Ministry of Education and Science of the Republic of Kazakhstan

M. Narikbayev KAZGUU University

"Approved for Defense"

Supervisor _____

_____

«__» _____ 20__

MASTER'S THESIS (PROJECT)

"Artificial Intelligence and Translation Technology"

specialty 6M020700 - Translation Studies

Written by Lyazzat Altynbekova

Supervisor Meiram Islam

Nur-Sultan, 2020

M. Narikbayev KAZGUU University

ARTIFICIAL INTELLIGENCE AND TRANSLATION TECHNOLOGY

Lyazzat Altynbekova

May 22, 2020

"Approved"

Supervisor's First and Last Name **Meiram Islam**

Supervisor's Signature _____

"__" _____ 20__

Nur-Sultan, 2020

Abstract of "Artificial Intelligence and Translation Technology"

By Lyazzat Altynbekova

HSE KAZGUU University

May, 2020

Many unprecedented changes in translation that can be seen today resulted from technological advances. In this thesis, the way in which translation technologies based on application of artificial intelligence (AI) make an impact on translation practice and translators is discussed. It is argued that translation technologies have increased significantly productivity and quality in some fields of translation. However, these technologies also lead to considerable uncertainties and challenges for translators.

The thesis discusses in detail two machine translation systems – "PROMT" and "Google Translate", which are currently the most popular among Internet users. The work of these systems is organized according to different principles: "PROMT" works according to the so-called principle of "translation according to the rules" and copies the human translation activity; Google Translate translates using technology based on the principle of statistics. During testing, typical mistakes made in machine translation have been identified, and the possibilities and areas of use of each of the translation systems have been determined. It has been established that these systems are unsuitable for working with texts containing a large number of complex sentences. These programs work mainly at the level of collocation, and they can be successfully used to translate formalized texts, such as technical documentation, consumer instructions, formal descriptions, etc., which are characterized by the use of simple common sentences and which do not contain sentences with complex syntactic constructions. It is concluded that professional translators can resort to the help of these systems in their practical activities when working with certain types of texts, which helps save time and optimize the translation process. However, a prerequisite for achieving high quality translations made using computer programs is their post-editing.

Key words: *artificial intelligence, translation technology, machine translation, translator, information technology, electronic dictionaries, translation automation, translation quality.*

TABLE OF CONTENTS

1 Introduction

The progress of translation technologies based on application of artificial intelligence (AI) (Bughin & Manyika, 2018) in different aspects have been affected by industrialization, globalization, social media and technology (Language Industry Survey, 2018). Translation technologies developed in recent years become an integral part of communication practice. This article is aimed to demonstrate how technological developments in translation being applications of AI (Foundations of Artificial Intelligence, 2018) have fundamentally changed the way translators work. The aim includes reviewing the process of translation. The product of translation, and the changes in the role of the translator. It is obvious that the development of translation technologies led to increased productivity of translators. However, some significant problems connected with the value of translation and the status of the profession of translator are also emerged.

The aim of the thesis is to investigate the current situation in the application of Artificial Intelligence and Machine Technologies to the work of translator and their impact on the quality of translation.

To reach the aim, the following tasks are to be implemented:

- Peculiarities of computer-aided translation are to be elucidated;
- Peculiarities of machine translation are to be investigated;
- Advantages and disadvantages of translation technologies are to be explained;
- The impact of translation technologies on translators is to be explored.
- The connection between translation technologies and quality of translation is to be distinguished;
- The need for post-editing after machine translation is to be explained.

As methodological tools, a comparative analysis of machine and human translation was used. In the practical part, an experiment was made. It consisted of translation into Russian by the PROMT and Google Translate systems, and the obtained results were analyzed.

2 Importance of artificial intelligence (AI) and translation technologies for translation practice

There are two main incentives for the development of machine translation in the modern world. The first is actually scientific; it is determined by the complexity of computer simulation of translation. As a type of linguistic activity, translation affects all levels of the language - from recognition of graphemes (and phonemes in the translation of oral speech) to the transmission of the meaning of a statement and a text. In addition, the translation is characterized by feedback and the ability to verify immediately the theoretical hypothesis about the structure of certain language levels and the effectiveness of the proposed algorithms. This characteristic feature of translation in general, and machine translation in particular, attracts the attention of theoreticians, because of whichever new theories of automation of translation and formalization of language data and processes continue to emerge.

The second incentive is social, and it is due to the increasing role of the practice of machine translation in the modern world as a necessary condition for ensuring cross-language communication, the volume of which is increasing every year. Other ways to overcome the language barriers to communication - the development or adoption of a single language, as well as the study of foreign languages - cannot be compared with translation in terms of effectiveness. From this point of view, it can be argued that there is no alternative to translation, so the development of high-quality and high-performance machine translation systems helps to solve the most important social and communicative problems.

The important benefits of machine translation can be presented as follows:

High speed of translation. Using a machine translation system can significantly reduce the time required to translate texts.

Low cost of translation. Using the services of professional translators, we are forced to pay money for each page of the translation. However, often there is no need to get a perfect translation of the text, but you need to grasp quickly the meaning of the sent letter or the content of the page on the Internet. In this case, the translation system will undoubtedly become a reliable and effective assistant.

Confidentiality. Many users regularly use the MT system to translate personal letters, because not everyone is ready to give personal correspondence to an outside translator or entrust the translation of financial documents.

Universality. A professional translator, as a rule, has specialization in translating texts of a certain subject. The translation program will cope with the translation of texts from a wide variety of fields: for the correct translation of specialized ones, connect the necessary settings. Enough terms online translation and translation of the content of web pages. The advantages of the online translation service are obvious. Online translation services are always at hand and will help at the right time to translate information quickly if you do not have a translator program. In addition, today with the help of translation systems it is possible to translate the content of Internet pages and search engine queries.

3 Peculiarities of computer-aided translation (CAT) and machine translation (MT)

3.1 Peculiarities of computer-aided translation

Analysis of the work of highly qualified translators allowed us to identify the main stages of their work:

1) Translation is based on the idea of passive grammar (analysis) and active (synthesis);

2) The analysis of the syntactic and semantic structures of the utterance occurs simultaneously. In translation, this is an analysis of parts of speech and sentence members;

3) The translator does not previously restore the complete syntactic structure of the sentence;

4) Analysis and synthesis occur according to some separate meaningful pieces ("translation units", syntagmas, etc.);

5) The translator works by trial and error;

6) During analysis and synthesis, there is a constant check for meaningfulness during which the analysed syntagma is compared with previous and subsequent ones. When translating, recursions constantly occur (Aranberri et al., 2016).

Based on these data, automatic analysis and synthesis algorithms have been developed for different MT systems. In systems based on translation correspondences, the sequence of stages of analysis and synthesis is approximately the following:

1) Dictionary search, identification of lexical units;

2) The exclusion from further analysis of closely related phraseological units, the grammatical analysis of which is not necessary due to the presence of strictly fixed translation equivalents that are in a special dictionary of turns or translation memory of the system.

3) A grammatical analysis of the most significant phenomena of the input text. Definition of the basic syntactic functions of sentence words. Definition of the basic grammatical parameters (cases, time, type, mood, gender, etc.) for the main lexicographic classes.

4) Translation of unambiguous and ambiguous words. Single words are translated according to a unique correspondence table. Translation of polysemous words can be done in various ways, including using context analysis;

5) Grammar analysis and synthesis;

6) Synthesis of the output text (Aranberri et al., 2016).

The basis for the classification of vocabulary and the creation of a machine dictionary is a special classification of words. In the initial classification of words, seven main lexical and grammatical classes are distinguished (nouns, adjectives, verbs, adverbs, prepositions, conjunctions, punctuation marks) and the class of homographs. Pronouns and numerals in MT refer either to nouns or to adjectives, depending on their syntactic functions.

The resolution of lexical ambiguity is one of the most pressing problems of computer linguistics. In the vast majority of cases, the lexical polysemy of a word is allowed within the phrase. In MT systems, lexical ambiguity is resolved using a contextual dictionary. The source material for its compilation is a concordance dictionary, which is a list of textual uses of each word taken in the context of a certain size. When using parallel texts, you can compare each use of the word form in the input text with its translation in the output text. The translation algorithm for a multi-valued word is a sequence of context requests for the presence of determinants in it, i.e. words that uniquely identify one particular translation of a multi-valued word. Context analysis operations are performed by standard operators, each of which represents a certain subroutine for computers. All operators are numbered in order. The operator number is an indication of which word should be checked and which operator should be returned when the answer is no.

Thus, based on comparative data, the following advantages and disadvantages of machine translation systems can be distinguished below.

3.2 Peculiarities of machine translation

At the heart of modern machine translation systems is a translation algorithm that uses formal grammar of languages and statistical data. To learn a language, the system compares thousands of parallel

texts containing the same information, but in different languages. For each text studied, the system builds a list of unique features. For example, rarely used words and special characters that occur in the text with a certain frequency. In machine translation systems, as a rule, there are three main parts: a translation model, a language model, and a decoder. A translation model is a table in which, for all words and phrases in one language, possible translations into another language are listed with the probability of these translations. The system compares not only single words, but also phrases from several words in a row. Translation models for each pair of languages contain millions of pairs of words and phrases. As for the language model, it is created by the system at the stage of studying texts.

The decoder is engaged in the translation. It carries out morphological and syntactic analysis of the text and for each sentence selects all translation options with sorting in descending order of probability. Then, the decoder evaluates all the obtained options using the language model for the frequency of use and selects the sentence with the best combination of probability and frequency.

Currently, types of machine translation are distinguished by the degree of automation:

- Fully automatic;
- Automated machine translation with human participation (with pre-, inter- or post-editing);
- Human-made translation using a computer (for example, using electronic dictionaries).

