The OSCAR PUBLISHING Services
American Journal of Philological Sciences

VOLUME Vol.05 Issue12 2025
PAGE NO. 200-204
DOI 10.37547/ajps/Volume05Issue12-55

The Evolution Of Machine Translation

Shamsiyeva Gulshoda
Teacher of Perfect-University, Uzbekistan

Received: 25 October 2025; **Accepted:** 16 November 2025; **Published:** 21 December 2025

Abstract: This article discusses the stages of the formation of machine translation. The concepts of automatic translation, translation memory, as well as types such as Rule-based machine translation, Example-based machine translation, Statistical-based machine translation, and Neuro-machine translation are analyzed. The scientific work done in the field of machine translation in Uzbekistan is also discussed. An analysis of the paratranslator platform created in Uzbek linguistics is also provided.

Keywords: Machine translation, translation memory, Rule-based machine translation, Example-based machine translation, Statistical-based machine translation, Neuro-machine translation. Paratranslator.uz.

Introduction: The first ideas about machine (automatic) translation belong to the English scientist Charles Babbage, who proposed the concept between 1836 and 1848. In his opinion, mechanical-electronic machines could perform coded automatic translation. According to Babbage's project, electronic machines with a memory capacity of fifty thousand words should be able to translate 100,000 words automatically. It took 100 years for these ideas of Charles Babbage to be realized. The first experiment on automatic translation was carried out by Georgetown University on January 7, 1954. The experiments were carried out using the IBM-701 computer. The experiment was successful, and for the first time in the history of computers, 60 sentences from Russian were automatically translated into English. This was the first public demonstration of a machine translation system, which aroused great interest in the media and the public.

According to the source, in the 18th century, the philosophers G. W. Leibniz and Descartes proposed the issue of encoding the relationship between sentences and words. The idea of the ability of a machine to translate was noted by Ch Babbage (1791-1871) in his project on the numerical analytical machine, a mechanical prototype of electronic digital machines that appeared 100 years later in 1836-1848. In the 1930s, “translation machines” were used in practice, while in Estonia, A. Vacher’s theoretical views on the implementation of mechanical translation were published in the newspaper “Vaba Maa” (1924). In France, J. Astrouni proposed the use of automatic bilingual dictionaries, and he received a patent for the development of a machine translation called “Mechanical Brain”. In 1933, in Russia, the author of the linguistic arithmetic, P.P. Smirnov Troyansky, created a system based on Esperanto, which included a method of distribution between languages with a grammatical control method and a bilingual dictionary.

In world linguistics, many scientists have conducted scientific research on machine translation, including N.D. Andryev, I.A. Melchuk , I.I. Revzin , V.Y. Rozensveig, Y.N. Marchuk , R.G. Piotrovsky , K.B. Bektaev , A.N. Belyaev, I.K. Belsky, A.V. Zubov, G.E. Miram, L.L. Nelyubin, V.I. Perebiynos, V.A. Chizhakovsky, Ye.A. Shingarev, G.G. Belonogov , R.G. Kotov, Babushkina N.V, Z.Shalyapina, O.Y. Mansurova, A.S. Panina, A.A. Khoroshilov (Russia); A. But, R. Richans, J. Hutchins, J. Allen, P. Brown (USA); M. Nagao (Japan); A. Vacher (Estonia); J. Astrouni in France; R. Sinha, A. Jain (India); B. Blazer, U. Schvol, A. Storrer (Germany) are among these researchers.

Related work

The first scientific and practical conference on machine translation was held at the Massachusetts Institute of Technology (1952), and the first IBM II machine translation program for translating from Russian into English was created in collaboration with Georgetown University in New York (1954). This research was

conducted under the leadership of L. Dorster and was limited to 250 lexical units and 6 grammatical rules within the framework of chemistry. ("Dorster experiment" or "Georgetown test"). The first machine translation in Turkic languages was also carried out under the leadership of Dorster (1961). The English-Turkish machine translation system consisted of 700 word forms, and the dictionary consisted of a list of bases and suffixes of their equivalents in Turkish and English.

