Journal of Theoretical and Applied Information Technology
28th February 2019. Vol.97. No 4

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

USE OF MACHINE TRANSLATION TECHNOLOGY FOR UNDERSTANDING SCIENTIFIC AND TECHNICAL TEXTS

[1]ILONA KOSTIKOVA, [2]ALONA SHEVCHENKO, [3]LIUDMYLA HOLUBNYCHA, [4]NATALIYA POPOVA, [5]VIKTORIIA BUDIANSKA

[1] Doctor, Full Professor, Head of the Department of Theory and Practice of the English Language, H. S. Skovoroda Kharkiv National Pedagogical University, Ukraine
[2] Ph.D., Associate Professor, Department of English Oral and Written Speech Practice, H. S. Skovoroda Kharkiv National Pedagogical University, Ukraine
[3] Doctor, Full Professor, Department of Foreign Languages #3, Yaroslav Mudryi National Law University, Ukraine
[4] Ph.D., Associate Professor, Department of Foreign Languages #3, Yaroslav Mudryi National Law University, Ukraine
[5] Ph.D., Associate Professor, Department of Pedagogic, Foreign Philology and Translation, Simon Kuznets Kharkiv National University of Economics, Ukraine

E-mail: [1]ilonakostikova@gmail.com, [2]alyonageorgievna@gmail.com, [3]golubnichaya11@gmail.com, [4]nataliapopova1971@gmail.com, [5]vikkbud@i.ua

ABSTRACT

The article deals with machine translation technology, which is being discussed by several sciences such as Computer Science, Programming, Linguistics recently. The purpose of this study is to investigate how the technology of machine translation can be used by scholars, scientists, users in order to understand scientific and technical texts.The principal results are the comparison of different systems (traditional, statistical, hybrid and neural) of machine translation, assessment their advantages and disadvantages, limitations, drawbacks, and errors; experimental analysis the quality of the scientific and technical text made by machine translation systems Google Translate and Pragma Online. Common and different in both translation systems have been proved experimentally. The obtained results were processed and analyzed. The methods of analysis, synthesis, comparison and classification were used as well as the method of mathematical statistics Student's criterion in the experimental research. The major conclusions are we definitely have all the grounds for advising Google Translate to use for general understanding scientific and technical texts.

Keywords: *Technology; Machine Translation; Neural Machine Translation; Google Translate; Pragma Online; Scientific and Technical Text*

1. INTRODUCTION

Information technologies create a single information environment of world importance. However, the language barrier becomes an obstacle to understanding the information. Seeking to overcome such an obstacle requires a reasonable choice of the most appropriate technology means. One of them is the rapid translation of necessary information from a foreign language with the help of innovative machine translation systems. They give possibility to understand the general essence of information and to work operatively with large volumes of documents.

The mentioned technology means facilitate the transfer of science technology. The latest scientific developments in various technical fields become available, their results can be used directly in modern developed production. Given this, the scientific and technical translation is a tangible sphere for the application of modern machine translation, since it is a translation of a formal and logical style, which is characterized by the accuracy of the wording and the uniqueness of the terminology.

Modern scholars face a contradiction, connected with misunderstanding the essence of translation made by machine means. On the one hand, it is known that machine translation is in-demand and is improving all the time in order to search the best connection of thinking and speech, providing the best translation options. In addition different big companies and corporations which can

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

hire translators exercise machine translation systems. On the other hand, scholars are required not to use any of these systems while sharing their researches.

So, the questions appear whether machine translation for scientific and technical texts can or cannot be used, when it may be advisable where and why it should not be applied. This suggests that the topics of research on the problem of machine translation are relevant. The significance of research contribution is without doubt, there are only practical researches of this question without theoretical analysis, comparison of main principals of machine translation for scientific and technical texts.

Considering the need to solve the problems of using machine translation systems for understanding scientific and technical texts in particular and for science in general, our research, aimed at their study, analysis and comparison, assessment of the quality of these systems is relevant.

So, we offer the clear research execution protocol. After the introducing the problem of the area in Section I, the recent works in Section II are reviewed. The rest of the paper is organized as follows: Section III investigates the explored materials and methods in the experiment of machine translation systems. Section IV goes through classification for machine translation systems including neural machine translation system, describes and compares translation quality of modern machine translation systems in practice namely Google Translate and Pragma Online. In Section V the discussion is highlighted. Section VI the conclusions, challenges, and perspectives are concluded.

2. RECENT WORK

As several sciences such as Computer Science, Programming, Linguistics investigate the benefits of information technology for educating people there are some papers about using technologies. They are for language learning, reading comprehension, online training platform by C. Lai, M. Shum, Y. Tian [1]; online reading by V. Zenotz [2]; online strategy instruction for integrating dictionary skills by J. Ranalli [3]; autonomous use of computer assisted language learning (CALL) by K. Smith, H. Craig [4]; translating constructivism into English as foreing language (EFL) classrooms through the use of technology by S. Lee [5] etc.

As for defining some features of modern machine translation it was done by J. Dean [6], G. Corrado [7], J. Markoff [8]. However, unfortunately, in their works, as in other studies [9], there are no comparisons of the modern translation systems with previous ones.

It means that the dynamics of the development of translation systems has not been determined. From a practical point of view, this may cause difficulties related to the analysis of the quality of translation of different systems. In order to overcome this problem, a study of previous models of different systems is conducted by J. Hutchins [10].

Some other works are related to sentence-based topic models by D. Wang, S. Zhu, T. Li, Y. Gong [11]; structuring texts and bringing order into texts by R. Mihalcea, P. Tarau [12]; correlation of NLP and machine learning by C. Aone, M. E. Okurowski, J. Gorlinsky [13] and others.