The first MT systems are characterized by a direct (word-by-word) translation strategy. The essence of this approach to the construction of MT is that the source text in the input language is gradually converted through a series of stages into the text of the output language. The transformations are reduced to the fact that the word (phrase) in the input language is replaced by its vocabulary equivalent in the output language. It is clear that in first-generation systems using a direct translation strategy, there is no need to model the functioning of the language system as a whole. For the operation of such systems, the vocabulary correspondence rules are sufficient. In rare cases, a context analysis is performed to translate ambiguous expressions, again presented in the system dictionary. It is important to keep in mind that the direct translation strategy does not distinguish between understanding (analysis) and synthesis (generation), since they are practically excluded from transformations according to the rules of dictionary

correspondences. Direct translation is always tied to a specific pair of languages. For example, the ambiguity of the expressions of the input language is allowed only to the extent that it is necessary for the output language. According to the time frame, the first generation systems were mainly created in the period from the late 1940s to the mid-1960s.

A significant modification of the direct translation strategy is found in systems with a transfer - the stage of interlanguage operations, not reducible only to the replacement of tokens of the input language with the dictionary correspondences of the output language. The presence of the transfer stage involves the construction of an “intermediate” or “internal” representation, which is further “adapted” to the structure of the sentence of the output language. Unlike the first strategy, in the architecture of MT systems with transfer, analysis (understanding) and synthesis exist as special procedures and are served by various algorithms. The development of the idea of transfer led to the appearance of a translation based on in-depth linguistic analysis. This strategy involves the analysis of the input text at all language levels (morphological, syntactic, semantic, pragmatic), as well as a multi-level synthesis of the output text. Criticism of the direct translation strategy has led to the creation of an intermediary language strategy (interlingua). The main feature of this strategy is that between the structures of the input language and the structures of the output language there is one or more intermediate languages into which the expressions of the input language are sequentially “copied” according to the relevant rules. Analysis and synthesis using the language of the intermediary are fundamentally separated. The analysis is carried out in the categories of the input language, and the synthesis in the categories of the output. Languages of syntactic and semantic-syntactic structures, purely semantic languages, languages of deep semantics approaching conceptual representation in the categories of knowledge theory (frames, scenarios, plans) can serve as a language (s) of intermediaries. Recently, a translation memory strategy has been developed. Translation memory - a database containing a set of previously translated texts.

One record in such a database corresponds to a “translation unit”, for which one sentence is usually taken (less often - part complex sentence, or paragraph). If the next sentence of the source text exactly matches the sentence stored in the database, it can be automatically substituted into the translation. The

new proposal may also slightly differ from the one stored in the database. Such a proposal may also be substituted in the translation, but the translator will have to make the necessary changes.

Algorithm of machine translation based on linguistic analysis can be presented as follows:

Step 1. Obtaining sentences of the source text from a file or from a buffer in memory.

Step 2. Splitting the sentence into words and defining the boundaries of the sentence.

Step 3. Morphological analysis of the source text - obtaining all possible lexical codes for each word found in the dictionary.

Step 4. Parsing the source text - grouping homogeneous adjectives and nouns, building a tree of main/dependent words.

Step 5. Semantic analysis of the source text.

Step 6. Translation of the constructed tree.

Step 7. Implementation of the coordination of the translated tree - semantic, syntactic and morphological synthesis.

Step 8. Writing the translated sentence to a file or to the buffer.

Step 1 is obvious.

Step 2. The task of breaking the text into words and sentences. Despite the apparent simplicity, the task of breaking the text into words and sentences in the general case is far from trivial. The main difficulty is represented by various abbreviations, initials, direct speech, words written with a hyphen, etc.

Word recognition is carried out using special templates. These patterns describe various alpha, numeric and alphanumeric groups and punctuation characters that will then stand out as separate words. For example, as separate words, dates recorded in their digital expression, paragraph and subparagraph numbers, abbreviations together with periods (according to a special dictionary), as well as words written with a hyphen in case they are recognized by a special dictionary module - a complex analysis module words. This module is used to recognize phrases such as black and white, spicy and spicy. Analysis and translation of such words is carried out on the basis of special rules for morphological transformations of

adjectives. As a result of the analysis of the selected words, some words (initials, abbreviations, etc.) will be assigned special markers that will allow to resolve ambiguity when recognizing the boundaries of sentences. Also at this stage, words are normalized in order to prepare them for a dictionary search.

Step 3. Morphological analysis. The solution to this problem is based on the dictionary of the source language. As a result of a dictionary search, a lot of lexical and grammatical classes are assigned to each word of a sentence: part of speech, case, number, gender, category, etc., which allows further comparison of classes based on certain characteristics (for example, to check the consistency of adjectives and nouns). The process of searching for words in a dictionary involves, in addition to searching for the original word if it was not found in the dictionary, searching for words with the removal of possible prefixes. For effective search for prefixes, a tree structure is used, the elements of which are the letters of prepositions. The search stops when either there is no further transition in the tree, or when a preposition is found and a word without this preposition exists in the dictionary. In addition to the dictionary of prepositions, for each of the languages there is a table of interlanguage correspondence, with the help of which the resulting word is obtained at the stage of text synthesis. At the stage of class recognition, phrases are also highlighted, which, according to the dictionary, are translated in one word (idiom dictionary): on the drum, go to a standstill, kick the bucket (literally - kick a bucket, “play in a box”). It is further believed that all such phrases are represented in one word. This ensures the correct coordination and translation of the phrase as a whole.

Step 4. Parsing. First, for each word, a search is made for the main word with which it must be coordinated as a result of the translation. It is not supposed that ambiguity should already be completely removed. In the process of searching for main words, the basic removal of ambiguity is performed. The syntax tree is constructed by sequentially recognizing predefined linguistic patterns and applying certain operations based on them.

The main operations in pattern recognition are:

a) Checking whether a word is a specific part of speech with specific characteristics (for example, a noun in the genitive);

b) Checking whether a certain word is a homonym, i.e. whether it can belong to different parts of speech (for example, fries, love);

c) Checking the coordination of two words (full agreement - adjective-noun, approval by case - noun-noun, etc.): beautiful girl, Anna Ivanova;

d) Obtaining semantic characteristics of the control of prepositions and verbs - each verb and preposition requires the presence of corresponding cases in the words they define (they control a specific case).

The set of these cases depends on the meaning of these prepositions and verbs (for example, being only controls the prepositional case, and writing controls the dative, accusative, instrumental and prepositional cases).

If any pattern is recognized, the following operations are possible on the elements that it covers:

- removing from the list of lexical and grammatical classes words of all classes that do not satisfy certain conditions (for example, deleting all classes except for classes that have among their nominative case characteristics);
- fixing the dependence between the words (for example, the verb noun in the nominative case);
- removing a word from the set of separate words of a sentence and adding it to one of the main words indicating the type of dependence, i.e. the word becomes dependent and is removed from consideration in the next steps of building a tree to simplify the rules of analysis.

Step 5. Semantic analysis of the source text. The main objective of this stage is to resolve ambiguity based on the resulting dependency tree. For this, the ambiguity of the basic words is initially resolved. As studies have shown, it is advisable to pairwise match adjacent base words in the reverse order of the words in the sentence. After all lexical words are matched with one lexical and grammatical class; the words “dependent on them” are “matched”. The selection parameters of the lexical and grammatical classes of dependent words are selected according to the type of dependence and the lexical and grammatical class of the main word. An example of a rule for removing ambiguity: Processing the homonym “adjective-participle”. If the homonym adjective-participle is preceded by a comma, then the word in question

removes the word form of the adjective. Otherwise, the participle word forms are removed: a strange sound was heard, whistling like the wind; c - the sound is whistling.

Step 6. Translation of the constructed tree. The translation process consists of the following steps:

a) A word-by-word translation of the base words of the tree (except for verbs) of the dependencies is performed with the preservation of the original lexicographic class (or the class closest to it in terms of characteristics);

b) for verbs from the list of basic words that have a gender attribute as initial characteristics, translation is carried out into many verbs of the same paradigm with singular and gender signs; for other verb forms - translation is carried out while maintaining the original lexical and grammatical characteristics;

c) for dependent words, the translation result is a set of words, which is determined on the basis of the type of dependence and the lexical and grammatical classes of the main words (for example, the translation of adjectives is the whole paradigm, most pronouns are translated while preserving the original cases) - the final lexical characteristics are determined at the synthesis stage. Also at this step, the analysis of words is carried out, the translation of which is a phrase. As a result of the translation, the phrase must be consistent. To do this, when this situation is detected, the dependency tree is completed on the basis of the main word phrases.

Step 7. Implementation of approval of the translated tree. The translation results in a partially consistent dependency tree. To obtain complete agreement, it is sufficient to use a procedure similar to the procedure for the final resolution of ambiguity, used at the stage of tree construction. Since the translation was carried out on the basis of the dependency tree, this procedure will allow you to get a consistent presentation of the proposal in the resulting language.

Then, based on the tree, the resulting sentence is constructed. To do this, for each word in the dictionary of the resulting language, a search is performed to obtain a specific word form corresponding to a fixed lexical and grammatical class. Also, the resulting words are supplemented by translations of prefixes if they were removed from the original word in the analysis.

Step 8 is obvious.

