8</td>
<td>0.2</td>
<td>6.29</td>
</tr>
</tbody>
</table>

Source: Prepared by the authors

According to the data presented in Table 2, the total penalty score is 6.29, which is less than the same indicator in the EG-1. The overall indicators for the group range from 3.6 to 11.1 penalty points.

Students in EG-2 also made mistakes of the first type, but there were fewer of them than errors of the second and third types. The average error rate of the first type is 4.64, which is significantly less than the same indicator in EG-1. One of the students has the highest penalty score – 10, which most likely indicates that the lexical units were unknown to him/her, and therefore caused certain difficulties in translation. Three students made only two mistakes of the first type, which is the best result in the group. Interestingly, the index of errors of the second type in EG-2 is almost twice as high as the results in EG-1: it is 1.8, while the maximum number of penalty points for errors of this type is 2.5.
The results of statistical analysis of the reliability of differences using Student's t-test for the EG-1 and EG-2 confirmed higher results in the EG-2 when making errors of the first type ($t_{\text{Emp}} = 2.9; \ p \leq 0.01$)

After a detailed analysis of the translations of the two experimental groups, we conclude that EG-1, which had the opportunity to use the terminology base with errors, made a greater number of errors of the first type, while errors of the second and third types were not significant in a number concerning errors of the first type. EG-2, which did not have a terminological base, made more mistakes of the second type, which can be explained by the choice of stylistic synonyms that were unsuccessful for the proposed context.

**Discussion**

The results of EG-1 (use of the terminology base with errors), presented in Table 1 led us to conclude that it is more likely that students who used a misspelled termbase when translating relied on the option suggested in that termbase instead of critically analyzing the resources provided, thoughtfully analyzing the meaning of the original message, and trying to reproduce its essence.

In turn, after analyzing the results of the EG-2 (did not use the terminology base with errors) presented in Table 2, we were convinced that this group coped better with the translation. Students relied on their knowledge, were critical of the translation text, analyzed its deep meaning, carefully checked and edited the translation text, and therefore received a better result.

This is consistent with the research results (MARCHENKOVA; KOLOBKOVA, 2020; POPOVA; BOBROVA; MALTZAGOV, 2021), according to which students, when starting to translate texts of various subjects in CAT programs, often overly trust the machine and do not pay attention to some mistakes made during translation.

To exclude such situations, the researchers propose to improve translation skills and abilities and the formation of information and computer competence when using machine translation systems by performing a system of exercises, the main task of which should be to translate authentic texts using automated translation programs.

Machine translation systems do not act as an end in themselves of training and professional activity but as a means of implementing these types of activities at a high-quality level (AFANASIEV et al., 2021). According to researchers (DOLZHENKOV et al., 2021; IVANOVA; LONTSKAYA; PAKHOMOVA, 2021), and we fully agree with this, translation
technologies in the course of mastering them by students should be positioned as one of many, although very important, components of professional competence.

We believe that based on several studies, the development of a system of exercises for teaching translation students should be based on the psychological structure of translation, which, being a type of activity, is formed by several operations (abilities) and actions (skills) (BAAR, 2012). Such a structure of translation activity, according to researchers (MAFULAH et al., 2018; STARLANDER, 2013), makes it necessary to include the formation of such competence as the possession of translation automation tools in the content of training, which includes:

- The ability to use the tool efficiently and quickly, integrate several programs to help with translation, editing, terminology processing, layout, search in documents (for example, text processing tools, spelling and grammar checks, translation databases, terminology databases, voice recognition programs);
- The ability to create and manage databases;
- The ability to adapt and master new tools, especially when translating multimedia and audiovisual material;
- The ability to prepare and carry out the translation in various formats and for various technical environments;
- Knowledge of the possibilities and limitations of the machine and automated translation.

**Conclusion**

The purpose of using CAT systems is to optimize the translator's work, reduce the time required for translation, coordinate the use of a terminological glossary during a long project and avoid re-translating similar material when working with technical and official-business style texts (namely, such texts make up a significant layer of the total work of translators).

However, the experiment proved that when using the terminology base, novice translators tend to trust it too much and actively rely on it without conducting a critical analysis of the resources offered to them. This, in turn, indicates that translation memory systems can hurt the professional activities of novice translators, and therefore a special organization of training in modern translation technologies is important, which would take
into account the results obtained and aimed at developing students' critical attitude to the mentioned software and the resources provided.

However, it should be taken into account that due to the small number of subjects in this experiment, the conclusions cannot be final, and therefore they should be tested on a sample of 150 students in the future.

ACKNOWLEDGMENTS: This paper has been supported by the RUDN University Strategic Academic Leadership Program.

REFERENCES


How to reference this article


Submitted: 28/12/2021
Required revisions: 19/01/2022
Approved: 25/02/2022
Published: 30/03/2022