However, they are not related to modern machine translation services. Without denying the significance of the results of predecessors, we note that their works do not adequately consider the advantages and disadvantages of each type of machine translation services.

Obviously, this is due to the complexity of determining the assessment of the quality of machine translation. There are some works [14], [15] that show an analysis and comparison of machine translation systems into their native language.

Though, such versions of machine translation services are already out of date today. Our previous works [16], [17], [18] provided recommendations on the implementation of machine translation systems in the professional training of translators, linguists, and philologists but did not consider peculiarities of modern neural systems.

The problem of systematization of previous and existing machine translation systems, analysis of their models, comparison of modern translation services, and the quality of their translation in the scientific and technical field as well as solving of the mentioned contradiction have not been the subject of a separate study yet.

Therefore, there are reasons to believe that the lack of certainty of such an issue necessitates the research in this direction. So, the conducted research purpose is to investigate how the technology of machine translation can be used by scholars, scientists, users in order to understand scientific and technical texts.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

To achieve this purpose, the following objectives were solved: to compare and highlight the basic advantages and disadvantages of different machine translation systems; to compare, analyze, and check experimentally the quality of most effective machine translation systems in practice, to discuss some open questions for further work.

3. MATERIAL AND METHODS

3. 1. Explored materials used in the experiment

To conduct the experimental research as for supplies Google Translate and Pragma Online systems were used. An experimental sample of the text for translation was taken as an article abstract, numbering 1142 characters [19]. Additional experiment was necessary to study the most general problems of translation of two main levels of language, included on lexical level (as basic) word combinations with the word "SYSTEM", on syntactic one (as advanced) a sample of complex sentences. They were also taken from the articles of technical sites.

3. 2. Explored methods used in the experiment

To archive the purpose of the research and to resolve the mentioned tasks in the theoretical part we have used the following cognitive-generalizing methods: analysis, synthesis, comparison, and classification in order to examine and compare different machine translation systems, determine advantages and disadvantages of each of them, develop classification, decide features of neural machine translation systems, develop own criteria for assessing the quality of machine translation of scientific and technical texts.

The basic indicators of the practical study of the analyzed translations of Google Translate and Pragma Online are: adequacy, the use of correct word equivalents, the accuracy of the translation of terminology, grammatical compliance.

An experimental study of samples was carried out using the method of mathematical statistics Student's criterion, the feasibility of which in our study is confirmed by the necessity of substantiation of the obtained results.

4. RESULTS

To assess the quality of machine translation and to be able to determine the possibility and advisability of using machines for translation of scientific and technical texts it is necessary to consider that today there are different machine translation systems, which use different principles of work and have their advantages and disadvantages.

4. 1. Classification for machine translation systems

There are some classifications, for example of well-known J. Hutchins' structure [10] for machine translation systems.

In our study we rely on the approaches that have become the basis for the creation of various machine translation systems; they are mostly divided into three groups:

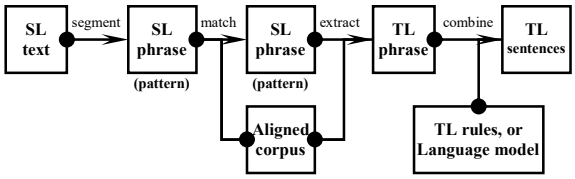
- rule-based systems, supplemented by transfer, interlingua, and the use of samples;
- statistic systems;
- hybrid systems.

Let's compare and analyze each of the systems in details.

1. Rule-based systems, supplemented by transfer, interlingua, and the use of samples

First, rule-based machine translation systems mean the creation and use of rules of linguistics. They are presented in Fig. 1, where ST is Source Text, TL is Target Language (Fig. 1).

Figure 1 : Rule-Based Machine Translation System.



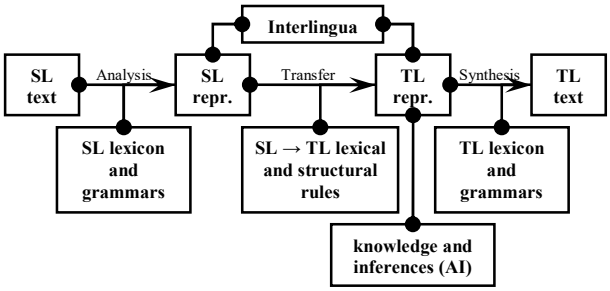
No doubt, the more precise the rules for the processing of language pairs (at different linguistic levels) are defined and the better the bilingual dictionaries are composed, the more efficient the translation made by such a machine translation system will be.

However, the task of creating such dictionaries and defining the rules takes a lot of time, which is a disadvantage of this system. In addition, the difficulties related to the exceptions in grammar lead to the need to develop new algorithms and to improve existing ones [14]. The advantage of this machine translation system is that it does not need the bases of parallel texts.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

In the process of development the machine translation systems were complemented by transfer and interlingua. They are demonstrated in Fig. 2, where AI is Artificial Intelligence (Fig. 2).

Figure 2 : Transfer and Interlingua Machine Translation Systems.



Then example-based translation was added. Its feature is the use of the bilingual body of texts as the main knowledge base. The mentioned was intended for translation by analogy, i.e. when the phrases are translated according to already existing translations, empirically.

Consequently, this system block of memory, where often repeated words, phrases or sentences are stored, is the basis of such translation. During the translation, the most similar sentences are translated and then combined. These systems are presented in Figure 3.