The structure of machine translation systems can be presented as follows:

1. Supporting software.

2. The composition of the linguistic database.

3. Linguistic processor.

1. Supporting software. At the stage of processing the source text, certain auxiliary software tools must be present in the machine translation system. To convert texts from one code representation to another, converters are needed. A typical converter supports a subset of code tables used for characters of a particular language. Therefore, for the Russian and Belarusian coding CP866, CP1251, and Unicode are relevant. When processing documents, the problem arises of converting the original document into plain text, for its further semantic and syntactic processing. At the same time, at the conversion stage, it is necessary to save useful information about the structure of the document and its style, about the relationship between paragraphs, about headings, etc. The task of the preformer is to recognize various formats of documents received at the input and to extract text information from these documents while maintaining its structure. Since there are a large number of different document formats, it becomes necessary to write preformers for each of these formats. This task is very time-consuming and, in addition, with the advent of a new format documents, it becomes necessary to write a separate transformer for the new format. To avoid such problems, in practice, as a rule, preformers are used that convert various document formats to one, which is usually the most easily structured. Usually HTML, XML, etc. are selected as such a format. Next, an analysis module is created for one format selected as the base, the output of which produces a structured text with a number of features selected for the text analysis depth by the preformer. When new formats appear, it becomes necessary only to create a preformer that converts this document format into a basic one. A separate block can be identified programs for interacting with databases, which are necessary for a complete linguistic analysis of the source text and synthesis of the text in the translated language. For convenient work with the machine translation system, it is necessary to provide a user-friendly interface that operates with a job management language accessible to any user.

It is with this interface that the user should be able to enter, edit and verify the text without programming, translate it, print it, carry out statistical processing of the material, obtain the necessary information, etc.

2. The composition of the linguistic database. The linguistic database for a machine translation system includes accumulated linguistic data objectified by texts, card indexes, dictionaries, grammars, and other linguistic sources. The typical composition of a linguistic database can be limited by the following components:

a) The lexico-grammatical classifier of the properties of the source language and the translated language (morphological coding system). When analysing the source text, each word in it should receive the corresponding morphological characteristics: a sign of a part of speech, gender, case, mood, number, etc. The coding system should be uniform for a specific machine translation system.

b) Basic bilingual morphological dictionary. This dictionary establishes the word-by-word correspondence of each word form of the source language to the word forms of the target language.

c) Dictionary of abbreviations and abbreviations. The dictionary is used at the stage of breaking the source text into words and sentences. Abbreviations and abbreviations should be decrypted, as far as possible, since they can be members of a sentence, therefore, they must be taken into account in syntactic and semantic analysis.

d) Dictionary of idioms. This dictionary is used before parsing, because very often the idiom is one member of a sentence and is considered as a whole; when translating an idiom in the source language can correspond to one word in the translated language.

e) Terminological dictionaries for subject areas. Additional dictionaries are included when you need to translate specialized texts.

f) Syntax dictionary. This dictionary should contain information on the syntactic compatibility of the sentence members in both the original language and the translated language, as well as the syntactic correspondences necessary for the translation.

g) Semantic dictionary (thesaurus, ontology). This component contains information about the semantic compatibility of lexemes, about lexico-semantic fields, it is used at the stage of constructing a semantic sentence graph.

h) Corpus of parallel texts. The corpus contains texts in the original language and their translations into another language. When a sentence or its fragment is found in the case of parallel texts, it is inserted into the text of the translation its correspondence in the translated language. The translation memory technology is built on the use of the text corpus.

3. Linguistic processor. The linguistic processor is designed for complete linguistic analysis of text in the source language, as well as synthesis of text in the target language. The linguistic processor includes the following components:

a) A program for breaking text into sentences and words.

b) A program for recognizing stable phrases. Idioms must be analysed and translated as an indivisible whole.

c) A program for deciphering acronyms and abbreviations.

d) The program of morphological annotation of the source text.

e) A parsing and dependency tree building program.

f) The program of semantic analysis and construction of the semantic graph of each sentence of the source text.

g) The program for selecting translational correspondence from a bilingual dictionary or corpus of parallel texts.

h) The program of semantic synthesis of text in a translated language.

i) A program for constructing the syntactic structure of a sentence and determining the word order in a synthesized sentence.

j) The program of morphological synthesis of word forms in the translated text.

In 1997, R. Nieko and M. Forkada proposed the idea of using the “encoder-decoder” model in machine translation — the source text is encrypted into a universal “representation”, and then decrypted

in the desired language. Reminiscent of interlinguistic systems, right? In 2003, a group of researchers from the University of Montreal, under the leadership of J. Benggio, developed a language model based on neural networks that helped overcome the problem of data parsing, which was typical for statistical systems that were popular at that time. This served as the starting point for the development of neural machine translation (NMT), which now occupies the minds of developers and translators.

NMT systems are taught in large text corps, which makes them similar to NSTs, but their approach to text processing is completely different. In 2013, N. Kalkbrenner and F. Blans developed a model capable of converting the source text into a continuous vector using a convolutional neural network encoder, and then using the recurrent neural network decoder to translate this vector into text in the target language. A year later, C. Cho and colleagues suggested using recurrent neural networks as an encoder - in their opinion; NMT is better suited for word processing. It works as follows.

It quickly decided to add another element to the "encoder-decoder" structure - the so-called "attention mechanism". Due to it, the decoder takes into account not only the last view proposed by the encoder - due to additional neural connections and layers. In general, depending on the characteristics, existing models can work with text in different ways: there are machines that can take into account the units following the word being processed, which makes the translation more accurate.

The capabilities of neural networks force large players to switch to NMTs. Back in September 2016, Google began to use the NMT instead of a statistical phrase translation for a Chinese-English pair, and after a while added a few more languages. It is curious that Google Translate is still switching from NST to NMT for some language pairs, and since NMT "pays attention to the context", you can find out which model is currently being used. In language pairs for which the NST is still applied, the system highlights individual elements of the target sentence if you hover over it; in the case of languages for which a neural network is connected, sentences are fully highlighted. Because of this, by the way, it is more difficult to track the origin of errors.

Now the RNA-based codec-decoder architecture, built-in attention mechanism and long-term short-term memory (LSTM) are a mandatory minimum for the average online translator (although in June 2017

a new model appeared that uses only attention mechanisms). In addition, the crowdsourcing mechanism has been introduced in some systems. The same Google Translate offers the user to note the most successful translation option and remembers the most frequently selected ones.

In order to compare the NST and the NMT, in 2016 we conducted an experiment on the parallel UN corps - these are 15 language pairs and 30 directions of translation. The results were evaluated using the BLEU scale (evaluates the proximity of a machine translation to a human reference made by a professional translator), and the quality of the translation did not yield to the neural network or exceeded the NST in all 30 directions. Researchers concluded that a neural network makes fewer morphological and syntactic errors.

Despite all the advantages of NMTs, errors are still inevitable, and the process of training neural networks takes a lot of time. In addition, statistical models also have their own strengths. Therefore, recently researchers have been combining various approaches to translation automation: in an attempt to achieve a more natural text, hybrid machine translation systems are born. You do not need to go far for an example of such a machine: “Yandex Translator” is familiar to everyone.

In “Yandex Translator”, both models generate their translation options, and then a special algorithm evaluates / selects / combines them and gives the result. Experiments show that switching a model often depends on the length of the text and whether there are complete sentences in it - with very short examples, neural networks sometimes still lose to classical statistics. An easy way to switch from a regular statistical model to a neural network is to add a point.

4 Perspectives of translation technologies

4.1 Advantages and disadvantages of translation technologies

In technologies based on translation correspondences, the sequence of stages of translation process consists of:

1) Dictionary search, identification of lexical units;

2) The exclusion closely related phraseological units from further analysis;

3) A grammatical analysis of the most significant phenomena of the input text;

4) Translation of unambiguous and ambiguous words;

5) Grammar analysis and synthesis;

6) Synthesis of the output text (DePalma et al, 2018).

One of the most pressing problems of translation technologies is the lexical ambiguity. In the vast majority of cases, the lexical polysemy of a word is allowed within the phrase. In translation technologies, lexical ambiguity is resolved using a contextual dictionary. The source material for its compilation is a concordance dictionary, which is a list of textual uses of each word taken in the context of a certain size. When using parallel texts, one can compare each use of the word form in the input text with its translation in the output text. The translation algorithm for a multi-valued word is a sequence of context requests for the presence of determinants in it, i.e. words that uniquely identify one particular translation of a multi-valued word.

Thus, based on comparative data, the following advantages and disadvantages of translation technologies can be distinguished as follows. Advantages of translation technologies:

1) Availability of dictionaries on speciality, instant search, arrangement in several windows, the ability to simultaneously review several translation options;

2) The ability to create your own user dictionary;

3) Connection to a text editor Microsoft Office, which allows, without interrupting the original and its translation, to choose the appropriate correspondence for the translation;

4) The ability to translate from different languages;

5) Compactness, the presence of an almost unlimited amount of information in the computer;

6) Speed of search.

Disadvantages of translation programs:

1) Limited review due to screen size, a large dictionary that is opened allows you to see a much larger number of word meanings, especially if it has many meanings;

2) Dictionaries-translators do not follow the rules of grammar, stylistics and vocabulary, do not take into account wordplay, artistic techniques;

3) most often, text translators choose one of the meanings of a multi-valued word, which may not correspond to the context;

4) In the absence of a word in the dictionary the word is not translated;

5) Electronic dictionaries-translators often give out several options for translating a word into another language without explaining the differences in one or another dictionary correspondence, which makes it difficult to choose the right one or another correspondence in this context (Harding & Cortés, 2018).

The use of statistics allows machine translation systems to change with the language. If people start writing a word in a different way, the system sees this as soon as new texts get to it. To improve the quality of the translation, the system is regularly updated and checks are carried out. However, high-quality machine translation of texts is still unattainable. However, it greatly facilitates and accelerates the work of translators.

Professional translator tools are automated translation systems. It is with their help that you can save time; ensure the uniformity of the text and its quality. At the same time, machine translation can help, for example, in an emergency situation when a person does not know the language, but an urgent translation is needed. Even if it is of poor quality.

