

14. Pavlina, S.Y. Separated by a Common Language: on the Usage of Idioms in British and American Electoral Discourse / S.Y. Pavlina // Journal of Siberian Federal University. – Humanities & Social Sciences. – 2017. – 10. – P. 1563–1572.

15. Salieva, R.N., Konopleva, N.V., Mirgalimova, L.M. Phraseological units in American and Russian governmental blogs / R.N. Salieva // Linguae Journal. – 2016. – №9 (4). – P. 58–67.

16. The Corpus of Contemporary American English. – URL: <https://www.english-corpora.org/coca/> (дата обращения: 17.09.2023). – Текст : электронный.

*А.В. Алина*

*Национальный исследовательский  
Томский политехнический университет*

### **From PROMT to ChatGPT: how memes about machines can become memes about people**

In this article, we analyze how early machine translation influenced the perception of non-manual translation in people's minds; what difficulties this technology has struggled with over the years; what trends exist in this study field now and to what extent machine translation is capable of replacing human translators.

Key words: machine translation; NLP; NMT; automatic translation program; neural networks.

Modernity has brought people the opportunity not only to simplify their lives with the help of computers, but also to communicate with their help. Moreover, communication does not necessarily have to be within the same language boundaries.

The first example of machine translation appeared in 1954 – when the Georgetown experiment was conducted in New York [8]. However, like any new technology at that time, it had taken machine translation another 60 years or so before it was commercialized and offered any significant result to the end user (it also happened due to the widespread accessibility to the Internet).

Machine translation is the automatic translation from one language to another by means of automatic programs. There are several methods of machine translation: rule-based machine translation and statistical machine translation [5, p. 403].

In the first case, linguistic information about the source and target languages is used: grammatical rules, sentence structure, phraseology, etc.

Then the program compares them according to pre-configured conversion parameters and outputs the source text according to the applied rules.

In the second case, a huge number of ready translations is analyzed in order to identify common patterns and regularities – how this or that phrase or even this or that paragraph should be translated. This processing model used to be based mostly on words, but as the model itself evolved, it began to operate on phrases and even learned to capture context to some extent.

In general, both of these principles led to one result: the rule-based statistical method, which is used in the vast majority of modern translation programs. At the dawn of the noughties, there were many of them, but the most popular, and therefore iconic, was PROMT, which predetermined the memetic nature of machine translation for the next few years in the minds of ordinary people.

It is not difficult to explain why this happened – in the vast majority of cases automatic translators were not a program in the full sense, but rather a large electronic dictionary with built-in functionality of matching words into a single sentence. Obviously, in such a case, homonyms in one language became a source for generating memes about machine translation in another.

As an example, we can cite a phrase from the instruction manual for a computer mouse, which in the original version looked like this: «Execute the installer by tray icon and insert ms windows binaries or another os with custom mouse driver support in current boot drive». However, the literal translation of this sentence into Russian is a perfect example of a «broken» translation: «Decapitate adjuster with image of salver and insert two-piece sets windows of mistress or another mouth with custom house support of car driver mouse in current motor of shoe». In this case, we have used a back-translation into English, and therefore we can compare how the meaning of the phrase is not just lost – it is absent.

A thorough comparison allows us to understand why the program translated specific words this way and not the other, but even then, the question remains how the operating system became a mouth.

Assessing the quality of such programs translations in retrospect, we can characterize it by their own translation example – «fable rejection», or, returning the original meaning – «epic fail».

Still, one cannot but recognize that machine translation, despite all its literalness, has become a very useful tool. Of course, it took a lot of human effort to fully develop them, because matching phrases and word combinations in a dictionary is a titanic job that requires constant tracking of new words, grammatical rules and distinguishing homonyms by the meaning of the words in a sentence. But at the moment, automatic translators have long ceased to be those «overminds» generating new meanings, and have become a tool – sometimes misfiring, but quite useful and complete [3, p. 310].

It is not possible to say that machine translation can replace human translation. First of all, because it is impossible to fit all variations of all known phrases, abbreviations and words into an electronic dictionary. As long as the language is alive, it is developing, and in such a scenario, dictionary development is an endless race to catch up. Secondly, it is easy to distinguish machine translation from human translation – the program does not adapt and adjust the meaning to the final language culture. Yes, for technical translation – manuals, documentation – a program may be enough. Nevertheless, artistic translation is not just about matching words, it is about adapting to the speaker of another language.