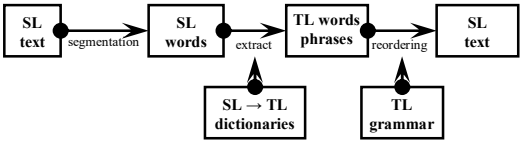


Figure 3 : Example-Based Machine Translation System.

2. Statistic machine translation systems

Second, statistic machine translation is a kind of machine translation, which is based on statistical models in accordance with the derivative analysis of the corpus of source and target texts (Fig. 4).

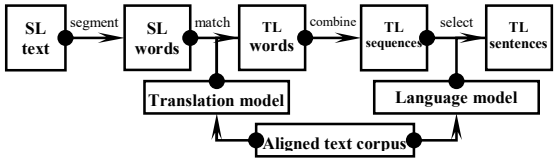


Figure 4: Statistic machine translation system.

The basic feature of statistical machine translation is the ability to "learning". That is, the system has access to the bases of texts presented in the source and target languages (parallel corps created by professional translators and linguists).

The principle of this machine translation system is as follows: the program analyzes a large body of texts in the source and target languages to ensure that the lexical-grammatical form of the original language and the translation language is adequate. When translating, the system is looking for examples in hundreds of millions of online documents to provide the best option done and preserved by translation professionals. Having found such examples in texts translated by people, the system can make intelligent assumptions about the correctness of the translation.

As we remember the model based on the Bayesian theorem is the basis of statistic machine translation. According to this theorem, the translation sentence is the most likely translation of the original sentence:

$$\max_{e} \mathrm{P}(e|f) = \max P(e) * P(f|e),$$

where e is a translation sentence, and f is an original sentence.

Thus, the statistic type of translation is based on statistical methods [20], [21], [22], which have been used since 2007 within their own closed systems of machine translation [23], [24], [25]. The use of words in different languages, which are characterized by the highest frequency, gives the opportunity to improve the translation, relying on the very frequency when translating the document. Its advantage when using, for example, on-line services is that it is looking for examples in large volumes of text to provide a better translation option.

3. Hybrid systems of machine translation

Third, based on the name of the translation system, it is natural to assume that it combines previously created translation systems that have already been described in the paper. These machine translation systems, on the one hand, have the accuracy of systems based on the rules, and on the other hand, they are able to learn, because they use statistical methods. The latter constantly is improving the quality of the translation.

Consequently, the availability of different systems of machine translation allows the user to

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

choose the best according to personal needs, purpose, conditions, sphere of translation.

Nowadays, the testing of so-called "neural machine translation" is going on. There are some papers about neural networks and automatic language identification [26], [27], [28], [29], [30], [31].

Neural machine translation is formed on the use of artificial neural network [32]; [33]; [34]. Neural machine translation allows users to translate more accurately, quickly, precisely and simpler.

The presentation of the system of neural synchronous machine translation was carried out in 2016 by scientists from three universities: Hong Kong, New York and Carnegie Mellon University. It testified that this translation system provides a fairly high quality translation without time delay.

This should be defined as one of the features of this system, because, unlike the traditional translation, where sentences are divided into separate words, phrases, parts of the sentence, which are translated sequentially. It sometimes causes unnatural and obscure translations, the technology of the neural machine translation can encoding the semantics of the whole sentence [35].

Consequently, thanks to the system of neural machine translation, based on the technology of neural networks, sentences will be translated as a whole: one of the networks analyzes the sentence in the language of the introduction, and the other one generates it in the desired language [36]. Thus, the quality of the neural translation will be closer to the usual language, which will allow the sentence to sound naturally.

The next essential difference of the neural machine translation is based on a self-learning system that learns on millions of examples, which has a set of algorithms behind the foundations. By not only studying millions of examples, but also taking into account context and associative relationships, the system improves the quality of translation. Then the result analyzed adapts to the human language according to its grammar.

The structural design of the system training of neural machine translation is being tested for 59 languages. Another feature of this machine translation system is the translation of the rarely used languages. The system translates from the original language to another directly, for example, from the Japanese to Korean, without the basic English language. Previously, the mentioned systems translated from the original into English, and then from English into the language of translation.

4.2. Ttranslation quality of modern machine translation systems

Today there are several companies working in the field of the development of these systems of all kinds for the translation of different languages of the world. Useful online services include: Google Translate, Pragma, PROMT, SYSTRANet, SDL FreeTranslation, Babel Fish, World Lingo, InterTran, Im Translator, Translate Oline.ua, Windows Live Translator.

These are offline translation services such as Google Translate Desktop, Client for Google Translate, Abbyy Lingvo, QTranslate, Dicter, Lingoes.

Definitely, there are a number of typical mistakes that arise during the machine translation. As a rule, these are errors in vocabulary and grammar. Mistakes in vocabulary, above all, include the incorrect correspondences, absence of translation, inadequate terminology. Mistakes in grammar include types of communication in the sentence, the order of the members of the sentence, the recognition of different grammar forms.

Our research relates to the quality of translation for its understanding scientific and technical texts by the main user – scholar. Therefore, the following criteria for assessing the quality of translation have been defined: the adequacy of the translation as a whole (understanding the content of the text); the use of adequate equivalents of words (words with many meanings and complex ones); accuracy of the translation of scientific and technical terminology; grammatical correspondence to the language of translation, namely: the coordination of the endings, the order of words.

Two of the most popular modern translation services, namely Google Translate and Pragma Online have been analyzed in the experiment. In the study, we did not apply additional translation functions for equal conditions of the experiment.