4.2 The impact of translation technologies on translators

Given the use of subject-oriented binary MT systems with a transfer in such a hybrid version, which are practical systems and based on preliminary terminological analysis of the corresponding subject area,

the user must take into account that the automatic dictionary is a nuclear part of any system, it is intended not only to convert text to lexical level, which is the lower level of analysis and transfer, but also to ensure the operation of the algorithm s of automatic parsing. When analyzing the results of the system's operation, it should be borne in mind that parsing is carried out within the framework of a single sentence, and not within the limits of over-phrasal unity, and moreover, not as a whole text. Therefore, with each new sentence, the MT system, as it were, starts the analysis anew, losing information about the boundaries of nominal and verb groups, functional segments, established in the analysis of the previous sentence. The developers of the MT systems are well aware of the flawedness of this approach, but it is rigidly determined by the requirement to switch at the initial stage of work from specific lexical units to their code designations. These semantic syntactic codes, the essence and variety of which depend on the system and the algorithms defined in it, are the basis for the application of analysis and synthesis algorithms that are universal for a particular language (Barbu, Eduard et al., 2016).

The translator’s practical work with the machine translation system provides for:

- preparation of the source text (array of texts) for translation - manual pre-editing of the text;

- editing the results of the MT system — manual post-editing of translations;

- maintaining your own (user) dictionary that records the results of working with machine translations and determines the configuration of the MT system for the tasks of a particular translator.

When implementing work at these stages, one should take into account the restrictions that are imposed on the results of the work of any MT system. These limitations are:

1) due to the local translation (translation by sentences), this feature leads to the fact that the system makes it difficult to analyze relationships within super-phrasal unity and search for antecedents, which leads to incorrect translation of substitute pronouns. Therefore, when preliminary editing the source text, it is necessary to pay attention to the use of such substituents and, if possible, replace them with relevant significant words;

2) due to the peculiarities of working with words that are absent in the dictionaries of the system (geo names and proper names, trademarks and rare words), which leads to possible violations in the

parsing of the input sentence. In addition, situations of incorrect recognition of proper names as common names and, accordingly, their translation are possible. When pre-editing, you should pay attention to the use of such names and label them so as to prevent their translation;

3) due to the variability of the use of terms in the source text, which may violate the unification of the translation of terminology within the same text. During preliminary editing, one should analyze the most frequent nominations (terms used), occasional abbreviations that may coincide in different terminological systems and languages for special purposes, as well as ways to use hyphen structures;

4) due to the fact that in real texts there are very long sentences, and in systems restrictions are introduced on the length of the sentence, in which the syntactic structure is recognized rather stably. This restriction can be removed by pre-editing very long sentences. Experience shows that with an average sentence length of 12 words, the MT result is optimal, but in the real text this indicator is often exceeded;

5) due to the linearity of recognition of stable collocations (machine revolutions), which make up most of the vocabulary of any machine translation system. Text pre-editing allows you to remove some of the limitations of MT systems in advance, it is necessary to establish the unity of the terminology used, for example, in data extraction systems.

Incorrect results often arise as a result of a discrepancy between the data extracted from the text and the nomination of the corresponding objects in the vocabulary (databases or ontologies). Pre-editing should be used to correct errors and in general to simplify the text in connection with the solution of problems of translation and knowledge engineering.

Pre-editing involves the following actions:

- introduction of articles in a foreign language text where necessary or grammatically justified;
- repetition of elements with a compound connection of phrases in a sentence;
- the introduction of unions using the union-free connection between proposals;
- elimination of constructions in brackets in the middle of the noun phrase or in the middle of the sentence;

- replacement of occasional abbreviations with full names or the introduction of special characters, preventing their translation as ordinary words;
- elimination of ellipses, informal constructions and metaphors;
- reduction to a single type of structure, which may have a different spelling.

The linguistic support of MT systems is usually implemented as a correlated system of automatic dictionaries (AD) and grammar rules.

In accordance with this approach, the automatic dictionary of the MT system can be functionally divided into 4 components:

1) dictionary articles of the so-called stop words, that is, service vocabulary that determines the involvement of specific parsing algorithms;

2) terminological vocabulary bases focused on the fixation of universal terms or multicomponent terms characteristic for use in specific subject areas or subdomains;

3) dictionary entries of general scientific vocabulary used in almost all scientific and technical texts;

4) dictionary entries of lexical units (words and phrases) added by the user to the so-called user dictionary.

This part of the AD is formed by the translator and / or terminologist within the framework of its own workstation and provides its finer tuning to the lexical spectrum of texts intended for translation. No matter how complete and focused on a narrow subdomain the AD is, the result of MT requires post-editing both at the level of the syntactic structure of the sentence and at the level of refinement and / or change of translations of individual words and phrases, as well as changes in morphological characteristics of gender, number, case, clarification of the forms of time and voice, changes in punctuation. When assessing the complexity of this process, stylistic changes are usually not considered. Paradoxically, process that causes the rejection of translators and a negative attitude towards the results of the MT as a whole. Studies (EU, 2017) showed that such a rejection is more common for professional translators than for those who are just getting this profession. Perhaps this is also connected with the level of computer literacy of the subjects, as well as with little translation experience. Long-term experience of

the author's own work shows that working with post-editing of MT results leaves room for solving creative and linguistic problems, however, teaching post-editing of MT results should be an integral part of the training of translators. Post-editing at the lexical level requires clarification and change of translations of specific lexical units, at the syntactic level - the transformation of the sentence structure. For example, in cases of translation from English into Russian, verification of agreement by gender, number and case, clarification of the place of the subject, sometimes a complete restructuring of the sentence or a transition to an indirect structure like *we have - we have - we have is needed*.

Post-editing the results of the MT and obtaining the final version of the translation of the text require an appeal to the dictionary and encyclopedic databases selected by the translator and included in the workstation, as well as to the pre-selected corpus of texts. As a result of work at the stage of the actual translation, a user dictionary is formed that clarifies the terminological features of a particular text. This dictionary at the stage of support of the selected machine translation system is included in its linguistic resources. Thus, after the translation of a specific text is completed, the linguistic resources must be reconfigured: the corpus of parallel texts is replenished by the source text and its translation, a user dictionary is formed and / or replenished, including the terminology identified and verified by the translator, and the dictionary database is replenished. Only the constant maintenance of our own machine translation system allows us to use it with maximum effect, setting up dictionaries for the necessary terminology and choosing convenient means and methods of post-editing.

Let us consider the features of the post-editing process and maintaining a user dictionary when the translator is working with the MT results.

The first part of the dictionary includes a special vocabulary that defines the reference information for the implementation of transfer algorithms, therefore this part of the dictionary - the stop-word dictionary is "untouchable" in the sense that no units (auxiliary and modal verbs, conjunctions, prepositions or homonyms with them) should not be included in the user dictionary even if the translator is not satisfied with the translation option selected in the system. This post-editing option can be fixed in replace mode throughout the text as the translator prefers, but not in the dictionary. The second

(terminological) part of the dictionary includes dictionary entries with verified descriptions of terminological units and their semantic-syntactic codes. This terminological base is focused on the subject area and, as a rule, the developers of the system or the terminologists responsible for this task are specially responsible in the team of those who prepare and translate the texts. The third part of the dictionary includes words of broad semantics, which cause the largest amount of post-editing. The fact is that the meaning and translation of these words are defined in AD by lexical units that determine the most generalized meanings that are included in the scope of the corresponding concept. The meanings of words of broad semantics are partially clarified by introducing phrasal verbs and phrases into the AD. Since the author has a great deal of freedom in the formation of new specifying phrases, the use of low-frequency expressions or not characteristic of the scientific style, no speakers are able to include all of them, respectively, when editing, these lexical units require special attention and creative tasks. A user dictionary is formed as a result of work at the post-editing stage, this dictionary captures the terminological features of specific texts with which the translator works. The study of the results of MT, scientific and technical texts, as well as the real volume of post-editing, allows you to recommend extreme caution when choosing new lexical units and their translations included in the user dictionary. It is necessary to analyze the entire text as a whole in order to understand how much these translations correspond to him (and not only to him). Only if compliance is established, it is possible to edit all uses of similar words and phrases in the replacement mode, and then enter them into a user dictionary for use in translating other texts from the same subject area. Today, the use of machine translation in scientific, technical and research projects, as well as for commercial purposes, is constantly growing. Serious advances in the quality of machine translation have led to the widespread use of MT by lay people to extract the essence of texts written in unfamiliar languages. Accordingly, there were special requirements for post-editing procedures and technologies (DePalma et al., 2018). At the same time, it can be argued that for the full use of the MT system, a professional translator should imagine in general terms the procedure analysis of the text in the system, which will allow him to prepare the text in advance so as to minimize the amount of post-editing. In addition, users of the MT system should well understand that the

quality of machine translation results depends on the tuning of the AD system to the tasks of a particular user. Taking into account the range and capabilities and limitations of the selected translation system will allow the translator to obtain a result that is easily edited using modern linguistic technologies. The correct use of the full range of these technologies is becoming particularly important today.

The use of statistics allows translation technologies to change with the language. If people start writing a word in a different way, the system sees this as soon as new texts get to it. To improve the quality of the translation, the system is regularly updated and checks are carried out. However, high-quality machine translation of texts is still unattainable. However, it greatly facilitates and accelerates the work of translators.

Today, translation needs to be done not only efficiently, but also quickly. Software products the common name of which is “Translation Memory” are designed for implementation of the principle “do not translate the same thing twice”. Translation Memory is a developed base of translated words, expressions, syntactic integers, which is created by the translator himself. In each new text, the program defines concepts that are already “familiar” to it and offers the specialist a translation from the database (Harding & Cortés, 2018). Thus, uniformity in the translation of documents is achieved and, of course, significant time savings.