As someone with experience in non-profit amateur comic book translation, I can tell you that adapting artistic translations is the most difficult task, and machine translation cannot become better at it. My favorite example from my practice is a comic book «Y: The Last Man»: «You split more infinitives than Gene Roddenberry» [6, p. 63]. This is one of the rare examples where neither literal translation nor adaptation will produce a good result. The original implies that the person, the main character refers to in his speech, uses a to-infinitive grammatical construction called «split infinitive» [11], the most popular example of which is the phrase «to boldly go where no man has gone before» sounding in the splash screen of the Star Trek TV series written by Gene Roddenberry. Moreover, there is also a play of words «split more infinitives than», designed to reinforce and show the degree of use of these infinitives in the culture of speech of the comic book character. In Russian, there is no similar construction to the split infinitive, and the ordinary infinitive does not fully convey the grammatical peculiarity of English, forcing to literally translate and explain the context of the joke to the reader. Of course, there are not so many such tricky examples, but when they occur, no translator program will be able to cope with the adequate interpretation.

We have understood that memes about machine translation are a product of time, and are currently rarely found (if ever), but what is happening with machine translation trends now? Ironically, humans are now the subject of machine translation memes. In the last two years, neural networks have become widespread: ChatGPT, DeepL, etc [9]. Notably, their original purpose is to summarize generalized information about specific topics, pictures, and other data source. However, text-based neural networks allow not just generating content, but also retrieving it in another language. Given the very principle of neural network operation – constant development and improvement by learning from huge data flows, we can assume that machine translation of the overwhelming majority of information will become available to anyone without any problems – there are already vivid examples of saving money and time by running texts through neural network translators (such as

DeepL). Automated translation still suffers from its predecessors' maladies: inability to work with context, occasional misinterpretation of homonyms. However, for the overwhelming number of tasks, it is quite enough.

As a result, we come to a topical consequence – many companies that used to employ translators now have the opportunity to save money by using machine translation [4, p. 235–237]. Of course, highly specialized technical or scientific texts are still the prerogative of professional translators with specialized skills. However, I would like to look at a precedent created with the help of neural networks in the localization market.

On September 6, 2023, the game «Starfield» was released on the video game market. According to the developers' statements, the game contains 1.7 million words and 150 thousand lines of text. Moreover, for undeclared reasons, Russian text localization in the game was not announced, because of which many buyers were upset. Some enthusiasts have calculated the approximate cost of localizing such a game in Russian – about 20 million rubles and several months of work. In other words, the game had no official Russian localization by the developers, and there was no possibility of amateur translation due to the huge scale and cost of the work.

On the day of release, thanks to the capabilities of the video game engine, the game began to be modified, and one of the modifications was an amateur localization project based on running all the in-game text through DeepL [2]. To the users' surprise, the modification was quite playable. As expected, it inherits all the same errors and flaws of machine translation – context, pronouns, lexical constructions, semantic pauses. However, the machine translation was more than enough to fully understand what was going on, and together with the subsequent manual edits, the amateur translation through the program-translator became the only way to play the game in Russian. We are not talking about a full-fledged artistic translation, but this precedent shows very clearly that a very clever electronic dictionary is enough for most of the text.

Yet, there are trends that really make you think about the future of machine translation, or, to put it more correctly, neuro-translations. As it was said earlier, the idea of neural networks is based on self-learning and continuous improvement. A neural network learns by analyzing huge data amounts, but even in this case the completely available text will be mostly recirculated, and in the end, it will not have a great effect. In other words – without changing the approach to neural network training, the potential of this training is limited and has an endpoint.

Interestingly, such a change of approach is already being actively studied and implemented in human life. We are talking about NLP – Natural Language Processing. This is a field of study of artificial intelligence, which

deals with the research of interaction between human language and computer systems [12].

Everything we express in writing or verbally carries a huge amount of information. The topic we choose, our tone, our choice of words – all this adds up some kind of information on its own that can be interpreted to extract some meaning from it. Theoretically, we can understand and even predict human behavior using this information.