It is known, Google Translate is the most popular free online machine translation service that automatically translates into different languages of the world. The Google Translate service is capable of self-learning. When translating, Google Translate analyzes texts edited by translators. On the one hand, it ensures translation quality, on the other, the translation is being improved.

Another popular online machine translation services is Pragma Online of Trident Software. Its feature is believed to be special attention to the linguistic part. The latest

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

development is Pragma 6.x for 56 directions of translation.

To analyze the translation (from English into Ukrainian), the article abstract, which has 1142 characters, was used [19].

The text of the translation by Google Translate is given in Table 1.

Table 1 : Translation by Google Translate

ORIGINAL TEXT	TRANSLATION BY GOOGLE TRANSLATOR
«A method for the nonlinear electric circuit analysis with the use of the orthogonal components of instantaneous admittance and resistance is proposed. The mechanism of the generation of the admittance orthogonal components of a nonlinear electric circuit is demonstrated. The analytical expressions for the components of instantaneous admittance and resistance are given and analyzed. The balance equations for the components of instantaneous admittance and resistance are analyzed. On their basis, the harmonic components of the current of the nonlinear electric circuit consisting of an active linear and nonlinear resistance, connected in series, are determined. The accuracy and adequacy of the developed method are proved by the comparative analysis of the current harmonic components calculated with the use of the proposed method and the ones obtained as a result of the numerical calculation of the researched circuit mathematical model. The advantages of the proposed method include its versatility in the use, the possibility to assess the circuit parameters influence on the current spectrum composition, good adaptation to the automation of the calculations in the frequency domain, the ability to obtain the predicted result independently of the degree of the approximating polynomial and the number of the analyzed harmonics.»	«Запропоновано метод аналізу нелінійної електричної схеми з використанням ортогональних компонентів миттєвого допуску та опору. Показано механізм генерації допустимих ортогональних компонентів нелінійної електричної схеми. Дано та проаналізовано аналітичні вирази для компонентів миттєвого входу та опору. Проаналізовано рівняння балансу компонентів миттєвого допуску та опору. На їх основі визначаються гармонічні складові струму нелінійної електричної ланцюга, що складаються з активного лінійного та нелінійного опору, з'єднаних послідовно. Точність та адекватність розробленого методу підтверджені порівняльним аналізом поточних компонентів гармоніки, розрахованих з використанням запропонованого методу, і результатів, отриманих в результаті чисельного розрахунку математичної моделі досліджуваної схеми. Переваги запропонованого способу включають його універсальність у використанні, можливість оцінити вплив параметрів ланцюга на склад струму спектру, гарну адаптацію до автоматизації розрахунків у частотній області, можливість отримання прогнозованого результату незалежно від ступінь апроксимуючого многочлена та число аналізованих гармонік.»

The text of the translation by Pragma Online is given in Table 2.

Table 2 : Translation by Pragma Online

ORIGINAL TEXT	TRANSLATION BY PRAGMA 6
«A method for the nonlinear electric circuit analysis with the use of the orthogonal components of instantaneous admittance and resistance is proposed. The mechanism of the generation of the admittance orthogonal components of a nonlinear electric circuit is demonstrated. The analytical expressions for the components of instantaneous admittance and resistance are given and analyzed. The balance equations for the components of instantaneous admittance and resistance are analyzed. On their basis, the harmonic components of the current of the nonlinear electric circuit consisting of an active linear and nonlinear resistance, connected in series, are determined. The accuracy and adequacy of the developed method are proved by the comparative analysis of the current harmonic components calculated with the use of the proposed method and the ones obtained as a result of the numerical calculation of the researched circuit mathematical model. The advantages of the proposed method include its versatility in the use, the possibility to assess the circuit parameters influence on the current spectrum composition, good adaptation to the automation of the calculations in the frequency domain, the ability to obtain the predicted result independently of the degree of the approximating polynomial and the number of the analyzed harmonics.»	«Метод для нелінійного електричного аналізу схеми із застосуванням ортогональних вузлів миттєвого доступу і опору запропонований. Механізм покоління ортогональних вузлів доступу нелінійного електричного кругообігу продемонстрований. Аналітичні вирази для вузлів миттєвого доступу і опору дані і розкладені. Balance вирівнювання для вузлів миттєвого доступу і опору розкладені. На їх базисі, гармонійні складові потоку нелінійного електричного кругообігу, що складається з активного лінійного і нелінійного опору, з'єднався в рядах, визначені. Точність і адекватність методу, що розвивається, доводить порівняльний аналіз поточних гармонійних складових вичислив із застосуванням запропонованого методу і ті отримали в результаті чисельного обчислення дослідженої схемної математичної моделі. Переваги запропонованого методу включають його різносторонність у використання, можливість оцінити схемні параметри впливають на поточну spectrum композицію, хороша адаптація до автоматизації обчислень в частотній області, здатність отримати передбачений результат незалежно від міри многочлена, що наближається, і числа розкладеного harmonics.»

Since this research is not linguistic, we, as stated above, did not intend to carry out a linguistic analysis of the text of the submitted translations, instead, the material was analyzed on the basis of certain criteria.

Thus, during the analysis it was found that: both translations are generally clear, although the version proposed by Google Translate is more adequate for the Ukrainian language, however, it is necessary for researchers to agree on and clarify certain terms; Google Translate is better than Pragma Online in translating lexical matches, which means the service has a better ability of finding a more precise match on the system and therefore offers a more affordable translation option; although Google Translate, unlike Pragma Online, has a wider base of terms and in most cases it is more likely to pick them up, however, both systems do not always correctly translate terms. Moreover, Pragma Online leaves untranslatable words that have no match in its system; in both translation systems there were problems with the construction of grammatically correct sentences.