However, there is no certainty concerning the particular roles translators will play in a globalized society, which is getting more and more technology-dependent (Mossop, 2017). Since there is a possibility that translation technologies can subsume the translation process, it is important to do our best to use effectively these technologies preparing for future changes. There is also a need to know what such tools as translation technologies can and cannot do as well as in which way users could use them to achieve the desired result.

So, I can state that, in spite of the fact that translation technologies have increased significantly productivity and quality in some fields of translation, there is also an increasing problem of translators to whom these technologies create considerable uncertainties and challenges.

5 Challenges in implementation

5.1 Translation technologies and quality of translation

In the general case, the perception and translation of text by a machine is determined by some model of automatic understanding, which, by the definition of a model, is a simplified (coarsened) representation of a person's real cognitive ability. This predetermines the fact that a machine, in comparison with a person, is able to recognize a much more limited subset of texts that fit into the used automatic processing model. This intuitive fact is also confirmed by experience working on prescriptive languages. For example, in (European Association for Machine Translation, 2020), the author indicates that when compiling rules for a prescriptive language focused on machine translation, a set of rules for understanding human text compiled by experts (editors, teachers, authors, translators) is only a subset of the set of rules for ensuring machine translation texts (42 of 59). Thus, machine translatability of texts is a particular, more rigorous case of translatability in general. Using machine translation to translate texts written in a machine-friendly language could provide high quality automatic translation. However, the bulk of the texts are created in unlimited natural language, i.e., the author of such a text does not use any prescriptive language that would take into account the peculiarities of using automatic translation as a whole or any separate machine translation system (while some such texts, for example, scientific and technical literature, it is still advisable to translate using automation tools, although with greater costs for pre- and post-editing). In this case, the machine translator is faced with problems that worsen the machine translatability of the text.

The complexity of the text for machine translation can be estimated in advance using the so-called translatability indicators or negative translatability indicator, that is, the linguistic features of the text that negatively affect the quality of the translation.

From the point of view of the machine translation system in (Harding & Cortés, 2018), the authors distinguish two types of such markers:

1. General markers: potentially cause difficulties for all machine translation systems. Obviously, the problems of text translation, characterized by these markers, are due to the internal structure of the

natural language, the polysemy of linguistic units. These markers are due to the fundamental impossibility (at least in the foreseeable future) in any formal model, discrete in nature, to take into account the continuity of the language (Robin, 2016).

2. Specific markers (for a specific MT system): cause problems for a particular machine translation system. The translatability problems associated with these markers may be due to two reasons:

a) the limitations of the natural language model in this system;

b) the specific features of a particular language. So, most of the markers (16 out of 20) are specific to the Japanese language and mainly represent certain lexical units and constructions used in the text in Japanese. It is also indicated for some markers that they are taken into account only for the German language. Another dimension to distinguish between translatability problems can be formulated depending on the stages of translation:

1) problems of understanding (analysis) of the original text;

2) problems of transferring (actually translating) the original text into the target language;

3) problems of synthesis of the text of the translation.

Thus, the problems of translatability of texts can be considered in several orthogonal dimensions:

1) subject of translation: person/machine;

2) stage of translation: analysis/actual translation / synthesis;

3) universality: dependence/independence on the translation system and a specific language. It should be noted that some markers can be assigned to several categories simultaneously.

Such a classification of markers indicates that the tools for pre- and post-editing should rely in their work, generally speaking, on various translatability markers. In addition, the discussion of the most complete assessment of translatability is appropriate only in conditions of limited analysis of a specific pair of languages and a specific machine translation system.

3. The problem of translatability of a scientific and technical text in Russian: we will focus our study on the body of scientific and technical articles and abstracts in Russian in the subject field “Mathematical Modeling”, therefore, we will consider features that depend on a specific natural language

in the context of this sublanguage. We will consider English as the language of translation. Let us consider in more detail the linguistic features of the texts of this sublanguage in order to identify problems that may cause difficulties in machine translation.

Translation problems (which can be identified by translatability markers) can manifest themselves at various linguistic levels:

1) the graphic level,

2) the lexical level,

3) the syntagmatic level.

Problems at the graphic level present a problem at the stage of analyzing the text of the original. Such problems are technical in nature. They, as a rule, are not difficult for a human translator, but can reduce the quality of automatic translation. The problems of the graphic level are associated with the variability of the use of various special characters and spaces, for example: “external” fragments vs. external fragments, direct / inverse heat equation vs. direct/inverse heat equation, Sturm- Liouville vs. Sturm-Liouville problem, input / output system vs. input-output system. Errors and spelling errors can also be attributed to problems of this kind.

The formal marker for the machine translatability of such a text is the lack of a unit in the dictionary and/or the inability to derive the unit according to the rules provided by the system (for example, the rules for generating forms of the morphological paradigm of a word). Problems of translatability at the lexical level are associated with:

a) insufficient vocabulary (vocabulary coverage);

b) the variability of lexical units;

c) asymmetry of the original and translation languages.

Vocabulary deficiency is a problem for both the human translator and the machine translation system. However, if a person in some cases can guess the meaning of a lexical unit based on context, for a machine translation system the absence of a lexical unit in a dictionary is critical. The variability of lexical units can take the following forms:

1) formal variability (variability in terms of expression),

2) semantic variability (variability in terms of content).

The problems associated with formal variability can be represented using morphological derivatives and transformations of phrases to denote the same concept.

Another aspect of formal variability is the use of paradigmatically related words and expressions when referring to the same referent (synonyms, hyperonyms), for example, a vertex of a graph/graph node, a bipartite graph/bigraph, a two-place function/binary function, the Gauss-Ostrogradsky theorem/theorem about divergence.

The problem for a person in the case of formal variability is caused by the fact that he can doubt whether these terms are morphological or lexical options to refer to one concept or indicate two different concepts. In machine translation, the problem of morphological variability actually comes down to the previously mentioned lexicon coverage problem: if both options are indicated in the dictionary, they have an interpretation given by the language model of the system. In a machine translation system, such an interpretation can be explicated by means of equivalents in the target language: the same translation is given for variants of one concept. Semantic variation is represented by polysemy and homonymy (including grammatical homonymy) of various units. A person, as a rule, can easily restore the lexical and grammatical and semantic characteristics of terms from the context, while for a machine it is necessary to determine the procedure for resolving the ambiguity due to semantic variability. The solution to this problem is determined by the method of modeling the language in the system. The restriction of the subject area can greatly contribute to reducing the polysemy of significant parts of speech. However, the service parts of speech, in particular, prepositions, are for the most part similar in different subdomains, and the ambiguity inherent in the service parts of speech is largely preserved. Some prepositions have different ways of translation.

Length (the number of words in a sentence) is not a linguistic characteristic of sentences. However, as is widely noted by researchers in the field of machine translatability, the length of a sentence is significantly correlated with syntactic complexity: the longer the sentence, the more likely it is to contain

subordinate clauses, appositive constructions, prepositional constructions, and other elements that complicate the syntactic structure. Too long sentences, even for a person can be difficult interpret. In machine translation, grammatical models embedded in the system are not always able to correctly process such language material.

When working with complex sentences, it is necessary to develop techniques for breaking them into parts in order to operate with syntactically simple blocks (Translating Europe Forum, 2018). For example, in scientific texts on mathematical modeling, the average length of a sentence is 14.84 words, with a sentence with a maximum length of 85 words. Of the 5353 sentences of the corpus, 1321 (24.68%) have a length of 20 or more words (Underwood & Jongejan, 2001). A sufficiently representative collection of such sentences indirectly indicates the syntactic complexity of the text.

Plug-in constructions are represented by words, phrases, sentences, which contain clarifications, explanations, corrections to what has been said, indicated in brackets or highlighted with commas. Structurally plug-in constructions can be simple (single-word) or common (consisting of several words). In the texts of annotations to articles on mathematical modeling, such constructions can be expressed: A. Adjectives: The question of the possibility of solving control problems with functional constraints in a narrower (classical) set of strategies - positional strategies also remained open. B. Nouns and nouns (including groups spread by participles).

G. Suggestions (including elliptic).

Plugin constructions, firstly, can break syntactically closely related blocks, thereby complicating the parsing of sentences and the search for contextually equivalent terms.

If the units are ambiguous, the distant arrangement of the context separated from the ambiguous unit by the positive insertion complicates the machine “understanding” of this unit and, as a result, the translation in the general case turns out to be incorrect. So, the inverse problem has a stable equivalent to the inverse problem, however, a sufficiently long positive insertion separating the phrase increases the likelihood that the machine will not be able to recognize the words inverse and the task as contexts for each other, and then each of the units can be translated any combination of equivalents defined by the

lexicon of the system for these polysemous words (for example, * converse task, * reverse problem, etc.) Secondly, the translation of inserted nominative words can be problematic construction.

Syntactically, such constructions act as:

A. Additions to the element of the main text.

B. Paraphrases syntactically equivalent to a noun phrase from the main text.

C. Constructions equivalent to a single title sentence not syntactically related to other elements of the text.

The case homonymy in the studied case is most often manifested for the genitive and instrumental case. Nouns in the genitive can function as a definition to another noun or fill in a separate valency with a predicate.

The prepositional group located in the postposition can fulfill the function of determining one of the preceding nouns or fill in a separate predicate valency.