In 1960, a special meeting was held at the Academy of Sciences of the former USSR dedicated to automatic translation. At the end of the meeting, a special laboratory for automatic translation was organized under the leadership of Professor A. Belenogov on machine translation. Collections dedicated to machine translation, such as "Mechanical Translation" (1954), "T.A. Informations" (1965), and "Communications of the Association for Computing Machinery" (1958) began to be published. In Kazakhstan, a special working group led by Professor A. Bektayev calculated the linguistic statistics of the novel "The Path of Abay".

A new phase of research on machine translation began in the 1970s. It was called "artificial intelligence" because of the modeling of certain problems on a computer. Also, during this period, TM (TM – Translation Memory), working with set theory, was created. According to it, the base segment (sentence) and the text translation are stored, forming a linguistic database. In these years, in the former USSR, translation programs such as AMPAR, ASPIR, which translate from English to Russian; NERPA, which translates from German to Russian; FRAP, RANO (Spanish-English), SYSTRAN, TAUM (English-French), POLA (Chinese-English), MIAN (French-Russian), CULT (Chinese-English), as well as STYLUS, SOCRAT, were created. Linguists such as A.I. Melchuk and Y.D. Aprisian worked on the creation of the ETAP processor.

The discussion of a report on automatic translation in the US Congressional Committee on Science in 1966 gave a negative result. After that, funding for automatic translation projects worldwide decreased sharply. Only in Japan, state funding for such research and experiments continue. In other countries, automatic translation projects are financed by large private companies. Such large organizations as "Phillips", "General Motors", "Siemens", and "LG" are among such companies, and they still invest in scientific research and experiments on automatic translation today.

In machine translation (automatic translation), texts are translated from one language (source language) to another (target language) using a computer. This process takes very little time. Another widely used type of translation system is the SYSTRAN automatic translation system, with the help of which 2.5 million pages of text are translated from English into German, French, Spanish, Greek, Italian, and from French into English, Spanish, and Italian per year.

In the USA, NASA specialists implemented the DIANA machine translation project in 1991. It was the largest automatic translation project in the world. Its database included 10 million words, based on the 6 largest languages of the world: English, Russian, German, Spanish, Arabic, and French. Russian research scientists led by Professor Lyakunov are implementing the Retrans Vista project belonging to the Vista Technology company. A total of 3.4 million words, including 1.6 million English-Russian and 1.8 million Russian-English words, are included in the translation memory of the program. Philipp Koehn, a scientist at the Massachusetts Institute of Technology, describes statistical machine translation models based on phrases using the "Pharaoh" machine translation engine.

By the 1990s, machine translation had undergone significant changes. This was due to the introduction of new technologies, the application of statistical methods, the advancement of the concept of connectionism, and the development of technology for using neural networks. The improvement of computer software in the 1990s-2000s also contributed to the creation of more advanced versions of machine translation systems. In later periods, MT systems based on corpus-based, neural network, and artificial intelligence technologies were created.

In the 1990s, this direction began to be called by various names: example-based, memory-based, analogy-based, and case-based machine translation, which is based on a database of previously translated examples or a corpus. Currently, dozens of companies around the world are engaged in the development of commercial programs: Systren (abbreviation of the phrase "System of translation"), IBM, L&H (Lernout & Houspie), Language Engineering Corporation, Transparent Language, Nova Incorporated, Trident Software, Atril, TRADOS, Caterpillar Co., Lingvo Ware, Ata Software, Lingvistica, Pragma, among others. Among such programs, it is appropriate to highlight the Retrans Vista version, which translates Russian texts into English. Its dictionary contains not only the meanings of millions of phraseological expressions, but also phrases used in everyday speech. In addition, the program also has a conceptual analysis that automatically extracts phrases from the text reflected in the dictionary. The dictionary of the Retrans Vista system consists of phraseological units and terms

related to natural sciences, technology, economics, business, and politics. The machine dictionary contains about 3.4 million. lexemes, of which 1.8 million. Russian-English, 1.6 million. English-Russian lexemes. 20% of the dictionary consists of lexemes, 80% of stable word combinations with an average "length" of 2.2 words. This program runs under the control of the OS Windows 98/NT/2000. Now, you can also use special sites for automatic translation via the Internet: www.alphaworks./hm/aw.nsf/html/mt, www.freetranslation.com, www.t-mail.com, www.translate.ru, www.foreiegnword.com/Tools/translate.dyn, etc.