However, Google Translate has better overcome this challenge by making errors only in matching endings, Pragma Online translates less correctly from the point of view of grammar, and also the shortcoming of Pragma Online are the types of connection in phrases. The correct word

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

order in complex sentences is one of the challenges for both systems, although the Google Translate service copes with the translation of complex sentences better.

To evaluate two translation services Google Translate and Pragma Online, and to confirm our analysis, samples of two levels of language: the basic and the advanced, were performed. In the first instance it was lexical level [37], in the second one the syntactic level [38] (as they contain the most number of errors in translation). The samples were taken from the technical journals.

It is necessary to mention that in research on applied linguistics, the typical task is to compare different samples (two or more variations) to determine whether there is a difference between them. Most often, for this purpose, Student's criterion is used, which is to compare the arithmetic mean of both sets. The standard error t defines the probability of the difference between the arithmetic mean.

Student's criterion helps to determine the significance or the insignificance of the difference in arithmetic mean frequency in the two samples, it is sufficient to know the absolute and average frequencies of the unit, which is investigated in both samples. Consequently, the criterion of Student's (t) is found by the formula:

$$t = \frac{\left|\bar{x} - \bar{y}\right|}{\sqrt{\frac{\sum (x_i - \bar{x})^2 n_i + \sum (y_i - \bar{y})^2 n_i}{n+m-2} \cdot \frac{n+m}{n \cdot m}}}, \quad (1)$$

where $\bar{x}$ is average frequency of the studied unit in the first sample; $\bar{y}$ is average frequency of the same phenomena in the second sample; x_i is absolute frequencies in the first sample; y_i is absolute frequencies in the second sample; n is the number of sub-samples in the first sample; m is the number of sub-samples in the second sample.

Consider Example 1.

Example 1. For the study we selected the word combinations with the word "*SYSTEM*", because they are the most frequent ones on the technical sites. The result of the translation by two machine translation systems Google Translate and Pragma Online is shown in the Tables 3 and 4 (sub-sample 500 word combinations).

Table 3 : The frequency of the word combinations with the word SYSTEM in the first sample

x_i	n_i	$x_i n_i$	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 n_i$
0	18	0	-1,57954545	2,49496384	44,9093492
1	29	29	-0,57954545	0,33587293	9,74031508
2	22	44	0,420454545	0,17678202	3,88920455
3	13	39	1,420454545	2,01769112	26,2299845
4	4	16	2,420454545	5,85860021	23,4344008
5	1	5	3,420454545	11,6995093	11,6995093
6	1	6	4,420454545	19,5404184	19,5404184
Σ	88	139			139,443182

Table 4 : The frequency of the word combinations with the word SYSTEM in the second sample

y_i	m_i	$y_i m_i$	$(y_i - \bar{y})$	$(y_i - \bar{y})^2$	$(y_i - \bar{y})^2 m_i$
0	20	0	-1,47297297	2,16964938	43,3929876
1	22	22	-0,47297297	0,22370343	4,92147553
2	17	34	0,527027027	0,27775749	4,72187728
3	10	зо	1,527027027	2,33181154	23,3181154
4	3	12	2,527027027	6,3858656	19,1575968
5	1	5	3,527027027	12,4399196	12,4399196
6	1	6	4,527027027	20,4939737	20,4939737
Σ	74	109			128,445946

So, in the first sample $\bar{x}$ =1.58, in the second $\bar{y}$ =1,47

$$t = \frac{\left|1{,}58 - 1{,}47\right|}{\sqrt{\frac{139{,}44 + 128{,}45}{88+74-2} \cdot \frac{88+74}{88 \cdot 74}}} \approx 0{,}54 \,. \quad (2)$$

According to the table of critical values of t, we evaluate its indicator as for the significant difference. For this purpose we need to determine the number of degrees of freedom. In our study $f = 88 + 74 - 2 = 160$. Since in our study t is approximately equal to 0.54, which is less than the smallest number in the series, we will accept the null hypothesis.

The mentioned means that we accept the null hypothesis, that is, the difference in the frequency in translation of the word combinations with the word SYSTEM in both of our samples is insignificant, it is statistically valid, and is related to the usual frequency fluctuations that are calculated. So, we can conclude that both systems translate these word combinations equally.

In another example, we investigated the translation of complex sentences by both of the mentioned translation systems. We chose complex sentences from the same sites.

It is necessary to mention that in research on applied linguistics the typical task is to compare different samples (two or more variations) to determine whether there is a difference between them. We investigated the translation of complex sentences by both of the mentioned translation systems.

To compare the arithmetic mean of the two sets, we use Student's criterion. We determine the probability of the difference between them, using the standard error t according to the formula:

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$t=\frac{d}{S_d}, (3)$$

where d is the difference between the arithmetic mean of the two sets, S_d is the average error of this difference. S_d is calculated by the formula:

$$S_d=\sqrt{\frac{\Sigma\left(x_1-\bar{x}\right)+\Sigma\left(y_1-\bar{y}\right)^2}{\left(n_1-1\right)\left(n_2-1\right)}\times\frac{\left(n_1+n_2\right)}{\left(n_1\times n_2\right)}}, (4)$$

where Σ is the sum, x_1 and y_1 are absolute frequencies of both samples, $\bar{x}$ and $\bar{y}$ are the arithmetic mean of the first and second sets, n_1 and n_2 are the number of sub-samples in the samples.

Consider Example 2.

Example 2. The results of the study of complex sentences from the sites made by machine translation systems Google Translate and Pragma Online are presented in Table 5.