When translating into English, this can cause a problem if there is a need for a permutation in the translated text, which corresponds to two cases:

1) the prepositional group is a definition of a noun that performs the function of the subject, or defines a noun, and in Russian there is a reverse order words. In this case, when translating, it is necessary that the predicate be located after the subject, so the problem of determining the boundary of the subject arises. In our body of texts, this situation is typical of sentences with reflexive verbs in the initial position.

2) a typical sentence structure in English for a predicate requires the presence of a circumstance expressed by a sentence group at the beginning of a sentence. In the corpus of scientific and technical texts, this mainly concerns circumstantial constructions expressing a goal or condition. In the text they are expressed using prepositional groups with prepositions for, with.

It should be noted that this problem can be partially solved at the vocabulary level - by calculating terminological phrases with prepositional definitions in the dictionary. For example, it is advisable to include in the dictionary such stable terms as average derivatives, a partial differential equation, etc.

Management causes problems in case of differences in two languages. This problem is of particular relevance with the discontinuous arrangement of the predicate unit (verb, verb noun) and the prepositional group (with a contact arrangement, it is advisable to lexicographically solve the problem).

Today, there are various machine translation systems, some of which perform translation at a very high level. However, in order to identify which systems are suitable for working with certain languages or types of texts, it is necessary to test these systems and analyze the quality of translation based on large fragments of texts. This is part of the task of linguistic translators, and the developers of automated translation programs try to correct errors in the system, develop and improve their products, based on the results and theoretical justifications of specialists in the field of translation.

To the main factors hindering machine translation the researchers include:

- linguistic ambiguity, which can be both lexical and grammatical in nature;
- the presence of complex syntactic structures that can vary significantly in the original language and in the language of translation;
- differences in word order in sentence (direct/reverse, strict/free);
- the presence of anaphoric connections in the text;
- the presence of idioms, the meaning of which cannot be conveyed through word-for-word translation;
- the presence of neologisms;
- the existence of cultural differences among language communities, etc. (Robin, 2016)

In literature, it is practically not noted that the quality of machine translation depends not only on the level of development of a particular MT system, for example Google Translate or PROMT, but also on specific pairs of languages, direction of translation and linguistic features of the translated text. The latter can act as the so-called translatability indicators (Zaretskaya et al., 2015). This term arose in connection with studies on machine translation and refers to graphic, lexical and syntactic elements of the text that provoke errors in machine translation. For the user of MT systems, the ability to recognize translatability indicators (TI) in the text has undeniable advantages. Firstly, the presence of TI in the text

predicts the quality level of machine translation, and with a large number of them, you can either refuse machine translation of the text, knowing that the translation will be deliberately erroneous or which is much more promising, eliminate the translatability indicators from the source text by editing “Bad” for machine translation of text fragments. There are two types of TI: universal and specific. Universal translatability indicators cause problems with machine translation of any kind of texts in any languages. Specific indicators of translatability are characteristic of specific languages, directions of translation, subject areas and systems of MT. For example, when evaluating the translatability of patent texts from English to Danish in the PaTrans system, it was found that the most problematic for this system are English prepositional groups, adverbs and subordinate clauses at the beginning of the sentence, as well as long sentences with adverbs. To improve the quality of MT, special techniques are being developed to eliminate translatability indicators from the source text, focused on specific machine translation systems. The most popular of these methods is the development of rules restricting the use of lexico-syntactic equipment, which is often called the “controlled language (Mossop, 2017)”.

Among the most well-known automatic translators can distinguish machine translation systems “PROMT” and “Google Translate”. The automatic translator “PROMT” works on the principle of “translation according to the rules”. The technology of this translation consists in the application of algorithms, according to which the program analyzes the text and on the basis of the analysis it synthesizes a translation variant. It is believed that the work of such a machine translator is similar to the process of human thinking (Mossop, 2017). The Google Translate machine translation system operates according to a fundamentally different technology based on a statistical calculation of the probability of matches. This system uses many databases of parallel texts in which word-combinations and their translations are stored in pairs. In the translation process, a statistical analysis is carried out: the system selects the equivalent for the translation based on the frequency of use, and as a result substitutes the option that has the highest percentage of matches. It should be noted that “Google Translate” was originally developed for translation from and into English, and still English is an intermediary language when working with other pairs of languages. In other words, the translation is not carried out directly: first, the text is transferred from the

original language to English, and only then - to the required translation language, which largely affects the quality of the translation.

The objectives of the experiment were to verify the following hypotheses:

1. Different MT technologies in the general case give different quality of translation of the same linguistic phenomena in the text, that is, they are characterized by a different set of indicators of translatability and / or different relevance the same IP for translation quality.

2. Based on the identified indicators of translatability, it is possible to formulate the rules of the controlled language, when used for the purpose of editing, the quality of machine translation of the text is significantly improved. To test the hypotheses from a section in the previous English section, a sub-corpus of sentences was compiled. At the first stage of the experiment, all the proposals of the sub-corpus were translated into Russian by the PROMT and Google Translate systems and the obtained results were analyzed. In this case, both Russian translations were analyzed (the lexical and grammatical correctness of the translation, the observance of stylistic norms and lexical coverage) were taken into account, as well as the linguistic phenomena of the original English sentences that caused the translation errors, i.e., actually English indicators of translatability.

At the second stage of the experiment, all previously obtained data on errors in the English-Russian translations of each of the PROMT and Google Translate systems, as well as on the translatability indicators that caused these errors, were analyzed and systematized.

Table 1. shows that the most frequent errors in translations of the PROMT and Google Translate systems are terminological.

Table 1. Comparative statistics of types of errors in the English-Russian MT.

Type of error (frequency)			
	Terminological	System Lexico-semantic	Syntactic
PROMT	104	61	99
GOOGLE	42	27	50

Terminology, especially multicomponent, is a very dangerous indicator of translatability and, therefore, the user must always check the correctness of automatically translated terms either in terminology dictionaries or with the help of Internet search engines.

As for electronic open online dictionaries, they give too many translation options that are difficult to navigate and may contain translation errors, because users who introduce new terms into these dictionaries, not 50, are usually caused by such translatability indicators as coordinating unions. The indicator of translatability “and” is especially dangerous when it is used repeatedly within the framework of one English sentence and the distant arrangement of dependent words. To help the translator in the process of this study, in parallel with highlighting the indicators of translatability and identifying incorrectly translated terms, we compiled a glossary of multicomponent terms of macro- and mega economics. Next in importance (but not in frequency) for the analyzed sublanguage and MT systems are lexical and semantic errors that distort the meaning of the text, which, of course, is unacceptable. Mistakes like terminology should be prevented first.

The indicators of the translatability of lexical and semantic errors in the Russian translation are mainly:

a) English multi-valued lexemes in the framework of one or different parts of speech and

b) Syntax ambiguity, which leads, for example, to incorrect choice of a support word in a phrase or to incorrect determination of homogeneous sentence members.

Last, syntax errors caused by a number of translatability indicators are very common, the main ones are: length and complex structure of sentences, distant arrangement of dependent members, union “and”, impersonal sentences, sentences in the passive voice, non-union adverbial clauses, chains nouns, as well as prepositions, which, as a rule, are ambiguous and may require nouns after themselves in various cases.

In the Table 1., stylistic errors are not included, since they are often quite a consequence of other types of errors and, in general, do not impede the understanding of the translated text. Stylistic errors can also be caused by indicators of translatability of the English text, such as an English word order different

from that adopted in the Russian language and a choice from all synonyms of words such that the Russian-speaking user gives the feeling that “they don’t say so”.

Based on the results of the second stage of the experiment, in order to eliminate the translatability indicators of the original English text by rephrasing (pre-editing) the problematic fragments, the rules of the controlled language were formulated. The following is a snippet of such rules for English texts:

1. Verify the translation of terminology according to the method described in (Moorkens & O’Brien, 2015).

2. Sentences longer than 20 words are divided into shorter ones.

3. Replace ambiguous tokens with unambiguous.

4. The chains of nouns are divided into groups with the pretext “of”.

5. Add an alliance before a union-free subordinate clause.

6. Complex sentences containing a multiple compositional connection, split into simple and / or duplicate prepositions.

The resulting translations were primarily checked for correct terminology, and if necessary, the translation of the term was corrected either using the glossary we compiled during the experiment, or using Internet search engines using the methodology (Robin, 2016). The experiment showed that the rules of a controlled language for pre-editing when translating texts from English into Russian using the Google Translate and PROMT MT systems significantly increase the quality of products of both systems, but not to the same extent. To obtain high quality translations, these systems generally require the use of different sets of rules for a controlled language. At the same time, for the English texts of the analyzed subject area, the Google Translate system requires significantly less effort for editing (and fewer rules of the controlled language) than PROMT, and can be recommended as a priority.

To describe translation errors, it is necessary to determine their classification. It is customary to distinguish between errors and translation transformations caused by typological differences between the source language and the target language; errors are spoken about in the case of unjustified transformations, logical inconsistencies, and violation of the norms of the target language (methodology

(Robin, 2016). Although the approach to the classification of errors is different in the academic environment and in translation practice, we can distinguish categories of errors that are not objectionable to either theorists or practicing specialists. Such a generally accepted classification is proposed in the work of D.M. Buzaji et al (Bughin & Manyika, 2018). The authors propose to distinguish between the following categories of errors:

1) denotative (distortion of the denotative text content, violation of the logic of the presentation of information),
2) discursive (distortion of genre and style features that impede the functioning of the text in the host culture),
3) expressive (distortion of the author's attitude to the object of description),
4) violation of the norm and Uzus of the target language. It is this classification that we take as a basis, realizing that we can apply it to the description of errors in translation made on the CAT platform with some reservations and additions developed in corpus linguistics for marking up a parallel educational building. The use of machine translation and translation memory suggests the likelihood of untranslated segments of the source text; it is precisely this kind of error that turns out to be the most frequent in machine translation (Bughin & Manyika, 2018).