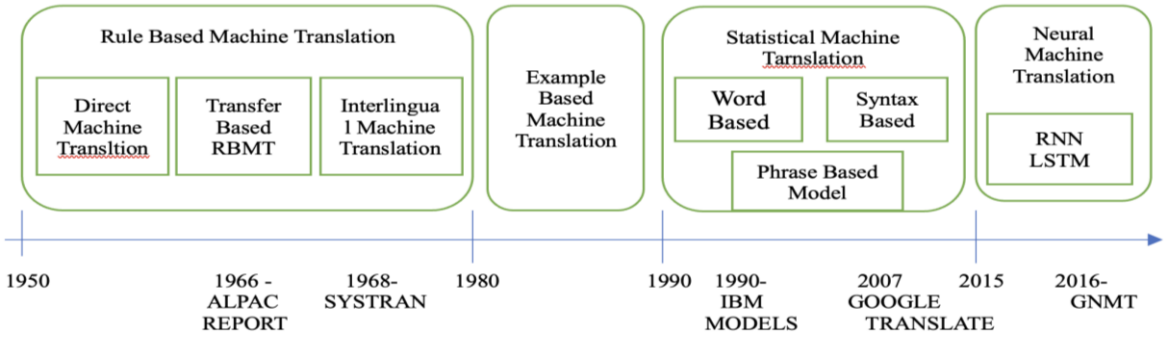
Since machine translation is a broad concept, there are currently four types of it: 1) machine-assisted human translation (MAHT) - the use of computer electronic dictionaries and instructions included in it when translating a text by a human; 2) computer-assisted translation (CAT) - serves to translate a text using a computer and separate the main concepts understood in it; 3) human-assisted machine translation (HAMT) - assigning the task of editing to a translation performed by computer software by a human resource; 4) fully automatic machine translation (FAMT) - performing both the translation and editing tasks of the text through a translator program. According to sources, an experiment on machine translation was first conducted in Turkology under the leadership of Dorster (1954). Since those years, two different directions of developing machine translation algorithms have been initiated in the study of the specific features of Turkic languages:

1. Machine translation between related Turkic languages.

2. Machine translation between languages belonging to other language families other than Turkic languages.

In this regard, these studies in Turkology have also given rise to various approaches and translation technologies. Although much research has been conducted on machine translation, most of the research in Turkology dates back to the last decade of our century. In particular, we find the research conducted by Güneyd Tantuğ, Ashref Adalı, and Kemal Oflazer on machine translation from Turkmen to Turkish. This machine translation technology is a rule-based system based on statistical language models and multi-meaning lexical and morphological transfer. Almost all MTs between sister languages have implemented similar modules: a morphological analyzer, a part-of-speech tagger, a bilingual transfer dictionary, and a morphological generator. In Czech->Latvian, syntactic structures are almost not used, since the sequence of words in pairs is almost the same. Translation from Turkmen to Turkish is based on direct morphological transfer. It is noted that the machine translation system consists of the following main blocks.

As a result of research on automatic translation, such types of machine translation as Rule-based MT, Sample-based MT, Statistical-based MT, and Neural MT have emerged.



Two factors influenced the development of machine translation:

1) the computer modeling of translation;

2) the breadth of machine translation's potential for ensuring interlingual communication for social purposes.