Table 5 : Calculation of Data for Determination S_d

x_1	$x_1-\bar{x}$	$(x_1-\bar{x})^2$	y_1	$y_1-\bar{y}$	$(y_1-\bar{y})^2$
29	-1,8	3.24	20	-1	1
25	-5,8	33,64	16	-5	25
38	7,2	51,84	27	6	36
30	0,8	0,64	21	0	0
32	1,2	1,44	18	-3	9
			24	3	9
		90,8			80

In the first case numbers *29, 25, 38, 30, 32* are obtained; in the second – *20, 16, 27, 21, 18, 24*. The arithmetic mean for the first set is $\bar{x}$ = *30,8*; the arithmetic mean for the second set is $\bar{y}$ = *21*; number of first sub-sample is n_1 = *5*; number of second sub-sample is n_2 = *6*.

Thus,

$$S_d=\sqrt{\frac{90,8+80}{(5-1)\cdot(6-1)}\times\frac{5+6}{5\times6}}=\sqrt{\frac{171}{20}\times\frac{11}{30}}=\sqrt{3.14}=1,77, (5)$$

$$t=\frac{30,8-21}{1,77}=5,54\,(6).$$

To determine the significance of *t*, we need to know the value, which is called the degree of freedom, and which is calculated using the formula: $d_f = (n_1 – 1) + (n_2 – 1)$.

Thus, $d_f = (5 – 1) + (6 – 1) = 4 + 5 = 9$. According to the table of critical values of *t* we find that when $d_f = 9$ and the degree of significance $P = 0,05$, *t* must be *2,26*, and when $P = 0,01$, $t = 3,25$. According to our calculations $t = 5,54$, this is more than *3,25* and *2,26*. This indicates that there is a significant difference between the two studied sets. So, we can conclude that both systems translate these complex sentences differently.

5. DISCUSSION

In determining the need for the rapid implementation of the best scientific ideas and technology transfer, the demand for understanding and translation of scientific and technical texts is increasing. The implementation of machine translation systems facilitates overcoming the language barrier [1], [2], [3], [4], [5], [39]. It should be noted that machine translation systems experienced significant dynamics during development from simple systems of direct machine translation to complex systems of neural machine translation.

Obviously, this way has allowed improving the quality of translation. There are many criteria that classify machine translation systems [10]. According to the approach on which the machine translation relies, we propose a classification of the systems, namely: traditional, statistical [25], hybrid, and promising neural [26], [27], [28], [29], [30], [31], [40].

All of them have their own advantages and disadvantages. Accordingly, the development of machine translation systems contributed to the creation of many modern machine translation services. In this sense today, effective services for translating texts are Google Translate and Pragma Online, so we have taken them for analysis. The analysis of translations made it possible to conclude that both systems of machine translation give an understanding of the essence of a scientific and technical text.

This indicates the expediency of their use for the general understanding of the text. The latter opens up the opportunity for the effective use of both systems for this purpose. However, the mentioned analysis of translations has also shown that the noted translation services make errors on different levels of the target language.

So, in spite of readability and sufficient clarity of translation made by modern machine translation services they can be used only for getting general information and not for sharing scholar's research as when using reverse translation (by the way, known as one of the methods for assessing the quality of translation) scholars all over the world will get misrepresented, inadequate ideas as mistakes made during the first machine translation will increase after the attempts of reverse translation.

When analyzing the quality of translation made by Google Translate and Pragma Online,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

from a theoretical point of view, it seams that both systems are equally useful for translation; and this is a certain advantage for users.

However, the practice shows the difference between the usages of both translation services. To prove this, we analyzed the quality of translation of both systems according to certain criteria. The results confirmed the practical data from another experimental works [14], [15]. Nevertheless, unlike the other published experimental results [15], [37], [40] the data we obtained about the assessing the quality of the translation, suggest the following.

Despite its ease of use, Google Translate is not only a very powerful translation tool that is capable of performing translation well enough in the general level and is very understandable to the user, but also showed better translation options in terms of the determined criteria, therefore, from our point of view, it may be better in translating scientific and technical texts for general information. The experimental applied research on various samples and according to Student's criterion was made; the different in both systems of translation have been proved.

Such conclusions can be considered useful from a practical point of view, as it is possible to advise reasonably to use the Google Translate service for translation with the purpose of getting general information.

6. CONCLUSIONS

The research has made it possible to establish the peculiarities of different systems of machine translation on the basis of the proposed classification (traditional systems based on rules supplemented by transfer, interlingua, on the basis of samples; statistic systems based on finding the most probable translation, comparing bilingual sets of texts; hybrid systems combining direct translation with the addition of statistic methods), it gave an opportunity to trace the dynamics of their development from the past to the present, to highlight the advantages and disadvantages of each system.

Features of a promising neural machine translation are that the system translates the sentences entirely, using millions of examples, rather than dividing them into separate words and phrases. The system will learn on its own and at all times will encode the semantics of the sentence, taking into account the context and associative relationships. This will manifest itself in the creation of better and more natural translations, so we hope to see the path from testing to implementation.

Comparison and analysis of the quality of translation of the selected scientific and technical text by two modern services of machine translation Google Translate and Pragma Online on the basis of certain criteria allowed to recognize that Google Translate is more correct in selecting lexical matches, translates technical terms more successfully, transmits links in phrases better. Due to such a comparison, it can be argued that it is advisable to use machine translation systems for getting general information.