Skipping segments of the source text in machine translation differs from the omission error that translators often make. Machine algorithms often do not "see" the end links of a chain of syntactic dependencies and miss them during translation, do not recognize inconsistent definitions and do not distinguish between nouns and adjectives derived from them (Sakamoto et al., 2017), while omitting information when translating "manually" is usually not related to the syntactic position of the segment in the source text.

5.2 Machine translation and post-editing

Let us compare the examples of translations made by the PROMT and Google Translate systems from English into Russian. When translating a fragment of a technical text - the operating instructions for an electronic device (game console) - the following results were obtained:

Original: *Keep food and beverages away from product. Do not spill liquid of any kind on this product as it may render it inoperative. Clean with a slightly damp cloth (cold water). Do not use soap, detergent or other chemicals. Never submerge the unit in water.*

PROMPT: *Держите отдельно еду и напитки от продукта. Не проливайте жидкость никакого вида на этом продукте, поскольку это может отдать его недействующий. Чистый с немного влажной тканью (холодная вода). Не используйте мыло, моющее средство или другие химикаты. Никогда не погружайте единицу в воду.*

Google Translate: *Держите еду и напитки от продукта. Не допускайте попадания какой-либо жидкости об этом продукте, как это может сделать его неработоспособным. Очистите со слегка влажной тканью (холодная вода). Не используйте мыло, моющие средства или другие химические вещества. Никогда не погружайте прибор в воду.*

It should be noted that both systems correctly recognized imperative forms used in the original text. The translator "Google Translate" correctly identified the cross-language affiliation of all language units. This is his advantage compared to PROMT, which chose the adjective clean to translate the tokens clean, rather than the verb clear. This problem is associated with the phenomenon of grammatical homonymy that is widespread in the English language. When translating according to grammatical rules, PROMT incorrectly decrypted the original syntax and made a mistake, because the program did not take into account the fact that English is an analytical type of language with a fairly strict word order and the predicate in the sentence: it is it the transfer was not recognized. "Google Translate" determined the presence of the verb in this sentence and its grammatical form (imperative), but did not recognize its specific meaning, since the category of the species in English is not obligatory for each form of the verb. In this sentence, a professional translator would definitely use an imperfective verb to emphasize the frequency of the prescribed action. The above translations show that both systems do not always cope well with the choice of controlling verbs and often work at the level of phrases standing in close proximity to each other.

At the same time, when the correct choice of control was made, we have the following examples:

PROMPT: *Держите отдельно еду и напитки от продукта. Не проливайте жидкость никакого вида на этом продукте, поскольку это может представить его недействующий. Чистый с немного влажной тканью (холодная вода). Не используйте мыло, моющее средство или другие химикаты. Никогда не погружайте модуль в воду.*

Google Translate: *Храните еду и напитки вдали от продукта. Избегайте попадания каких-либо жидкостей об этом продукте, как это может сделать его неработоспособным. Очистите со слегка влажной тканью (холодная вода). Не используйте мыло, моющие средства или другие химические вещества. Никогда не погружайте прибор в воду.*

Google Translate: *Храните еду и напитки вдали от продукта. Избегайте попадания каких-либо жидкостей об этом продукте, как это может сделать его неработоспособным. Очистите со слегка влажной тканью (холодная вода). Не используйте мыло, моющие средства или другие химические вещества. Никогда не погружайте прибор в воду.*

Differences are also found in the translation of adverbs. So, “PROMT” translated the dialect away using the dialect separately and “Google Translate” omitted this dialect, limiting itself to translating the preposition from. However, using the function of viewing and selecting available translation options allows us to see that in Google Translate the units away and from are considered as a couple and more accurate translation options are *вдали от, далеко от*.

The PROMT system has the function of choosing a special terminological field. When translating this passage with the connection of the thematic area “technology: gadgets”, the following text was received:

The above translation indicates that the user is most often offered options that practically do not differ lexically or grammatically. However, if we compare the original and the translation of the second sentence, we see that the system made grammatical and lexical transformations, replacing the imperative of the verb spill in a negative construction (Do not spill liquid of any kind) on the construction “verb + noun” (do not allow / avoid getting into it), but could not recognize the correct control for the

noun (getting somewhere). When translating the second part of this sentence, the correct version was not given for the multi-valued union as used to express a causal relationship.

We can conclude that both systems recognize the inclination well; “PROMT” makes fewer mistakes in determining the part-membership of lexical units; both systems do not always cope with the choice of the correct control, the selection of the correct lexical equivalent and the norms for using the verb form in the Russian language. However, despite the presence of errors and inaccuracies made by the programs, we can conclude that the general meaning of the translated fragment is clear, and after making corrections to the translations made by each of the systems, you can get a text that will fully comply with the norms of the Russian language.

Let us cite as another example translations of the text of another genre - an encyclopedic article.

Cybernetics arrived in the late 1940s out of information theory and the understanding of the importance of feedback in communication and control. It has had a profound effect on the communications and information technology industries where it has been seen as a key to designing sound telecommunications and information systems. Cybernetic comes from the Greek word for `steersman' and the idea of `steering' an organisation in the right course has become a metaphor for good management. However, Stacey (1999) has argued that importing an engineering model of control built on cybernetic theory, which draws heavily on ideas of stability, predictability and consensus which go back to Fayol, into management ignores our lived experience of not knowing and of not being in control, of seeing things as unpredictable and yet, in practice, even in unpredictable situations, coping as human beings.

“Google Translate”: *Кибернетика пришла в конце 1940-х годов из теории информации и понимания важности обратной связи в коммуникации и управлении. Он оказал глубокое влияние на отрасли связи и информационных технологий, где он рассматривается как ключ к созданию надежных телекоммуникационных и информационных систем. Cybernetic происходит от греческого слова «рулевой», и идея «управления» организацией в правильном направлении стала метафорой для хорошего управления. Тем не менее, Стейси (1999) утверждает, что импорт инженерной модели управления, основанной на кибернетической теории, которая в значительной*

степени опирается на идеи стабильности, предсказуемости и консенсуса, которые восходят к Файолю, в управление игнорирует наш жизненный опыт незнания и отсутствия контролировать, воспринимать вещи как непредсказуемые, а на практике, даже в непредсказуемых ситуациях, справляться как люди.

The initial text fragment consists of two narrative sentences, the first of which is complex and contains a large number of constructions based on subordinate connections between their components. The second is a simple common sentence with homogeneous members, gerund and complex syntactic constructions. Translations allow you to see that both systems do not recognize complex syntactic relationships, incorrectly determine the predicate, object, attribute and other relationships in the sentence. Completed translations are, rather, an arbitrary set of words and phrases that are not united by semantic links. They, unlike the previous example, cannot be edited without referring to the original.

After analyzing the text material, we identified formal markers that correspond to the identified linguistic problems of the text and can adversely affect the translatability of texts. It should be noted that to identify formal markers, preliminary marking of the text in parts of speech with the help of a lexicon is required, including for each unit of text a set of all possible labels, without removing the ambiguity.

Highlighted markers and the corresponding translation problems include:

1) the presence of sentences with a length of more than 20 words: syntactic complexity;

2) the number of words with lexical and grammatical labels from various entries in the dictionary: lexical homonymy;

3) the number of words with lexical and grammatical labels from one article dictionaries: grammatical homonymy;

4) the presence of prepositions with, from, from: lexical ambiguity;

5) the number of brackets in the sentence: separation of context-sensitive elements;

6) the presence of nouns marked with a instrumental case: case homonymy;

7) the presence of nouns marked with a genitive: case homonymy;

8) the presence of the sequence “noun in the indirect case + preposition”, which is after the predicate*, and before the predicate there is no noun in the nominative case*: polysemy of the syntax of the prepositional group;

9) the presence of verbs and nouns for which there is an article in the dictionary that includes the preposition (for example, it recognizes in the text it is recognized, and in the dictionary there are two entries - it depends on discontinuous arrangement of the controlling preposition;

10) the number of combinations of the adjective * (but not the participle *) + preposition: ellipsis;

11) the presence of an adjective * or participle at the end of a sentence: ellipsis;

12) the quantity and / or + noun in the genus; and / or + noun in tvp: case homonymy;

13) the presence of and / or after the predicate *, and before the predicate there is no noun in the nominative case *: the boundary of the subject, including the composition;

14) the presence of active predicates in the form of plural: the need for syntactic transformation;

15) the presence of two or more verbs in finite form: syntactic complexity, in particular, the boundaries of composing predicate constructions or complex valency.

An asterisk indicates that a word is recognized as a given form / part of speech if there is at least one corresponding mark in its mark up.

Highlighted markers allow you to formalize and automate the assessment of the "suitability" of the text for machine translation: the presence of too many markers indicates the need to pre-edit texts or indicates the need to abandon the use of a machine translator. Today, various automation tools are being developed for evaluating the translation (targeting certain systems and languages), allowing to give a numerical characterization of the text's translatability (to evaluate the translatability index) using an automatic assessment of the presence of markers in the text. The procedure for calculating such an index is applied, which consists in the fact that a priori each sentence is evaluated by the maximum number of points, and for each marker, taking into account its weight, penalty points are calculated that are subtracted from the initial assessment. In this case, the weights of the markers may depend on the pair of analyzed languages or the machine translation system used. In addition to weight, the number of markers of a

certain type (the degree of their "expression") in each sentence can also be taken into account. A separate area of research in the field of machine translatability is represented by methods of preparing text for machine translation. Markers highlighted by us can assist in compiling a prescriptive language for annotation authors to improve the quality of their machine translation. The creation of texts for machine translation using prescriptive languages (Underwood & Jongejan, 2001) is one of the key strategies for ensuring high quality automatic translation. The use of machine-oriented prescriptive languages aims to remove (as far as possible) translatability markers from the original text in order to improve machine translation and reduce post-editing labour. As a result of the analysis of markers at various linguistic levels, rules are determined that require the removal (avoid), replace or add certain elements in problematic cases. Along with the rules of the prescriptive language, automated pre-editing tools may also be involved in preparing the text for translation. A unique approach is that it is possible to compose text for translation in an interactive mode, which avoids a number of analysis problems at the syntagmatic level. Another use of translatability markers is the creation of tools for automating the post-editing of machine translations.