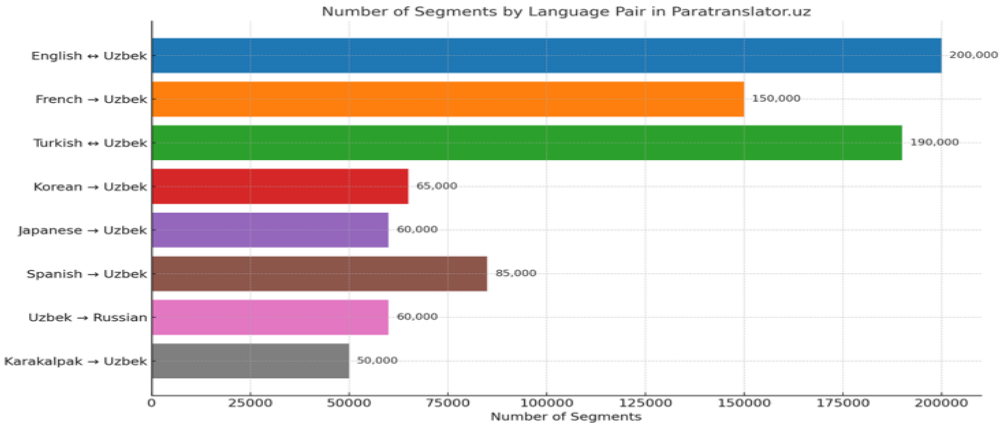
METHODOLOGY

In Uzbekistan, scientists such as A. Toshpulatov, M. Khakimov, M. Aripov, S. Muhamedov, and N. Abdurakhmonova conducted research on machine translation and language models. In particular, N. Abdurakhmonova conducted research aimed at creating linguistic support for a program that translates a foreign text into Uzbek and vice versa, and the scientist published several scientific works on corpus linguistics, morphological analysis, linguistic support for machine translation, and corpus-based language teaching. The monograph “Linguistic Support for Machine Translation” discusses the stages of modeling for automatic translation, formal language theory, grammatical coordination of English and Uzbek, the theory of algorithm development for software, and important aspects of building a linguistic base. In recent years, a number of scientific works have been carried

out in Uzbek linguistics within the framework of computational linguistics, including machine translation. It is appropriate to cite the dissertation for the degree of Doctor of Philosophy (PhD) by Nilufar Abdurakhmonova on the topic “Linguistic support of the program for translating English texts into Uzbek (on the example of simple sentences)” (2018), the dissertation for the degree of Doctor of Philosophy (PhD) by Shakhnoza Abidova on the topic “Models and algorithms of electronic translation for Turkic languages (on the example of translation from Uzbek to Karakalpak)” (2021), and the dissertation for the degree of Doctor of Philosophy (PhD) by Anora Akhmedova on the topic “Study of lexical-semantic relations of analogies in a parallel corpus (based on English-Uzbek and Uzbek-English parallel texts)” (2022) as research conducted in this area.

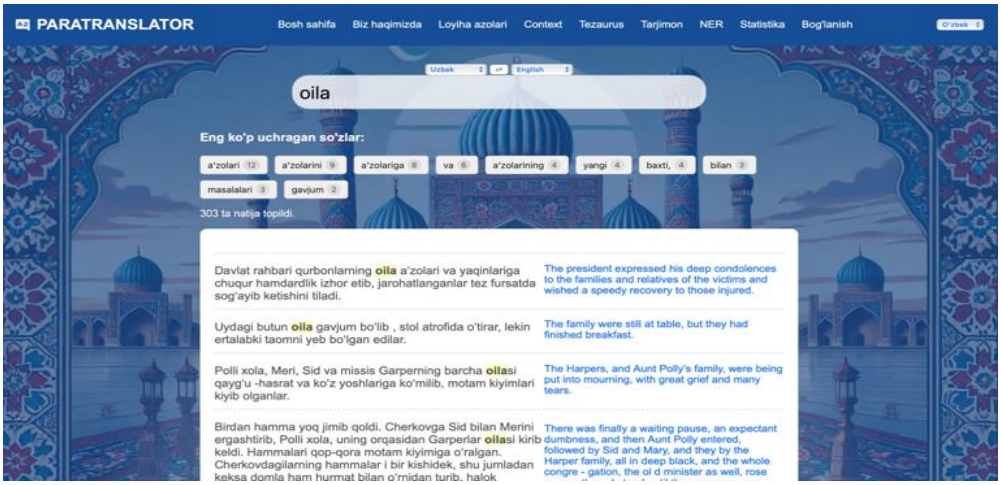
The Paratranslator platform was developed as part of a research project that aimed to put contextual translation technology based on a parallel corpus into practice.