The research was carried out in the samples (according to the word combinations in the first example, according to the complex sentences in the second example), due to Student's criterion it is proved that the systems translate complex sentences in different ways. Thanks to this comparison it is found that Google Translate translates complex sentences more accurately.

Consequently, based on the results of comparative analysis, experimental research and the proof of its results by methods of mathematical statistics, we consider using the Google Translate service for the general understanding of the texts of the scientific and technical sphere to be more advisable.

Therefore, there are some limitations for the study as machine translation systems can be used for getting overall content of a scientific and technical text but it cannot be applied for presenting academic research.

To evaluate the research outcomes, it is necessary to say, that the paper analyses and compares different machine translation systems, its advantages and disadvantages, scientific and technical text examples translated with Google Translate and Pragma Online, and especially it focuses on recent efforts and advances. Moreover, we present quick discussion on text translation quality evaluation.

On the one hand, while, there are many researchers focusing on improving translation quality, there are also other researchers working toward the improvement neural machine system. On the other hand, understanding scientific and technical texts is still representing a challenging task and needs more updates as due to the variety of modern machine translation systems, it's required to find the best evaluation methods that work effectively with any system; but, still we don't know and it's a great challenge for mankind whether machine translation system can be absolutely sufficient.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

There are some open issues for further work in this area. The results of the experiment proved that analyzed machine translation systems make errors on different levels of the target language. Although the target text is readable, it is necessary to warn scholars against applying these systems for sharing their research, as when doing reverse translation foreign academic circles will get misrepresented information.

The rapid updating of machine translation systems allows asserting that the results obtained at a certain time can change. The inability to predict the latest developments in the neural machine translation within the framework of this study generates a potentially new and interesting direction for further research.

They can, in particular, be focused on neural machine translation, for example, from Google service. Such intelligence will allow exploring the features of the influence of neural networks on the translation system itself, to identify its advantages and disadvantages, to perform translation analysis according to certain criteria.

REFRENCES:

[1] C. Lai, M. Shum, and Y. Tian, "Enhancing learners' self-directed use of technology for language learning: the effectiveness of an online training platform," *Computer Assisted Language Learning*, Vol. 29, No. 1, 2016, pp. 40-60.

[2] V. Zenotz, "Awareness development for online reading," *Language Awareness*, Vol. 21, 2012, pp. 85-100.

[3] J. Ranalli, "Online strategy instruction for integrating dictionary skills and language awareness," *Language Learning & Technology*, Vol. 17, No. 2, 2013, pp. 75-99.

[4] K. Smith, H. Craig, "Enhancing the autonomous use of CALL: A new curriculum model in EFL," *CALICO Journal*, Vol. 30, No. 2, 2013, pp. 252-278.

[5] S. Lee, "Translating constructivism into EFL classrooms through the use of technology," *Multimedia-Assisted Language Learning*, Vol. 8, No. 1, 2005, pp. 220-249.

[6] J. Dean, A. Ng, "*Using large-scale brain simulations for machine learning and A.I.,*" 2012. URL: https://googleblog.blogspot.com/2012/06/using-large-scale-brain-simulations-for.html

[27 G. Corrado, "Google-RE.WORK Deep Learning", *Summit Google's Large Scale Deep Neural Networks Project*, 2015. URL: https://www.youtube.com/watch?v=KELYHjq9Gbs

[8] J. Markoff, "How Many Computers to Identify a Cat? 16,000," *New York Times,* 2012. URL: https://www.nytimes.com/2012/06/26/technology/in-a-big-network-of-computers-evidence-of-machine-learning.html?pagewanted=all

[9] Ch. McDonald, "*Commenting on Gil Fewster's January 5th article in the Atlantic,*" 2017. URL: https://medium.com/@chrismcdonald_94568/ok-slow-down-516f93f83ac8

[10] J. Hutchins, "*Machine Translation systems and applications: past and present,"* Imperial College London, 2011. URL: http://www.hutchinsweb.me.uk/ImpColl-2011-ppt.pdf

[11] D. Wang, S. Zhu, T. Li, and Y. Gong, "Multidocument summarization using sentence-based topic models," *Proceedings of the ACLIJCNLP 2009 Conference Short Papers,* 2009, pp. 297-300.

[12] R. Mihalcea, P. Tarau, "Textrank: Bringing order into text," *Proceedings of the 2004 conference on empirical methods in natural language processing*, 2004, pp. 404-411.

[13] C. Aone, M.E. Okurowski, and J. Gorlinsky, "Trainable, scalable summarization using robust NLP and machine learning," *Journal of Theoretical and Applied Information. Technol Proceedings of the 17th international conference on Computational linguistics*,Vol 1, 1998, pp. 62-66.

[14] A. L. Mischenko, "Machine translation in the context of modern scientific and technical translation," *Bulletin of V. N. Karazin Kharkiv National University. Roman ad German Philology. Methodology of Teaching Foreign Languages*, Vol. 10, Issue 51, 2013, pp. 172–180.