Such tools can reduce the cost of finding errors in the translation and their correction. We also note that the assessment of the effort required to post-edit machine translation is one of the approaches — along with the previously noted weighting of translatability markers in the original text — used to evaluate the translatability of the text.

The presented analysis allows us to conclude that the machine translators "Google Translate" and "PROMT" are unsuitable for translating texts that are replete with complex and complex sentences, common definitions, metaphorical comparisons, complex syntactic structures, etc. Translation of such texts can only be performed man, since only he is capable of carrying out deep cognitive speech-thought processes. However, machine systems can be used to translate texts from English into Russian that are sufficiently formalized in the speech plan (technical documentation, consumer instructions, contracts, etc.). However, in this case, the translations made by the machine require compulsory editing and refinement "manually". Professional translators, from whom society usually expects high-quality

translations, can resort to the help of PROMT and Google Translate when translating texts that are limited in scope, since these computer systems often offer quite adequate, editing-free options translations for individual sentences or small fragments of text, or options requiring minimal proofreading, which takes an experienced translator little time. Thus, the PROMT and Google Translate systems, when used correctly, help speed up the translation process and optimize the work of practicing translators.

Under the influence of Internet access and the use of computer tools, translators often have to choose from the proposed translation options rather than generate their own. This shift does not abolish the "classical" requirements for the work and professional competence of the translator, but adds new tasks, the solution of which requires new knowledge, skills. The shift from generation to selection of ready-made translation options in essence implies the need for serious work to edit the completed text of the translation, removal from the "bottom-up" mode used in the translation and analysis of the source text in the "top-down" mode to understand the semantic structure of the text and the role of each its components in achieving the communicative goal set by the author (Underwood & Jongejan, 2001).

6 Conclusion

In fact, the technological and psychological prerequisites for the occurrence of specific translation errors of MT and on the CAT platform are somehow related to new competencies in the field of technology and organization of project work in the context of intercultural communication, which are necessary for a modern professional translator (European Masters in Translation Competence Framework - 2017). On the other hand, the presence of specific translation errors enables software developers to identify problem areas and see prospects for improving MT and CAT tools and the platform user interface.

Basic linguistic and professional competencies of the translator play a decisive role, and new ones can increase the speed and efficiency of work. “SmartCat” training can be introduced into the training program for future translators (Translating Europe Forum, 2018). When working on a CAT platform, the translator needs to learn to master technological innovations in the course of work, to have the skills of searching and extracting information, to develop an analytical cognitive style, which is manifested in attention to details and the ability to create clear algorithms of actions and adhere to them. As the most important abilities of the modern translator can be designated the ability to independently master the new, the ability to verify information and the ability to process the text without losing sight of the details.

New professional requirements are determined by social factors characteristic of the current state of intercultural communication in various professional fields:

1) the generally accepted norm of network communication in English;
2) joint remote work in a multinational team;
3) the spread of computer-mediated professional communication;
4) improvement and widespread use of machine translation systems and CAT tools.

Computer translation tools cannot completely replace a professional translator. Their use reduces the time required to work on large projects. The key figure on the CAT platform is the translator-editor, who is able to choose the appropriate translation options, to correlate the content of the translated segment with the meaning and communicative purpose of the text in the target language.

During the experiment, research hypotheses were verified and confirmed that, on the one hand, despite the difference in technologies that automate the translation, both systems cannot cope with the translation of certain linguistic phenomena of the sublanguage (translatability indicators), on which the quality of the translation depends, and on the other hand, the difference in technologies of text processing naturally leads to the fact that, in addition to general systems, these systems are characterized by translatability indicators specific to each of them. It is shown that translatability indicators can be eliminated from the source text using pre-editing based on a controlled language that takes into account translatability indicators, and thus significantly improve the quality of automatic translation without resorting to system developers. In cases where the translated text must meet high criteria for the quality of the translation, it may be necessary to combine pre-editing and post-editing.

Bibliography

Aranberri, Nora et al. (2016). Tools and Guidelines for Principled Machine Translation Development. In LREC.

Barbu, Eduard et al. (2016). The first automatic translation memory cleaning shared task. Machine Translation, 30 (3—4), 145—166. doi.org/10.1007/s10590-016-9183-x. Burchardt, Aljoscha et al. (2017). A linguistic evaluation of rule-based, phrase-based, and neural MT engines. The Prague Bulletin of Mathematical Linguistics, 108(1), 159—170. doi: 10.1515/pralin-2017-0017.

European Masters in Translation Competence Framework — 2017. European Commission, EU, 2017. Retrieved from: https://ec.europa.eu/info/sites/info/files/emt_competence_fwk_2017_en_web.pdf

Bou, Gaelle. (2016). "How does Neural Machine Translation work?" In Systran Blog (blog), 17/10. http://blog.systransoft.com/how-does-neural-machinetranslation-work/. (Last accessed 11/02/2020).

Bughin, J., & Manyika, J. (2018). "The promise and challenge of the age of artificial intelligence". McKinsey Global Institute, https://www.mckinsey.it/idee/the-promise-and-challenge-of-the-age-ofartificial-intelligence, 30/10. (Last accessed 11/02/2020).

DePalma, Donald. A. & Pielmeier, Hélène, Stewart, R. G. (2018). The Language Services Market: 2018. Cambridge, MA, USA: Common Sense Advisory, Inc.

European Association for Machine Translation: http://www.eamt.org/mt.php. (Last accessed 11/02/2020).

Foundations of Artificial Intelligence - Neural Networks Building Artificial Brains http://www.cs.nott.ac.uk/~pszrq/files/2FAIANNa.pdf. (Last accessed 11/02/2020).

Harding, Sue-Ann & Cortés, Ovidi. C. (eds.) (2018). The Routledge handbook of translation and culture, London; New York: Routledge, Taylor & Francis Group.

Ivleva, Marina A. et al. (2017, October). Cloud Platform SmartCAT in Teaching Future Translators. International Conference on Linguistic and Cultural Studies Springer, Cham. 155—160. doi: 10.1007/978-3-319-67843-6_19.

Gollner J. Information 4.0 for Industry 4.0 // Towards a European Competence Framework: Tekom-Jahrestagungund tcworld conference in Stuttgart. Zusammenfassungen der Referate. – Stuttgart: tcworld GmbH Verantwortlich, 2016. – P. 93–94.

Lacroix F. Writing for the 21st Century // Towards a European Competence Framework: Tekom-Jahrestagungundtcworld conference in Stuttgart. Zusammenfassungen der Referate. – Stuttgart: tcworldGmbHVerantwortlich, 2016. – P. 102–106.

Mossop, Brian. (2017). Conflict over Technology in the Translation Workplace. https://www.youtube.com/watch?v=aAPQ_4_LRwM. (Last accessed 11/02/2020).

Moorkens J., O'Brien S. Post-Editing Evaluations: Trade-offs between Novice and Professional Participants // EAMT 2015. Proceedings of the 18th Annual Conference of the European Association for Machine Translation. – Antalya, Turkey, May 11–13, 2015. – P. 75–81.

Muscan M. On the translatability of cultural specifics in economic texts. URL: https://www.researchgate.net/publication/298973806_On_the_translatability_of_cultural_specifics_in_economic_t exts (accessed 21.02.2020).

Online translator PROMT. URL: https://translate.ru (accessed 20.02.2020).

Siranush M. Translation Peculiarities of Economic Texts. URL: http://publications.ysu.am/wpcontent/uploads/2015/12/Siranush_Meloyan.pdf (accessed 21.02.2020).

Robin, Edina. (2016). The Translator as Reviser. The Modern Translator and Interpreter. Budapest: ELTE Eötvös Kiadó, 45—56. Retrieved from: http://www.eltereader.hu/media/2016/04/HorvathTheModernTranslator.pdf#page=46. (accessed 22.02.2020)

Sakamoto, Akiko et al. (2017). No news is good news? Translation Spaces, 6 (2), 333—352. doi.org/10.1075/ts.6.2.08sak.

Tagkas P. Translation of Economic Texts: Challenges and Limitations. In: Arslan, F., ed. Contemporary Issues on Linguistics and Language: LILA‘14/Linguistics and Language Conference Proceedings Istanbul: DAKAM Publishing, June 2014, pp. 279–287.

Translating Europe Forum. 2018. How is artificial intelligence transforming society the translation industry? Brussels, https://www.youtube.com/watch?v=2qOC1BlZGvc&feature=youtu.be (Last accessed 11/02/2020).

Underwood, N.L., Jongejan B. Translatability Checker: A Tool to Help Decide Whether to Use MT. Proceedings of MT Summit VIII (Santiago de Compostela, Spain), 2001, pp. 125–133.

Web service Google Translate. URL: https://translate.google.ru (accessed 20.02.2020)

Zaretskaya Anna et al. (2015). Integration of Machine Translation in CAT tools: State of the art, evaluation and user attitudes. Skase Journal of Translation and Interpretation, 8 (1), 76—89. Retrieved from: http://www.skase.sk/Volumes/JTI09/pdf_doc/04.pdf. (accessed 24.02.2020).