The main linguistic achievement of the platform is the possibility of implementing “multilingual translation” based on “contextual parallel corpora”. The platform is based on parallel corpora in Uzbek, English, Russian, Turkish, French, Karakalpak, Japanese, Korean, and Spanish of scientific, official, artistic, and journalistic styles. There are a total of 960,000 parallel segmented texts in the following language pairs:

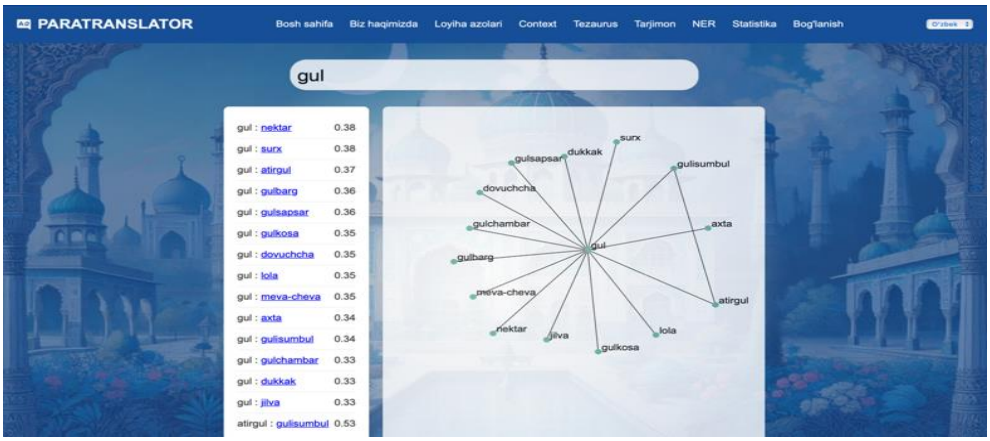


Practical possibilities: usage, interface, and examples

These corpora allow for the identification of “contextual alternatives” to the units being translated. For example, when a user enters a word or phrase on the platform, parallel contexts related to it are displayed, as well as statistical information on how they are used in the language (word frequency, collocations, synonyms, etc.).



The Thesaurus section of the platform graphically displays words that are related based on semantic similarity for each word. For example, when we search for the word “gul” (flower), the following semantic network is generated:



This figure depicts a contextual-semantic network automatically generated for the lexeme “gul” (flower) on the Paratranslator.uz platform, and the words with the closest contextual connections are visually presented in a graphical form. In the graphic image, words related to the term “gul” (flower) (atigul-rose, gulsumbul-hyacinth, etc.) are arranged according to the degree of semantic proximity. For example, the word “gul” (flower) is in the center of the network, and words such as “lola”-tulip, “gulchambar”-garland, “gulbarg”-petal, “nektar”-nektar are connected to it with varying degrees of semantic proximity. Each connection coefficient (in the range of 0.33–0.38) is based on the frequency of use in the corpus and contextual proximity. This visual and statistical approach provides translators, lexicographers, and linguists with important opportunities to identify connections between words in real-world contexts, analyze synonyms, and analyze associations.

Through the interface, users can analyze the context, determine word frequency, and choose the most appropriate translation option. This is a practical tool for translators, philologists, researchers, and students.

CONCLUSION

Context-based translation technology has become an urgent need in today's globalized language environment. The PATRANSLATOR platform is an advanced product that can respond to this need on a scientific and technical basis, and it is organized as a means of semantic transfer in a contextual sense, not just a literal reflection of language changes.

Scientifically based parallel corpora, NLP tools, the platform's multilingual potential, functional capabilities in the interface, and research aimed at strengthening the linguistic resources of the Uzbek language have led to its high appreciation on a regional and global scale. On this basis, the PARATENTRANSLATOR platform can be recognized as an advanced model offering practical and theoretical solutions in the fields of linguistic technologies, translation studies, computational linguistics, and education.

REFERENCES

1. Abdurakhmonova N. Linguistic support of machine translation. (Monography) – Tashkent, 2018.
2. Abdurakhmonova N. Computer linguistics. – Tashkent: Nodirabegim, 2021. –B.225.
3. Rahimov A. Fundamentals of computer linguistics. –T., 2011. – B.143.
4. Philipp Koehn. A Beam Search Decoder for Phrase-Based Statistical Machine Translation Models. Machine Translation: From Real Users to Research 6th Conference of the Association for Machine Translation in the Americas, AMTA 2004. Washington, DC, USA, September/October 2004. – P. 65
5. https://paratranslator.uz/en/