[15] S. V. Palii, "The method of incorporating the machine translation system into the information and organizational environment for foreigners' training," *Managing the Development of Complex Systems*, Vol. 15, 2013, pp. 143-151. URL: http://journals.uran.ua/urss/article/view/40090/36213 (in Ukrainian)

[16] I. Kostikova, "Language games in teaching English," *Theory and Methods of Learning and Upbringing*, Vol. 43, 2017, pp.174-184. URL: http://doi.org/10.5281/zenodo.1243599

[17] I. I. Kostikova, "Information and communication technologies in students' language learning," *International Journal of*

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Education and Science, Vol. 1, Issue 1-2, 2018, pp. 7-14. URL: DOI: 10.26697/ijes.2018.1-2.01

[18] L. Holubhycha, O.Shatrava, "Using interactive whiteboard at English lessons for Law Students," *Science and Education*, Vol. 12, pp. 24-30, 2017. URL: http://scienceandeducation.pdpu.edu.ua/doc/2017/12_2017/3.pdf

[19] A. S. Al-Mashakbeh, "Development of the analitical method for nonlinear circuits analysis with the use of the admittance and resistance orthogonal components," *Eastern-European Journal of Enterprise Technologies*, Vol.1/8, Issue 91, 2018, URL: https://doi.org/10.15587/1729-4061.2018.121874

[20] K. Quach, "*Google's neural network learns to translate languages it hasn't been trained on: First time machine translation has used true transfer learning",* 2016, URL: https://www.theregister.co.uk/2016/11/17/googles_neural_net_translates_languages_not_trained_on

[21] G. Lewis-Kraus, "The Great A.I. Awakening", *The New York Times*, 2016, URL: https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html

[22] Q. Le, M. Schuster, "A Neural Network for Machine Translation, at Production Scale", *Google Research Blog*, 2016, URL: https://ai.googleblog.com/2016/09/a-neural-network-for-machine.html

[23] "*Google Switches to its Own Translation System"*, 2007. URL: http://googlesystem.blogspot.com/2007/10/google-translate-switches-to-googles.html

[24] B. Schwartz, "*Google Translate Drops SYSTRAN for Home-Brewed Translation"*, 2007. URL: https://searchengineland.com/google-translate-drops-systran-for-home-brewed-translation-12502

[25] K. Li, T. Edwards, "Statistical models for automatic language identification," *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'80*, 1980, pp. 884-887.

[26] J. Braun, H. Levkowitz, "Automatic language identification with perceptually guided training and recurrent neural networks," *International Conference of Spoken Language Process*, Sydney, Australia, 1998, pp. 289-292.

[27] I. Lopez-Moreno, J. Gonzalez-Dominguez, O. Plchot, D. Martinez, J. Gonzalez-Rodriguez, and P. Moreno, "Automatic language identification using deep neural networks," *Acoustics, Speech and Signal Processing (ICASSP)*, 2014 IEEE International Conference, 2014, pp. 5337-5341.

[28] J. Gonzalez-Dominguez, I. Lopez-Moreno, P.J. Moreno, J. Gonzalez-Rodriguez, "Frame-by-frame language identification in short utterances using deep neural networks," *Neural Networks*, Vol. 64, 2015, pp. 49-58.

[29] A. Lozano-Diez, R. Zazo Candil, J. González Domínguez, D.T. Toledano, and J. Gonzalez-Rodriguez, "An end-to-end approach to language identification in short utterances using convolutional neural networks," *Proceedings of the Annual Conference of the International Speech Communication Association*, INTERSPEECH, 2015, pp. 403-407.

[30] R. Zazo, A. Lozano-Diez, J. Gonzalez-Dominguez, D.T. Toledano, J. Gonzalez-Rodriguez, "Language identification in short utterances using long short-term memory (LSTM) recurrent neural networks," *PloS one*, Vol. 11, 2016, p. 6014-6017.

[31] W. Geng, W. Wang, Y. Zhao, X. Cai, B. Xu, "End-to-End Language Identification Using Attention-Based Recurrent Neural Networks", *Proceedings of the Annual Conference of the International Speech Communication Association*, INTERSPEECH, 2016, pp. 2944-2948.

[32] A. Imada, "Is artificial neural network intelligent?," *International Journal of Computing*, Vol. 10, Issue 1, 2011, pp. 66-76. URL: http://computingonline.net/computing/article/view/738

[33] H. Muhlenbein, "Artificial intelligence and neural networks the legacy of Alan Turing and John von Neumann," *International Journal of Computing*, Vol. 5, Issue 3, 2014, pp.10-20. URL: http://computingonline.net/computing/article/view/405

[34] Q. Zhao, "Learning and understanding based on neural network trees," *International Journal of Computing*, Vol. 3, Issue 1, 2014, pp. 84-92. URL: http://computingonline.net/computing/article/view/257

[35] Ch. Boitet, H. Blanchon, M. Seligman, V. Bellynck, "*MT on and for the Web,"* 2010. URL: http://www-clips.imag.fr/geta/herve.blanchon/Pdfs/NLP-KE-10.pdf

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[36] "*What is a neural network based translation?,"* 2018, URL: https://translator.microsoft.com/help/articles/neural/

[37] V.M. Zayacz, M.M. Zayacz, "Comparison methods of statistical characteristics in sample formation in Linguistics," *Bulletin of the National University 'Lviv Polytechnic': Information systems and tips*, Vol. 673, 2010, pp.296-305. URL: http://ena.lp.edu.ua/bitstream/ntb/6753/1/33.pdf (in Ukrainian)

[38] S.K. Romanyuk, "Use of statistical methods in linguistic research," *Scientific Notes of the National University Ostroh Academy. Philological Series*. Vol. 54, 2015, pp. 134–137. URL: https://eprints.oa.edu.ua/4185/1/52.pdf (in Ukrainian)

[39] K.-O. Jeong, "Reframing task performance with technology to promote positive interdependence in language learning", *Theoretical and Applied Information Technology,* Vol. 96, No 21, 2018, pp. 7303-7311.

[40] M. O. Elfahal, M. E. Mustafa, R. A. Saeed, "Automatic spoken language recognition for multilingual speech resources", *Theoretical and Applied Information Technology,* Vol. 96, No 24, 2018, pp. 8081-8096.