However, there is a problem – the information a person generates is a typical example of unstructured data. Unstructured data does not fit into the traditional row and column structure of relational databases and represents the vast majority of data available in the real world [1]. They are cluttered and difficult to process. Nevertheless, thanks to the advances in machine learning and the development of neural networks, this direction has been actively developed in practice. We speak not about trying to interpret a text or speech based on key words (the old-fashioned mechanical way), but about understanding the meaning of these words (the cognitive way). Modern researches allow us to identify figures of speech, such as irony, for example, or even analyze the tone of a text.

The most vivid example of NLP technology use are chatbots that were presented by several companies not so long ago. They can imitate the tone, semantics and personality of specific fictional or historical characters based on the information analyzed [4, p. 235–237]. That is no longer an example of re-interpretation but, in a way, a full-fledged information generation.

This brings us to the idea that such a technology can deepen the quality of machine translation, smooth out the shortcomings and introduce a new stage of their development – neural machine translation (NMT).

The future of neural machine translation is very promising, and its capabilities will only grow over time as artificial intelligence advances and neural networks become more complex. Most experts agree that in the not-too-distant future, translation will be a combination of NMT and humans: AI will provide scalable translation capabilities, and humans will provide creativity, critical thinking, and fine-grained interpretation [10, p. 272].

To summarize, machine translation now still does not pose a significant threat to the professional translator, nor is it likely to become one. On the contrary, in most cases it becomes a useful timesaving tool. Nevertheless, at the same time, the development of neural networks and machine translation, as well as neural machine translation on the horizon, increases the requirements to the competencies of a translation specialist, and if you do not meet them, then you may well find yourself the object of memes. However, this time it will be the category of memes where the machine, after years of being the target of ridicule, has taken away the work from a human.

## Литература

1. Ваш путеводитель по миру NLP (обработке естественного языка). – URL: <https://habr.com/ru/companies/otus/articles/705482/> (дата обращения: 20.10.23.). – Текст : электронный.
2. Для Starfield вышел русификатор всех диалогов на основе нейросети DeepL. – URL: <https://dtf.ru/u/254020-nik-jun/2079483-dlya-starfield-vyshel-rusifikator-vseh-dialogov-na-osnove-neyroseti-deepl> (дата обращения: 17.10.23.). – Текст : электронный.
3. Ладожинская, Д.А., Малышко, Р.В. Влияние машинного перевода на переводчиков/ Д.А. Ладожинская, Р.В. Малышко // Инновационные аспекты развития науки и техники. – 2021. – № 8. – С. 305–312.
4. Ландо, Т.М., Митренина, О.В., Николаев, И.С. Прикладная и компьютерная лингвистика / Т.М. Ландо, О.В. Митренина, И.С. Николаев. – Москва : Ленанд, 2017. – 310 с.
5. Утробина, А.А. Компьютерная лингвистика и машинный перевод: об истории становления / А.А. Утробина // Вестник Башкирского университета. – 2022. – № 2. – С. 401–404.
6. Brian, K.V. Y: The Last Man Vol. 7: Paper Dolls / K.V. Brian. – New York : Vertigo, 2006. – 144 p.
7. Exploring the Future of Translation in 2023. – URL: <https://www.marstranslation.com/blog/exploring-the-future-of-translation> (дата обращения: 19.10.23.). – Текст : электронный.
8. IBM archives: 701 Translator. – URL: [https://www.ibm.com/ibm/history/exhibits/701/701\\_translator.html](https://www.ibm.com/ibm/history/exhibits/701/701_translator.html) (дата обращения: 18.10.23.). – Текст : электронный.
9. Modern Technology and The Future of Language Translation. – URL: <https://www.smashingmagazine.com/2023/07/modern-technology-future-language-translation/> (дата обращения: 17.10.23.). – Текст : электронный.
10. Sin-wai, C. The Future of Translation Technology: Towards a World Without Babel / C. Sin-wai. – New York ; London : Routledge, 2017. – 316 p.
11. Split infinitive. – URL: [https://en.wikipedia.org/wiki/Split\\_infinitive](https://en.wikipedia.org/wiki/Split_infinitive) (дата обращения: 19.10.23.). – Текст : электронный.
12. The future of translations: exploring the latest translation technologies and trends. – URL: <https://www.artonetranslations.com/translation-technologies-and-the-future-of-translation/> (дата обращения: 20.10.23.). – Текст : электронный.

Науч.рук.: Сыскина А.А., к-т филол. н., доц.